

ТВ
ТЕОРИЯ ВЕРОЯТНОСТЕЙ
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
МС

М.В.БОЛДИН, Г.И.СИМОНОВА
Ю.Н.ТЮРИН

ЗНАКОВЫЙ
СТАТИСТИЧЕСКИЙ
АНАЛИЗ
ЛИНЕЙНЫХ МОДЕЛЕЙ

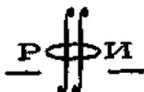


МОСКВА
НАУКА • ФИЗМАТЛИТ
1997

ББК 22.172

Б79

УДК 519.23 + 519.24



Издание осуществлено при поддержке
Российского фонда фундаментальных
исследований по проекту 96-01-14047

Болдин М. В., Симонова Г. И., Тюрин Ю. Н. Знаковый статистический анализ линейных моделей.— М.: Наука. Физматлит, 1997.— 288 с. — (Теория вероятностей и математическая статистика.) — ISBN 5-02-015222-6.

В монографии изложен новый непараметрический подход к анализу статистических данных, когда закон распределения наблюдений неизвестен и выводы основываются не на самих данных, а на знаках определенных функций от них. Рассмотрены важные для приложений статистические модели регрессии и авторегрессии, для которых единым знаковым методом решены основные статистические задачи. Свойства знаковых правил изучены для конечных и растущих объемов выборок, показана их высокая устойчивость к грубым ошибкам. Предложены численные алгоритмы знакового анализа.

Для специалистов, аспирантов, студентов, изучающих и использующих методы математической статистики.

Табл. 23. Ил. 35. Библиогр. 111 назв.

Научное издание

БОЛДИН Михаил Васильевич
СИМОНОВА Галина Ивановна
ТЮРИН Юрий Николаевич

ЗНАКОВЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ ЛИНЕЙНЫХ МОДЕЛЕЙ

Серия «Теория вероятностей и математическая статистика»

Редактор *Е.Ю.Ходак*
Компьютерный набор авторов

ИБ № 41842

ЛР № 020297 от 27.11.91. Подписано в печать 15.08.97. Формат 60×90/16.
Бумага офсетная № 1. Печать офсетная. Усл. печ. л. 18. Уч.-изд. л. 19,8.
Тираж 1000 экз. Заказ тип. № 2182 . С-024.

Издательская фирма «Физико-математическая литература» РАН
117071 Москва В-71, Левянский проспект, 15

Отпечатано в Московской типографии № 2 РАН
121099 Москва Г-99, Шубинский пер., 6

Б 1802090000—024 73-97. Наука, II полугодие
053(02)-97

ISBN 5-02-015222-6

© М.В.Болдин, Г.И.Симонова,
Ю.Н.Тюрин, 1997

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	6
ВВЕДЕНИЕ	9

ЧАСТЬ I

ЛИНЕЙНЫЕ МОДЕЛИ НЕЗАВИСИМЫХ НАБЛЮДЕНИЙ

Глава 1. Знаковый анализ однопараметрической регрессии	21
§ 1.1. Закон Хаббла: история и современность	21
§ 1.2. Определение постоянной Хаббла знаковым методом	25
§ 1.3. Асимптотические результаты	36
1.3.1. Состоятельность (36). 1.3.2. Асимптотическая нормальность (39).	
§ 1.4. Функция влияния	45
Глава 2. Знаковые критерии	48
§ 2.1. Общая линейная модель	48
§ 2.2. Локально оптимальные знаковые критерии в задаче регрессии	51
§ 2.3. Вычисление критических значений. Асимптотическая теория	59
§ 2.4. Пример. Двухфакторные таблицы	62
§ 2.5. Вычисление критических значений. Конечные объемы выборок	65
Глава 3. Знаковые оценки	68
§ 3.1. Знаковые оценки и их вычисление	68
§ 3.2. Знаковое оценивание. Асимптотическая теория	80
3.2.1. О роли асимптотической теории (80). 3.2.2. Состоятельность знаковых оценок (81). 3.2.3. Асимптотическая нормальность знаковых оценок (86). 3.2.4. Асимптотическая ковариация знаковых оценок (90). 3.2.5. Равномерный закон больших чисел (90). 3.2.6. Теорема о равномерной линейности (94). 3.2.7. Асимптотическая мощность знаковых критериев (100). 3.2.8. Функция чувствительности (101).	
§ 3.3. Сравнение оценок	102
3.3.1. Как сравнивают оценки (102). 3.3.2. Ранговое оценивание (104). 3.3.3. Оценки наименьших квадратов и наименьших модулей (107). 3.3.4. Асимптотическая эффективность знаковых оценок (109).	
Глава 4. Проверка линейных гипотез	111
§ 4.1. Знаковые критерии для линейных гипотез	111
§ 4.2. Асимптотические свойства знаковых критериев для линейных гипотез	113

§ 4.3.	Примеры	117
§ 4.4.	Проверка линейных гипотез в однофакторных и двухфакторных таблицах	122
	4.4.1. Однофакторный анализ (t выборки, которые могут отличаться сдвигом) (122). 4.4.2. Двухфакторный анализ (125).	
§ 4.5.	Вычисление критических значений в задачах проверки линейных гипотез	127
	4.5.1. Однофакторный анализ (127). 4.5.2. Двухфакторный анализ (132).	

ЧАСТЬ II

АВТОРЕГРЕССИОННЫЕ МОДЕЛИ ВРЕМЕННЫХ РЯДОВ

Глава 5.	Процедуры наименьших квадратов и наименьших модулей в простейшей авторегрессионной модели	133
§ 5.1.	Введение	133
§ 5.2.	Простейшее стационарное уравнение авторегрессии и его решения	135
§ 5.3.	Процедуры наименьших квадратов	137
	5.3.1. Оценка наименьших квадратов (138). 5.3.2. Тесты наименьших квадратов (141).	
§ 5.4.	Оценка наименьших квадратов в нестационарной авторегрессии	145
§ 5.5.	Процедуры наименьших модулей	149
	5.5.1. Оценка наименьших модулей (149). 5.5.2. Тесты наименьших модулей (153). 5.5.3. Взвешенные оценки наименьших модулей (157).	
§ 5.6.	Функционалы влияния оценок наименьших квадратов и наименьших модулей	159
	5.6.1. Функционал влияния оценки наименьших квадратов (161). 5.6.2. Функционал влияния оценки наименьших модулей (163). 5.6.3. Функционал влияния взвешенной оценки наименьших модулей (164).	
§ 5.7.	О проверке стационарности авторегрессионного уравнения	165
§ 5.8.	Приложение. Доказательства теорем	170
Глава 6.	Знаковый анализ однопараметрической авторегрессии	175
§ 6.1.	Введение в знаковый авторегрессионный анализ	175
§ 6.2.	Знаковые тесты	179
§ 6.3.	Знаковые тесты в нестационарной авторегрессии	185
§ 6.4.	Теорема о равномерном стохастическом разложении. Мощность знаковых тестов при близких альтернативах	189
§ 6.5.	Сравнение знаковых тестов с другими непараметрическими тестами	192
§ 6.6.	Знаковые оценки параметров	197
	6.6.1. Знаковая оценка $\hat{\beta}_{n,S}$ (197). 6.6.2. Знаковая оценка $\beta_{n,S}^*$ (201). 6.6.3. Знаковая оценка $\tilde{\beta}_{n,S}^*$ (202).	
§ 6.7.	Функционалы влияния знаковых оценок	204
	6.7.1. Функционал влияния знаковой оценки $\hat{\beta}_{n,S}$ (204). 6.7.2. Функционал влияния знаковой оценки $\tilde{\beta}_{n,S}$ (208). 6.7.3. Функционал влияния знаковой оценки $\beta_{n,S}^*$ (208).	
§ 6.8.	Результаты моделирования: оценивание квантилей, доверительное оценивание, засоренные выборки	210

	6.8.1. Оценивание квантилей (210). 6.8.2. Доверительное оценивание параметра β (215). 6.8.3. Знаковое оценивание по засоренным выборкам (220).	
	§ 6.9. Приложение. Доказательство теоремы 6.4.1	222
Г л а в а	7. Знаковый анализ многопараметрической авторегрессии	235
§	7.1. Введение	235
§	7.2. Тестовые статистики и их распределения при гипотезе	238
§	7.3. Теорема о равномерном стохастическом разложении: мощность знаковых тестов при близких альтернативах	245
§	7.4. Проверка линейных гипотез	250
§	7.5. Знаковые оценки параметров	254
	7.5.1. Знаковая оценка $\hat{\beta}_{n,S}$ (254). 7.5.2. Знаковая оценка $\tilde{\beta}_{n,S}$ (256). 7.5.3. Знаковая оценка $\beta_{n,S}^*$ (257).	
§	7.6. Функционалы влияния оценок в многопараметрической авторегрессии	260
	7.6.1. Функционал влияния оценки наименьших квадратов $\hat{\beta}_{n,LS}$ (260). 7.6.2. Функционал влияния оценки наименьших модулей $\tilde{\beta}_{n,LD}$ (262). 7.6.3. Функционал влияния взвешенных оценок наименьших модулей $\tilde{\beta}_{n,LDW}$ (264). 7.6.4. Функционал влияния знаковой оценки $\tilde{\beta}_{n,S}$ (265). 7.6.5. Функционал влияния знаковой оценки $\beta_{n,S}^*$ (265).	
§	7.7. Эмпирическая функция распределения и эмпирические процессы, построенные по остаткам	267
§	7.8. Приложение. Доказательство теоремы 7.7.1	274
СПИСОК ЛИТЕРАТУРЫ	281

ПРЕДИСЛОВИЕ

Для непараметрической статистики последние полвека были временем зарождения ранговых методов, их бурного развития, достижения ими зрелости и обретения широкой известности. Ранговый подход в статистике состоит в замене элементов выборки их рангами и в дальнейшем анализе этих рангов, а не самих исходных наблюдений. Для статистических моделей, более сложных, чем выборка, рангами заменяют видимые остатки. Лет десять назад мы начали изучать возможности, которые открывает для непараметрического анализа аналогичный переход от наблюдений или остатков к их знакам. Некоторые знаковые правила были известны издавна. Вспомним хотя бы критерий знаков. Мы развивали этот подход в наших теоретических статьях и практических исследованиях. Когда же все мы (авторы этой книги) собрались в Московском университете и когда наши исследования получили поддержку Российского фонда фундаментальных исследований, мы решили изложить эту тему систематически. Так появилась эта книга. Знаковый метод исходит из предположения, что случайные ошибки могут быть положительными или отрицательными с равными вероятностями. При этом условии знаковые выводы оказываются свободными от распределения. Как дополнительное достоинство следует отметить, что эти выводы проявляют высокую устойчивость при отступлении от модельных предположений. Например, по отношению даже к значительной доле грубых ошибок. Может показаться удивительным, но относительная асимптотическая эффективность знаковых правил оказывается довольно высокой, несмотря на явную потерю информации, которая происходит при замене чисел их знаками.

Наибольшее внимание мы уделили малым выборкам. При этом мы стремились к точным (не асимптотическим) результатам: точным уровням значимости при проверке гипотез, точным доверительным вероятностям при доверительном оценивании. Отметим, что необходимые вычислительные процедуры оказались достаточно простыми и быстрыми, чтобы их можно было провести на рядовых компьютерах. Мы даже разработали для этих целей пакет прикладных программ. С его помощью и сделаны вычисления в тех примерах, которые рассмотрены в книге. Мы исследовали также асимптотические свойства знаковых правил, когда число наблюдений неограниченно растет. Асимптотическое исследование знаковых процедур сводится к исследованию случайных процессов (и полей) типа эмпирических и взвешенных эмпирических, но построенных по видимым остаткам.

Технически эти параграфы в книге наиболее сложные. Знаковые правила, представленные в книге, образуют непараметрическую схему обработки данных, с помощью которой можно решать основные статистические задачи: проверку гипотез (в том числе линейных), вычисление точечных оценок и построение доверительных множеств для неизвестных параметров. По своим возможностям и кругу доступных им задач знаковые методы оказываются вполне сопоставимы с ранговыми.

Мы развиваем знаковые методы применительно к линейным моделям. В первой части книги это регрессионные и факторные модели, т.е. статистические модели независимых наблюдений. Во второй части мы рассматриваем линейные модели для временных рядов, в основном для модели авторегрессии. Во введении мы на примерах показываем возможности знаковых методов и разъясняем основы знакового непараметрического анализа.

Мы старались сделать изложение доступным как можно более широкому кругу читателей, которые интересуются не только собственно статистической теорией и ее непараметрическим аспектом, но и приложениями этой теории. Сказанное в особенности относится к первой части. Здесь мы обсуждаем и численные методы для построения знаковых процедур. Для чтения книги достаточно знакомства с базовыми курсами теории вероятностей и математической статистики. Требуется также некоторый навык чтения математической литературы; это больше относится ко второй части. Желательно знакомство с классическим гауссовским регрессионным и дисперсионным анализом.

Вклад авторов в разные главы книги был различен. Первая часть в основном написана Ю.Н.Тюриным, вторая — в основном М.В.Болдиным. Г.И.Симоновой выполнены все расчеты и написаны вычислительные параграфы в обеих частях книги.

Следующие строки для нас самые приятные во всей книге. В них мы выражаем признательность всем, кто так или иначе способствовал нашим исследованиям и помог нам изложить их здесь. К сожалению, не всех мы можем назвать поименно. Но это не значит, что мы неблагодарно забыли остальных.

Мы благодарны Московскому государственному университету и его механико-математическому факультету, из которого все мы вышли и где сейчас продолжаем работать, за стимулирующую к исследованиям обстановку, за поощрение исследований и за их поддержку.

Мы благодарны чтим память академика Б.В.Гнеденко, при заботливом внимании которого были проведены эти исследования.

Мы благодарны профессору Н.М.Сотскому за обсуждения теоретических и прикладных вопросов статистического анализа временных рядов.

Мы благодарны профессору А.С.Шарову за консультации по астрономическим проблемам, затронутым в главе 1.

Мы благодарим профессоров А.П.Коростелева, А.А.Макарова, Я.Ю.Никитина и Э.В.Хмаладзе за внимание к работе и полезные обсуждения.

Мы благодарны профессору Д.М. Чибисову за перевод этой книги на английский язык. В процессе этой работы он сделал много ценных замечаний, способствовавших прояснению и уточнению текста книги, за что мы также ему глубоко признательны.

Мы благодарим наших студентов, слушателей наших лекций и участников наших семинаров, в сотрудничестве и в общении с которыми мы проводили наши исследования.

Мы благодарны Российскому фонду фундаментальных исследований, который своим грантом поддержал наши исследования и издание этой книги.

Мы благодарим наших близких за всемерную поддержку и безграничное терпение.

Москва,
июль 1997 г.

*М.В. Болдин
Г.И. Симонова
Ю.Н. Тюрин*

ВВЕДЕНИЕ

Книга посвящена знаковому непараметрическому подходу к статистическому анализу данных. Мы излагаем этот метод применительно к линейным моделям независимых наблюдений и к авторегрессионным моделям временных рядов.

Непараметрические методы, и знаковый в их числе, расширяют область приложения статистических методов по сравнению с классическими параметрическими методами, в которых считается, что распределение случайных ошибок следует определенному закону, известному с точностью до параметров. Среди прикладных методов статистики метод наименьших квадратов, несомненно, является наиболее важным. Этот метод позволяет получить глубокие статистические результаты при предположении, что случайные ошибки распределены по нормальному (гауссовскому) закону. Это модельное предположение отражает в общих чертах поведение реальных статистических наблюдений во многих задачах. Замечательное свойство гауссовской модели состоит в том, что на ее основе создана целостная система статистической обработки, включающая в себя точечное и интервальное оценивание параметров, проверку статистических гипотез и т.д. Именно это (вместе с простотой вычислений) обеспечило гауссовским методам стойкую и долгую популярность в прикладных исследованиях.

Но теоретические богатства гауссовской модели не должны заслонять от нас тот факт, что положенные в ее основу математические предположения на практике могут не выполняться. В этом случае выводы, полученные средствами гауссовской статистики, рискуют оказаться ошибочными. К тому же не существует методов, позволяющих надежно проверить, следуют ли случайные ошибки нормальному закону. Более того, часто вообще не удается указать какого-либо параметрического семейства, в рамках которого находится конкретное распределение ошибок. Поэтому на практике нужна статистическая теория, ориентированная не на какое-либо параметрическое семейство, а применимая к более широкому классу статистических ошибок, т.е. непараметрическая. При этом мы хотели бы, чтобы по кругу решаемых задач и по своим возможностям предлагаемый метод не уступал гауссовскому, т.е. тоже предоставлял естествоиспытателю целостную систему статистической обработки, включая упомянутые точечное и доверительное оценивания и проверку гипотез.

Один из таких непараметрических методов — ранговый, т.е. основанный на рангах наблюдений — был разработан в последние

десятилетия и сейчас широко известен. Он основан на предположении, что случайные ошибки независимы и распределены непрерывно и одинаково.

В этой книге мы развиваем другой непараметрический метод, в основе которого лежат предположения, что независимые случайные ошибки с равными вероятностями принимают положительные и отрицательные значения. При этом часто они даже не обязаны быть одинаково распределенными. Так как в этом методе используются не сами наблюдения, а их знаки (или знаки остатков), естественно называть этот метод знаковым.

Отдельные знаковые правила были известны в статистике издавна. Общеизвестным примером может служить критерий знаков для проверки гипотез о медиане независимых одинаково распределенных наблюдений. Наши знаковые процедуры применимы к более сложным моделям данных, в частности, к регрессионным схемам (в том числе факторным) с независимыми ошибками и к линейным моделям временных рядов типа авторегрессии, скользящего среднего и их модификациям. Для всех таких (достаточно разнообразных) моделей знаковые процедуры сохраняют несколько характерных особенностей, из которых здесь выделим лишь несколько.

Важнейшим свойством нам представляется то, что при упомянутых выше предположениях знаковые статистики *свободны* от распределения исходных данных. Это позволяет не только строить тесты с фиксированным уровнем значимости для малых выборок, но и получать точные доверительные множества для неизвестных параметров. Такая возможность — редкое и ценное свойство даже в схемах независимых наблюдений, не говоря о более сложных моделях временных рядов.

Другое привлекательное свойство знаковых процедур — их асимптотическая нормальность (в случае, когда объемы выборок стремятся к бесконечности) при минимальных ограничениях на распределение данных. Это полезное и важное свойство. Оно, к слову, выполняется вовсе не для всех широко используемых на практике процедур. В частности, для схемы авторегрессии общеизвестные и обычно используемые оценки наименьших квадратов и наименьших модулей являются \sqrt{n} -асимптотически нормальными лишь при конечной дисперсии наблюдений; если же эта дисперсия бесконечна, то оценки могут сходиться к истинным значениям быстрее, чем с обычной скоростью $n^{-1/2}$, но, к сожалению, их асимптотические распределения неизвестны. Разумеется, переход от самих наблюдений к их знакам сопровождается потерей информации. Замечательно то, что асимптотическая эффективность знаковых правил по сравнению даже с оптимальными не слишком мала. В гауссовской регрессии эта асимптотическая эффективность равна $2/\pi$ (как в схеме повторной выборки), а в гауссовской авторегрессии она равна $(2/\pi)^2$. Вообще, если случайные ошибки имеют симметричную плотность, достигающую в нуле своего максимального значения, то асимптотическая эффективность знаковых оценок в линейной модели относительно оценок наименьших квадратов не опускается ниже $1/3$, при тех условиях, когда знаковые оценки и оценки наименьших квадратов асимптотически

нормальны. Для данных с «тяжелыми хвостами» асимптотическая эффективность знаковых процедур относительно процедур наименьших квадратов может быть и больше единицы и вообще сколь угодно большой.

Знаковые оценки устойчивы к выбросам в данных, в отличие от оценок наименьших квадратов (а в схеме авторегрессии неустойчивы к выбросам и оценки наименьших модулей).

Знаковые процедуры можно использовать и в случае разнораспределенных ошибок в схемах линейной регрессии и авторегрессии.

Наконец, знаковые процедуры составляют полную систему статистической обработки данных, сравнимую по кругу решаемых задач с процедурами наименьших квадратов для гауссовских ошибок.

Разумеется, существуют и иные способы обработки данных (например, ранговые), покрывающие весь спектр решаемых задач и обладающие некоторыми из перечисленных свойств. Привлекательность знаковых процедур в том, что для них эти свойства выполняются все вместе. На нескольких конкретных примерах мы покажем сейчас, как действуют знаковые методы. Первые два примера относятся к моделям независимых наблюдений, третий посвящен временным рядам. Мы начинаем с иллюстрации устойчивости, хотя это не самое главное и вполне ожидаемое свойство знаковых процедур. Зато оно очень наглядно.

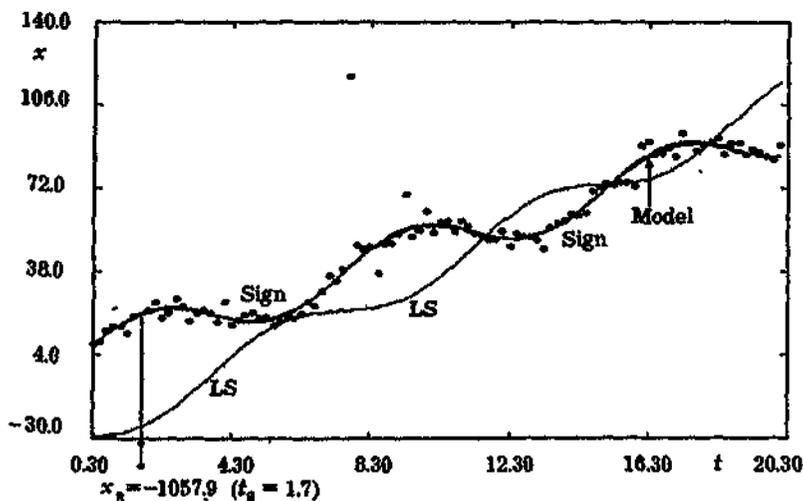


Рис. В.1. Графики функции, оцененной знаковым методом (Sign) и методом наименьших квадратов (LS)

Пример 1. На рис. В.1 показаны искусственные данные, подчиняющиеся регрессионной модели $x_i = \theta_1 + \theta_2 t_i + \theta_3 \sin(\omega t_i) + \xi_i$, $i = 1, \dots, n$. Они имитируют наблюдения над зависимостью $x = \theta_1 + \theta_2 t + \theta_3 \sin(\omega t)$ в моменты t_1, t_2, \dots, t_n . Эти наблюдения сопровождаются случайными ошибками ξ_i , $i = 1, \dots, n$, которые мы

считаем независимыми. Относительно распределения этих ошибок мы принимаем

Условие В.1. $P\{\xi_i < 0\} = P\{\xi_i > 0\} = 0.5$.

В рассматриваемом примере мы взяли следующую функцию распределения ошибок ξ_i , $i = 1, \dots, n$:

$$F(x) = (1 - \varepsilon)\Phi(x/\sigma) + \varepsilon K(x/\sigma),$$

где $\Phi(x)$ — функция распределения стандартного гауссовского распределения, $K(x)$ — функция распределения Коши,

$$K(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\sqrt{\frac{\pi}{2}} x\right), \quad \sigma = 3, \quad \varepsilon = 0.2.$$

Точки на рис. В.1 изображают пары чисел (t_i, x_i) при некоторых определенных значениях параметров $\theta_1, \theta_2, \theta_3$ (из табл. В.1) при числе наблюдений $n = 101$ и $\omega = 4\pi/15$. На рис. В.1 не изображен единичный выброс $x_8 = -1057.9$ при $t_8 = 1.7$.

Предположим, что параметры $\theta_1, \theta_2, \theta_3$ нам не известны (значение ω задано), и попытаемся восстановить эту зависимость по наблюдениям (t_i, x_i) , $i = 1, \dots, n$. Кривая LS — результат восстановления зависимости по методу наименьших квадратов. Несмотря на то, что основная масса наблюдений близко следует истинной зависимости, восстановленный результат имеет мало общего с действительностью. Так проявилась в этом примере хорошо известная неробастность метода наименьших квадратов. В массе ошибки невелики и ведут себя (приблизительно) как совокупность независимых одинаково распределенных гауссовских случайных величин. Однако примерно 20% наблюдений содержат гораздо большие ошибки и потому выпадают из общей массы. Такие наблюдения часто рассматривают как грубые ошибки, засоряющие основной массив данных. Именно это засорение и оказывает в данном случае решающее влияние на конечный результат, полученный методом наименьших квадратов.

Кривая Sign получена иным методом — знаковым. Она почти не отличается от графика истинной зависимости x от t , отмеченной на рисунке меткой Model. В табл. В.1 приведены оценки параметров, полученные двумя методами — наименьших квадратов и знаковым.

Т а б л и ц а В.1

Параметр	Model	LS	Sign
θ_1	3.50	-29.77	3.54
θ_2	4.50	6.87	4.47
θ_3	10.00	-7.38	9.32

Из рис. В.1 и табл. В.1 видно, что знаковый метод дал хорошие результаты, несмотря на засорение. Практика (и теория) показывает, что знаковые методы вообще обладают высокой устойчивостью к засорению.

П р и м е р 2. Пример относится к анализу факторов. Мы рассматриваем таблицу с двумя входами, где каждой комбинации

уровней факторов (i, j) соответствует несколько значений отклика x_{ijk} , $k = 1, \dots, m$ (одинаковое число наблюдений в каждой клетке). Таблица содержит факстуственные данные, порожденные аддитивной моделью действия факторов, когда

$$x_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}.$$

Здесь $\mu, \alpha_1, \alpha_2, \dots, \beta_1, \beta_2, \dots$ — соответственно генеральное среднее и эффекты факторов; величины e_{ijk} рассматриваются как ошибки, из-за которых отклики x_{ijk} отклоняются от ожидаемых значений $\mu + \alpha_i + \beta_j$. Совокупность ошибок в этой таблице e_{ijk} имеет те же свойства, что и в примере 1: к однородной выборке из гауссовского распределения примешано немного резко отличающихся значений. В этом примере модельная функция распределения ошибок имеет вид

$$F(x) = (1 - \varepsilon)\Phi(x/\sigma_1) + \varepsilon\Phi(x/\sigma_2),$$

где $\sigma_1 = 0.5, \sigma_2 = 20, \varepsilon = 0.15$. Количество строк таблицы равно 5, столбцов — 7. Число наблюдений в каждой клетке равно 12.

Рассмотрим задачу оценивания неизвестных μ, α_i, β_j по наблюдениям x_{ijk} . Вновь применим методы наименьших квадратов и знаковый. К сожалению, факторный анализ нельзя представить столь же наглядно, как регрессионный, и поэтому нельзя наглядно судить о качестве полученных тем или иным методом результатах (т.е. об оценках). Поэтому прибегнем к частотному анализу остатков, т.е. величин $x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$. Здесь $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j$ обозначают оценки соответствующих параметров. Если оценки близки к истинным значениям, остатки ведут себя примерно как (ненаблюдаемые) ошибки e_{ijk} . По их гистограмме в некоторой степени можно судить о качестве оценок и тем самым о методе оценивания.

Гистограммы на рис. В.2 и В.3 построены по остаткам $x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$. На рис. В.3 представлены результаты, полученные методом наименьших квадратов; на рис. В.2 — знаковым методом. На обоих рисунках показаны лишь остатки, попавшие в отрезок $[-3, 3]$. Здесь же изображена плотность гауссовского распределения с $\sigma_1 = 0.5$, которому подчиняется примерно 85% ошибок модели.

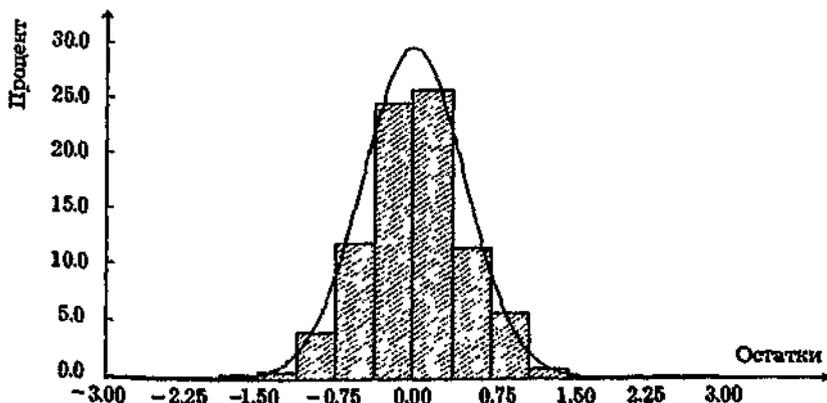


Рис. В.2. Гистограмма остатков в знаковом методе

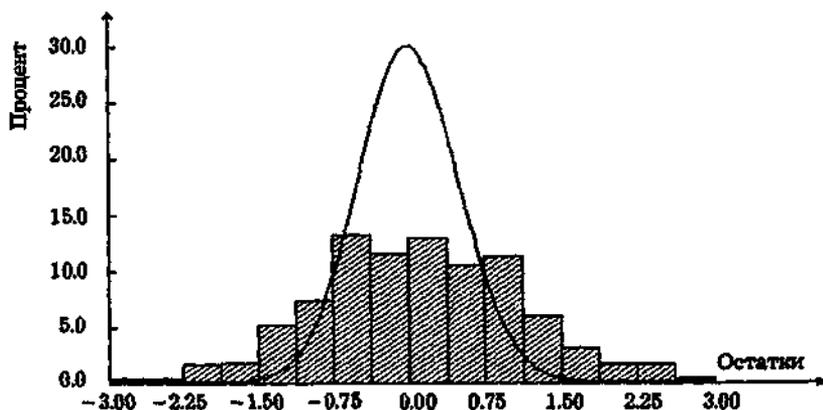


Рис. В.3. Гистограмма остатков в методе наименьших квадратов

По гистограммам видно, что разброс остатков в знаковом методе оказался меньше. Это означает, что знаковые оценки близки к истинным значениям параметров и мало подвержены влиянию грубых ошибок. В табл. В.2 приведены результаты оценивания параметров модели.

Таблица В.2

Параметр	Model	LS	Sign
α_1	0.50	0.99	0.56
α_2	2.00	2.39	2.08
α_3	7.80	7.48	7.66
α_4	-5.30	-4.68	-5.30
α_5	-5.00	-6.19	-5.00
β_1	-6.40	-5.62	-6.36
β_2	7.30	7.32	7.26
β_3	13.50	13.94	13.39
β_4	-7.00	-7.82	-6.87
β_5	5.40	5.47	5.42
β_6	-2.60	-3.27	-2.52
β_7	-10.20	-10.02	-10.31
μ	1.00	0.86	0.96

Пример 3. Пример иллюстрирует устойчивость знаковой оценки в линейной стационарной модели авторегрессии. Модели авторегрессии, модели скользящего среднего и более общие модели авторегрессии-скользящего среднего являются наиболее популярными в приложениях моделями временных рядов. Итак, пусть наблюдаются величины y_0, \dots, y_n , где

$$y_i = u_i + z_i^2 \xi_i, \quad i \in \mathbb{Z} = \{\dots, -1, 0, 1, \dots\};$$

величины $\{u_i\}$ удовлетворяют уравнению авторегрессии первого порядка

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i \in \mathbb{Z}, \quad |\beta| < 1,$$

где $\{\varepsilon_i\}$ суть независимые одинаково распределенные случайные ошибки. Величины $\{z_i^\gamma\}$ — независимые одинаково распределенные случайные величины, принимающие значения 1 и 0 с вероятностями γ и $1 - \gamma$ соответственно; параметр γ , $0 \leq \gamma \leq 1$, характеризует уровень засорения; $\{\xi_i\}$ — последовательность независимых одинаково распределенных величин, моделирующих грубые ошибки, засоряющие основной массив наблюдений; последовательности $\{u_i\}$, $\{z_i^\gamma\}$, $\{\xi_i\}$ предполагаются независимыми между собой.

На рис. В.4 мы приводим модельные данные y_0, y_1, \dots, y_n для $n = 400$; $\beta = 0.2$. Здесь $\{\varepsilon_i\}$ — последовательность стандартных гауссовских величин; $\{\xi_i\}$ — последовательность гауссовских величин с нулевым средним и среднеквадратичным отклонением $\sigma = 10$; $\gamma = 0.05$.

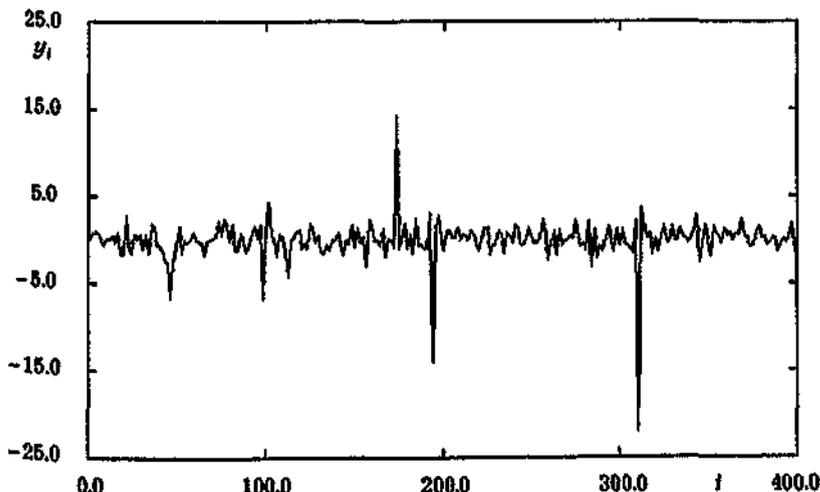


Рис. В.4. Наблюдения, подчиняющиеся модели $y_i = u_i + z_i^\gamma \xi_i$, $i = 1, \dots, 400$, $\beta = 0.2$; $\varepsilon_i \sim N(0,1)$; $\gamma = 0.05$; $\xi_i \sim N(0,100)$

Простейшей и наиболее распространенной оценкой параметров в моделях авторегрессии (как и в модели линейной регрессии с независимыми ошибками) является оценка наименьших квадратов. Оказывается, как и в схеме линейной регрессии, оценка наименьших квадратов очень чувствительна к выбросам. В то же время наша знаковая оценка является устойчивой. Мы иллюстрируем сказанное расчетами по модельным данным, описанным выше. В табл. В.3 приведены результаты расчетов для трех реализаций.

Таблица В.3

$N = 400$, $\beta = 0.2$, $\varepsilon_i \sim N(0,1)$, $\gamma = 0.05$, $\xi_i \sim N(0,100)$

Номер реализации	Оценка наименьших квадратов	Знаковая оценка
1	0.04	0.17
2	0.02	0.23
3	0.03	0.21

Устойчивость знаковых оценок — наглядное, но, как мы уже говорили, не главное привлекательное свойство знаковых процедур. К тому же, в настоящее время предложены и изучены многие другие методы устойчивого оценивания (см., например, [42, 81; 36, 73]). Главным достоинством знакового метода мы считаем то, что с его помощью можно делать точные (не асимптотические) статистические выводы (построение статистических тестов и доверительных множеств) для малых выборок. При этом наши выводы оказываются возможным сделать независимо от конкретного вида распределения данных, и в этом смысле предлагаемые знаковые процедуры являются непараметрическими.

Проследим основные этапы знакового анализа на примере моделей, которые мы рассматриваем в этой книге. В них наблюдаемый статистический материал, скажем, X имеет структуру

$$X = f(\theta, \xi),$$

где $f(\cdot, \cdot)$ — известная функция, ξ — совокупность случайных ошибок, θ — неизвестный (конечномерный) параметр. Для регрессионных моделей (и для других линейных моделей для независимых наблюдений) формула упрощается. Здесь

$$X = f(\theta) + \xi. \quad (B.1)$$

Так, в примере 1 параметр $\theta = (\theta_1, \theta_2, \theta_3)$. В примере 2 параметр θ в совокупности образуют средний уровень μ и главные эффекты $\alpha_1, \alpha_2, \dots$ и β_1, β_2, \dots . Наконец, в примере 3 это одномерный параметр β .

Если в гауссовском случае анализ обыкновенно начинают с оценивания неизвестного θ , то непараметрический анализ схемы (B.1) мы начинаем с построения разумного (еще лучше — в каком-либо смысле оптимального) статистического теста для проверки гипотезы

$$H_0: \theta = \theta^0, \quad (B.2)$$

где θ^0 задано. Такая проверка оказывается возможной, если от исходных наблюдений X перейти к статистике Y (функции от X и от θ^0), распределенной свободно при гипотезе H_0 из (B.2). Имеется в виду свобода от распределения ошибок, означающая, что распределение Y одно и то же при любых распределениях ошибок, которые допустимы по предположениям модели. В знаковом анализе это условие B.1. В наших примерах статистики Y суть соответственно

$$\text{sign}(x_i - \theta_1^0 - \theta_2^0 t_i - \theta_3^0 \sin(\omega t_i)), \quad i = 1, \dots, n,$$

$$\|\text{sign}(x_{ij} - \mu^0 - \alpha_i^0 - \beta_j^0)\|,$$

$$\text{sign}(u_k - \beta u_{k-1}), \quad k = 1, 2, \dots, n.$$

Когда выполнено условие B.1 и θ^0 — истинное значение, перечисленные выше случайные величины независимы и принимают значения 1 и -1 с вероятностями 0.5.

Уместно отметить, что в ранговом подходе принимают, что случайные ошибки независимы и распределены одинаково (и непрерывно). При этом предположении свободно распределенными оказываются совокупности рангов приведенных выше остатков.

После перехода от наблюдений X к статистике Y сложная гипотеза $H_0: \theta = \theta^0$ превращается в простую. Благодаря этому для проверки H_0 оказывается возможным указать статистические критерии (основанные на Y), уровень значимости которых можно контролировать и который сохраняется постоянным при всех допустимых распределениях случайных ошибок. Часто при построении таких критериев (знаковых, ранговых, знако-ранговых) удается воспользоваться теми или иными концепциями оптимальности. Это улучшает мощность критериев. В тексте книги мы рассказываем, как это сделать. Сейчас приведем лишь результаты. Скажем, в примере 1 локально оптимальная знаковая тестовая статистика для проверки $H_0: \theta_1 = \theta_2 = \theta_3 = 0$ есть

$$\left(\sum_{i=1}^n \text{sign } x_i \right)^2 + \left(\sum_{i=1}^n t_i \text{sign } x_i \right)^2 + \left(\sum_{i=1}^n \sin(\omega t_i) \text{sign } x_i \right)^2.$$

В примере 3 локально оптимальная тестовая статистика для проверки $H_0: \beta = \beta_0$ против односторонней альтернативы $H^+: \beta > \beta_0$ имеет вид

$$\sum_{m=1}^{n-1} \beta_0^{m-1} \left(\sum_{k=m+1}^n \text{sign}(u_k - \beta_0 u_{k-1})(u_{k+m} - \beta_0 u_{k+m-1}) \right). \quad (\text{B.3})$$

Против двусторонней альтернативы тестовой статистикой служит абсолютное значение статистики (B.3).

Распределения тестовых статистик могут зависеть от конкретного значения θ^0 . Это не мешает тому, чтобы вычислить эти распределения и их квантили сколь угодно точно при каждом значении θ^0 . Например, методом статистических испытаний. Проверяемую гипотезу следует отвергнуть при больших значениях критериальной статистики.

Путем «обращения» статистического теста для неизвестного параметра θ можно строить доверительные множества и даже точечные оценки. А именно, доверительное множество для истинного значения θ есть совокупность таких β , что гипотеза $H: \theta = \beta$ упомянутым знаковым критерием не отвергается. При этом если проверка производится на уровне значимости α , то соответствующая доверительная вероятность равна $1 - \alpha$. Пусть, к примеру, $c_{n,\alpha}(\beta)$ и $c_{n,1-\alpha}(\beta)$ суть квантили уровней α и $1 - \alpha$ случайной величины (B.3). Тогда доверительное множество для β уровня не меньше $1 - 2\alpha$ есть

$$\left\{ \beta: c_{n,\alpha}(\beta) \leq \leq \sum_{m=1}^{n-1} \beta^{m-1} \sum_{k=m+1}^n \text{sign}(u_k - \beta u_{k-1})(u_{k+m} - \beta u_{k+m-1}) \leq \leq c_{n,1-\alpha}(\beta) \right\}.$$

В тексте мы показываем, как такие множества можно строить явно. При больших n вместо точных квантилей можно использовать приближенные, получаемые с помощью гауссовской аппроксимации.

Для регрессионных моделей доверительные множества выглядят проще. Так, если гипотеза $H_0: \theta = 0$ принимается, когда $X - f(\theta) \in \mathbb{A}$, то доверительное множество для неизвестного истинного значения θ есть

$$\{\theta: X - f(\theta) \in \mathbb{A}\}. \quad (\text{B.4})$$

Чтобы перейти от доверительных множеств к точечным оценкам, надо сделать еще лишь шаг. Наши статистические критерии таковы, что порождаемые ими системы доверительных множеств упорядочены по вложению. Поэтому в качестве точечной оценки параметра можно взять их общую часть. Иными словами, в качестве точечной оценки параметра мы выбираем то значение переменной θ , при котором минимальна величина критериальной статистики. В одномерном случае (пример 3) этому соответствует решение задачи

$$\sum_{m=1}^{n-1} \beta^{m-1} \sum_{k=m+1}^n \text{sign}(u_k - \beta u_{k-1})(u_{k+m} - \beta u_{k+m-1}) \div 0;$$

символ \div означает переход функции через нулевой уровень.

Для регрессионных моделей дело опять упрощается, так как можно ограничиваться гипотезами вида $H_0: \theta = 0$. С помощью знаковых тестов эта гипотеза отвергается, если соответствующая критериальная статистика, скажем, $q(Y)$, оказывается чрезмерно большой. Поэтому в качестве оценки неизвестного θ в схеме (B.1) мы берем

$$\hat{\theta}_n = \arg \min_{\theta} q(X - f(\theta)). \quad (\text{B.5})$$

В частности, для примера 1 оценка находится решением задачи

$$\begin{aligned} & \left[\sum_{i=1}^n \text{sign}(x_i - \theta_1 - \theta_2 t_i - \theta_3 \sin \omega t_i) \right]^2 + \\ & + \left[\sum_{i=1}^n t_i \text{sign}(x_i - \theta_1 - \theta_2 t_i - \theta_3 \sin \omega t_i) \right]^2 + \\ & + \left[\sum_{i=1}^n (\sin \omega t_i) \text{sign}(x_i - \theta_1 - \theta_2 t_i - \theta_3 \sin \omega t_i) \right]^2 \Rightarrow \min_{\theta_1, \theta_2, \theta_3}. \end{aligned}$$

Мы исследуем асимптотические свойства знаковых оценок и показываем, что они состоятельны и \sqrt{n} -асимптотически нормальны. Для распределений с «тяжелыми хвостами» их асимптотическая эффективность относительно, например, оценок наименьших квадратов может быть сколь угодно большой. Знаковые оценки устойчивы к засорениям, как мы иллюстрировали выше.

Перечислим основные этапы статистического анализа, когда он основывается на знаках (эти этапы сложны для многих статистических моделей).

— Построение новой знаковой последовательности.

— Построение основанного на этой последовательности локально оптимального и непараметрического знакового теста для проверки гипотезы о фиксированном значении параметра.

— Построение доверительных множеств для конечных объемов выборок.

— Построение соответствующих непараметрических знаковых точечных оценок параметров.

— Асимптотическое исследование построенных тестов и оценок при нулевой гипотезе и близких альтернативах.

— Проверка линейных гипотез (для многопараметрических схем).

— Исследование устойчивости процедур по отношению к выбросам.

Уместно отметить, что описанным выше способом (новая последовательность — тест — доверительные оценки — точечные оценки — линейные гипотезы) можно анализировать данные не только с помощью знаков, но и рангов. Мы обсуждаем ранговый подход в книге только кратко и сравниваем его со знаковым.

Впервые описанный выше способ действий был предложен Дж.Ходжесом и Э.Леманом (см. [78]) в 60-х годах для задачи о двух выборках, отличающихся неизвестным сдвигом. Они применили для этого статистику ранговых сумм Уилкоксона. С тех пор для ранговых методов эта непараметрическая программа реализована в полном объеме усилиями многих авторов. Мы реализуем такую же программу для знаковых методов.

Книга состоит из двух частей. В первой части мы описываем непараметрические знаковые правила для линейных моделей с независимыми ошибками. Эта часть состоит из четырех глав. В главе 1 мы рассказываем о знаковом методе, избрав для этого простейший пример — однофакторную линейную модель. Мы выбрали для этого однофакторную линейную регрессию как математическую модель и закон Хаббла из астрономии как ее актуальное приложение. На этом объекте мы показываем основные этапы знакового (непараметрического) анализа: вывод рационального статистического критерия, основанный на постулируемых свойствах случайных ошибок; построение оценок параметров на основе статистических тестов; изучение асимптотических свойств упомянутых процедур при увеличении числа наблюдений.

Глава 2 посвящена выводу статистических знаковых критериев для общих линейных моделей и их изучению. Особое внимание уделено практическому вычислению критических значений для малых выборок.

Глава 3 отведена знаковому оцениванию. Мы определяем знаковые оценки с помощью знаковых тестовых статистик и затем обсуждаем вопросы их практического вычисления. Большую же часть главы 3 занимает асимптотическая теория знакового оценивания. Из относительно новых для непараметрического оценивания явлений можно отметить единственность (глобальную, во всем пространстве) знаковых оценок. Из технических новшеств упомянем доказательство асимптотической линейности знаковых статистик на расширяющихся компактах. Далее мы сопоставляем знаковые оценки с другими

оценками и устанавливаем их близость к оценкам наименьших модулей. Последнее обстоятельство может иметь важные последствия, так как оно позволяет дополнить метод наименьших модулей знаковыми правилами интервального оценивания и знаковыми статистическими критериями.

Последняя в первой части глава 4 посвящена проверке линейных гипотез и их многочисленным конкретным примерам.

Во второй части книги мы описываем непараметрические знаковые процедуры для моделей авторегрессии. Эта часть книги состоит из трех глав. В главе 5 мы даем обзор наиболее известных и употребительных статистических методов анализа авторегрессии — метода наименьших квадратов и метода наименьших модулей. Мы изучаем их свойства на примере авторегрессии первого порядка. Эти результаты затем используются в главах 6 и 7 для сравнения этих классических процедур с непараметрическими знаковыми. Впрочем, лишь часть результатов главы 5 является известной. Многие результаты являются новыми. Например, поведение тестов и оценок при близких альтернативах, метод исследования оценок и тестов наименьших модулей, некоторые факты о поведении оценок при засорении, результаты о проверке стационарности авторегрессии и т.д. Все заимствованные результаты снабжены ссылками. Отсутствие ссылки означает, что результат является оригинальным. Таких принципов цитирования мы придерживаемся во всей книге.

Собственно знаковому анализу однопараметрической стационарной авторегрессии посвящена глава 6. Мы предпослали ей подробное введение. Знакомясь с ним, читатель сможет составить мнение о сути знакового подхода. В главе 7 мы распространяем действия знакового метода на многопараметрическую авторегрессию. Эту главу тоже предваряем подробным введением. В главе 5 мы даем обзор наиболее употребительных процедур статистического анализа авторегрессии — наименьших квадратов и наименьших модулей. Эти результаты понадобятся нам в главах 6 и 7 для сравнительного анализа общепотребительных процедур с непараметрическими знаковыми. Впрочем, лишь часть материала главы 5 является общеизвестной. Многие результаты (например, поведение тестов и оценок при близких альтернативах, метод исследования оценок и тестов наименьших модулей, некоторые факты о поведении оценок при засорениях, результаты о проверке стационарности авторегрессии и т.д.) являются оригинальными.

Все заимствованные результаты снабжены ссылками; отсутствие ссылки означает, что ранее этот факт в литературе не обсуждался.

ЛИНЕЙНЫЕ МОДЕЛИ НЕЗАВИСИМЫХ НАБЛЮДЕНИЙ

Глава I

ЗНАКОВЫЙ АНАЛИЗ ОДНОПАРАМЕТРИЧЕСКОЙ РЕГРЕССИИ

В этой главе мы разъясняем существо знакового метода на примере однопараметрической линейной регрессии. Знаменитый закон Хаббла из астрономии описывается именно такой статистической моделью. Чтобы показать действие знакового метода на реальных данных, мы выбрали закон Хаббла.

§ 1.1. Закон Хаббла: история и современность

В 1929 г. американский астроном Э. Хаббл (Edwin P. Hubble) опубликовал статью [80], которая сыграла важнейшую роль в современном естествознании. В этой статье он сообщил об удивительной закономерности, которую он обнаружил, изучая внегалактические туманности — галактики, огромные звездные системы, аналогичные Галактике, в которой мы живем. Он заметил, что эти туманности удаляются от Земли, и тем быстрее, чем они дальше. Более того, скорость этого удаления оказалась пропорциональной расстоянию.

Сейчас эту пропорциональность называют *законом Хаббла*, а коэффициент пропорциональности в законе — *постоянной Хаббла*. О том, какова ее величина, в астрономии до сих пор идет дискуссия, хотя сама линейная зависимость признается безусловно. Теперь явление «разбегания» галактик истолковывается как свидетельство расширения пространства.

Закон линейного расширения (закон Хаббла) должен быть статистическим по самой природе. Дело в том, что галактики, помимо общей вовлеченности в расширение пространства, имеют и присущие им собственные движения относительно друг друга. В силу этого их лучевые скорости могут отличаться от предписанных законом Хаббла. Мы будем считать эти отклонения от линейной зависимости случайными и взаимно независимыми.

В 1917 г. в США был введен новый телескоп — 100-дюймовый рефлектор, который по своим возможностям значительно превосходил все существовавшие до того времени инструменты. Вскоре на нем начал работать вернувшийся с войны молодой Эдвин Хаббл. Его интересовали туманности. В то время так называли все небесные объекты, которые при фотографировании не распадались на звезды.

Сначала Хаббл изучал туманности, чтобы разобраться в их многообразии. Тогда много спорили, как их классифицировать, и потому больше обращали внимание на их формы в «целом». Но затем

Хаббл занялся детальным изучением некоторых из них — тех, на фотографиях которых, сделанных с помощью телескопов, проявлялись отдельные звезды. Прежде всего он обратил внимание на крупные и, как можно было думать, близкие туманности в созвездиях Андромеды и Треугольника. Ключевой шаг к открытию закона Хаббл сделал в 1923–1924 гг.: в туманности Андромеды он нашел цефеиду. Цефеиды — звезды замечательные. Они периодически изменяют свой блеск, и логарифмы их периодов пропорциональны абсолютной звездной величине в максимуме блеска, характеризующей их истинную светимость (т.е. блеск звезды, находящейся на некотором стандартном расстоянии). Эту закономерность незадолго до этого (в 1918 г.) открыл Х.Шепли. Наблюдая цефеиду, можно определить ее период и видимый блеск. Сравнивая же абсолютную и видимую звездные величины, можно оценить расстояние до цефеиды. А стало быть, и до галактики, в которой она находится. Посмотрите, например, как выглядела построенная Хабблом кривая блеска первой открытой им внегалактической цефеиды (рис. 1.1.1).

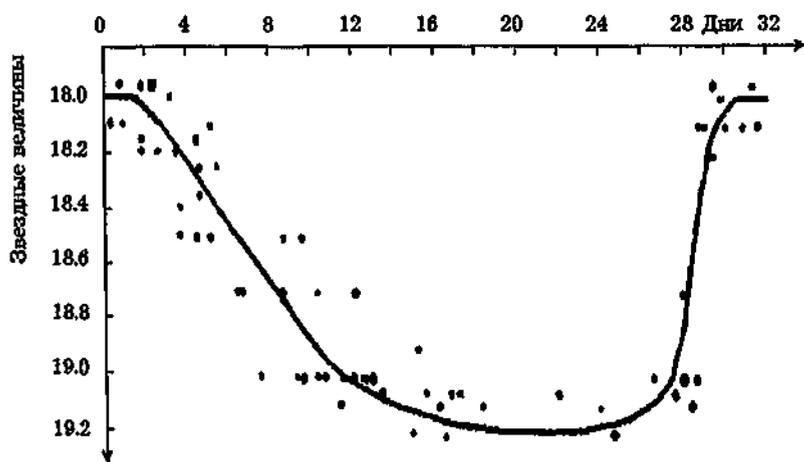


Рис. 1.1.1. Кривая блеска первой открытой Хабблом цефеиды в туманности Андромеды (с любезного разрешения авторов воспроизведено из [43])

Впервые в науке были найдены основанные на надежных индикаторах расстояния до туманности Андромеды и до туманности Треугольника. Они оказались огромными — порядка миллиона световых лет. Одновременно Хаббл фотографировал и другие туманности. Он заметил, что в галактиках, расстояния до которых ему уже стали известны, наиболее яркие звезды имели практически одинаковую абсолютную светимость. Поэтому по видимому блеску наиболее ярких звезд можно было также судить о расстояниях до галактик. К сожалению, менее точно, чем по цефеидам. Так или иначе, но к 1929 г. (времени появления упомянутой статьи) Хаббл определил расстояния уже до двух десятков галактик.

Отметим, что в дальнейшем расстояние до туманности Андромеды не раз уточнялось. Выяснилось, что светимости тех цефеид, которые были основой для упомянутой зависимости Шепли, были измерены неправильно. Это привело к появлению систематических ошибок во всех последующих оценках расстояний. Были обнаружены и другие неточности. Естественно, что исправление этих ошибок изменило и оценку постоянной Хаббла. Впрочем, на линейном характере зависимости «расстояние—скорость удаления» эти ошибки и поправки не отражаются. Мы приводим здесь данные, которыми оперировал Хаббл. Они взяты из [109]. Нумерация туманностей соответствует NGC (New General Catalog).

Таблица 1.1.1

Туманность	Расстояние (Мпс)	Скорость (км/с)
SMC	0.032	170
LMC	0.034	290
NGC 6822	0.214	-130
598	0.263	-70
221	0.275	-185
224	0.275	-220
5457	0.450	200
4736	0.500	290
5194	0.500	270
4449	0.630	200
4214	0.800	300
3031	0.900	-30
3627	0.900	650
4826	0.900	150
5236	0.900	500
1068	1.000	920
5055	1.100	450
7331	1.100	500
4258	1.400	500
4151	1.700	960
4382	2.000	500
4472	2.000	850
4486	2.000	800
4649	2.000	1090

При измерении расстояний до туманностей на результат оказывают влияние многие ошибки — как общие для всех измерений (систематические), так и относящиеся к отдельным галактикам или их группам. Последние ошибки мы расцениваем как случайные и независимые.

Теперь пора напомнить, что лучевые скорости небесных объектов определяют по смещению в их спектрах. В основе этого метода

лежит эффект Доплера. Из-за эффекта Доплера для движущегося тела частота спектральной линии уменьшается, если тело удаляется от наблюдателя, и увеличивается в случае приближения. Поэтому спектральные линии сдвигаются в длинноволновую или коротковолновую части спектра, для оптической области — в красную или голубую. Изучая спектры светил, можно весьма точно определить их лучевые скорости. Первое измерение лучевой скорости туманности было сделано В.Слайфером в 1912 г. Это была туманность Андромеды. Оказалось, что она приближается к Земле со скоростью 300 км/с. К 1917 г. Слайфер довел число туманностей с измеренными лучевыми скоростями до 25; к 1925 г. таких измерений было проделано уже 45. Дело шло медленно, так как для каждого измерения требовались многочасовые и даже многосуточные экспозиции.

К 1929 г., к моменту написания статьи, Э.Хаббл располагал скоростями для 46 туманностей. Лишь немногие из этих скоростей были отрицательными, т.е. соответствующие туманности приближались к Земле. Они, к слову, были из числа ближайших. У остальных скорости были положительными, достигая 1000 км/с. Простое сопоставление скоростей и расстояний, как это сделано на рис. 1.1.2, убеждает нас, как убедило и Хаббла, что зависимость между ними существует.

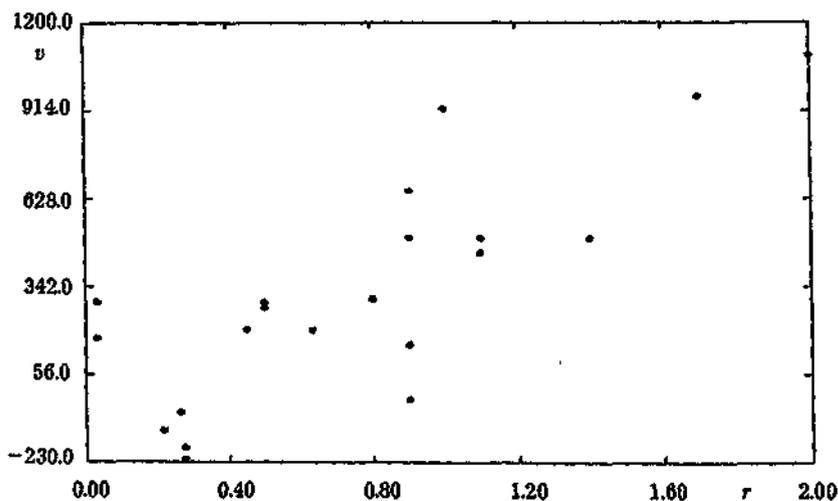


Рис. 1.1.2. Данные из статьи Э.Хаббла (1929 г.). По горизонтальной оси отложены расстояния в Мпс, по вертикальной — скорости в км/с

Представление о расширении пространства (о разбегании галактик) сейчас принято наукой и вошло в основы современных космологических теорий и в интерпретацию наблюдений. Однако о численном значении постоянной Хаббла идут споры. Обсуждаются два возможных значения: $(50 \pm 10) \frac{\text{км}}{\text{с} \cdot \text{Мпс}}$ и $(80 \pm 10) \frac{\text{км}}{\text{с} \cdot \text{Мпс}}$. Оценки постоянной Хаббла извлекаются из обширного наблюдательного материала. Слабым местом остается оценка базовых расстояний, через которые затем

определяют все остальные. Из-за этого и возникают разные оценки. Остроту проблеме придает то, что значение постоянной Хаббла при обратной (во времени) экстраполяции наблюдаемого сейчас расширения пространства приводит к разным оценкам возраста Вселенной. При большем из двух названных выше чисел возраст оказывается меньше, чем возраст самых старых звезд. Этот возраст, конечно, тоже гипотетичен. Тем не менее современная теория развития звезд устанавливает его именно таким, так что статистическая задача оценки коэффициента наклона линии регрессии оказалась связана с острейшими проблемами космологии.

Нелишним будет заметить, что одна из целей, ради которых был запущен в пространство оптический телескоп «Хаббл», состоит в уточнении постоянной Хаббла, для чего необходимо обнаружить и исследовать цефеиды в существенно более далеких галактиках, чем туманность Андромеды.

В одном из последних (ко времени написания этой книги) сообщений [66] говорилось о том, что с помощью этого телескопа такие цефеиды были обнаружены. Это позволило заново оценить постоянную Хаббла: $(80 \pm 17) \frac{\text{км}}{\text{с} \cdot \text{Мпс}}$. Полученное значение согласуется с так называемой короткой хронологией. Так что острота научной проблемы еще увеличилась.

§ 1.2. Определение постоянной Хаббла знаковым методом

Обозначим через r удаление небесного объекта от Земли. Через v обозначим его лучевую скорость, т.е. мгновенную скорость его движения вдоль направления «наблюдатель—объект». Предположим, что для некоторого числа n небесных объектов мы располагаем измерениями упомянутых величин. Пусть эти объекты каким-либо способом занумерованы числами от 1 до n . Через r_i, v_i мы обозначаем соответственно удаление и лучевую скорость для объекта $i, i = 1, \dots, n$. Ввиду отмеченной ранее статистической природы закона Хаббла и неизбежных ошибок при измерениях r_i и v_i связь между ними не сводится к чистой пропорциональности. Соотношения между v_i и r_i (закон Хаббла) надо записать в виде

$$v_i = r_i \theta + \xi_i, \quad i = 1, \dots, n. \quad (1.2.1)$$

Здесь θ — коэффициент пропорциональности, а поправки ξ_1, \dots, ξ_n выражают отклонения от линейной зависимости. В величину ξ_i входят и возможные ошибки в измерениях r_i, v_i , и упомянутые ранее отступления от пропорциональности из-за собственного движения туманностей. Мы будем считать эти отклонения независимыми случайными величинами. Это предположение составляет основу статистической модели. Затем мы предположим, что в наших наблюдениях нет систематических ошибок. Для знакового анализа это предположение принимает форму основного условия

$$P\{\xi_i > 0\} = P\{\xi_i < 0\} = 1/2. \quad (1.2.2)$$

В статистической теории часто принимают предположение о том, что случайные ошибки распределены одинаково. Если это предположение кажется неподходящим, то приходится принимать какую-либо конкретную схему, по которой изменяются распределения случайных ошибок. (Скажем, в (1.2.1) можно было бы предположить, что при сохранении формы распределения ξ_i его масштаб изменяется как некоторая определенная функция от r_i .) К сожалению, выбор такой схемы трудно сделать обоснованно. Между тем ясно, что всякая ошибка при конструировании статистической модели явления неизбежно порождает ошибки при ее применении. Для знакового анализа предположение об одинаковой распределенности случайных ошибок ξ_1, \dots, ξ_n не является обязательным.

Перечислим предположения, которые составляют статистическую модель, с помощью которой мы будем анализировать данные табл. 1.1.1. Между величинами v_i и r_i , $i = 1, \dots, n$ ($n = 24$), существует зависимость (1.2.1). Поправки ξ_1, \dots, ξ_n мы считаем независимыми случайными величинами, для которых верно (1.2.2). Коэффициент θ в (1.2.1) не известен и подлежит оценке по наблюдениям. Истинное значение θ — искомую постоянную Хаббла — мы обозначим как θ^0 . В качестве приближения для θ^0 мы далее возьмем такое значение θ , при котором наблюдения (r_i, v_i) , $i = 1, \dots, n$, находятся в наилучшем согласии с линейной зависимостью (1.2.1). При этом это согласие мы станем проверять не по самим видимым остаткам $v_i - \theta r_i$, а по совокупности их знаков

$$(\text{sign}(v_1 - \theta r_1), \text{sign}(v_2 - \theta r_2), \dots, \text{sign}(v_n - \theta r_n)).$$

Как предлагает описанная во введении «непараметрическая программа», статистический анализ (1.2.1) надо начать с гипотезы $H_0: \theta = 0$. Для ее проверки против альтернативы $H: \theta \neq 0$ нам нужен разумный статистический критерий.

Мы воспользуемся знаковым критерием (2.2.9) из гл. 2, точнее, его двусторонним вариантом (2.2.11). В гл. 2, § 2.2 мы показываем, что для одинаково распределенных ошибок и против односторонних альтернатив критерий (2.2.9) со статистикой

$$\sum_{i=1}^n r_i \text{sign } v_i \quad (1.2.3)$$

является локально наиболее мощным знаковым критерием. Согласно (2.2.9) гипотезу $H_0: \theta = 0$ надо отвергать в пользу $H^+: \theta > 0$, если

$$\sum_{i=1}^n r_i \text{sign } v_i \geq z, \quad (1.2.4)$$

где критическое значение z надо выбрать так, чтобы при гипотезе вероятность

$$P\left\{\sum_{i=1}^n r_i \text{sign } v_i \geq z\right\} \quad (1.2.5)$$

была равна заданному значению (уровню значимости, малому положительному числу). Против двусторонних альтернатив мы предлагаем применять двустороннюю версию этого критерия, т.е. отвергать гипотезу $H_0: \theta = 0$ в пользу $H: \theta \neq 0$, если

$$\left| \sum_{i=1}^n r_i \operatorname{sign} v_i \right| \geq \operatorname{const}. \quad (1.2.6)$$

Для неодинаково распределенных ошибок, что, по-видимому, свойственно разбираемой задаче, критерий (1.2.4) не является оптимальным. Однако он сохраняет предписанный уровень значимости и в этом случае, если выполнено (1.2.2). Поэтому им можно пользоваться и в данном случае.

Для выбора критических значений в (1.2.4) или (1.2.6) надо знать распределение при гипотезе H_0 тестовой статистики (1.2.3). Введем случайную величину

$$T_n = \sum_{i=1}^n r_i \zeta_i, \quad (1.2.7)$$

где случайные величины ζ_1, \dots, ζ_n принимают значения $+1$ и -1 с вероятностями $1/2$ и независимы в совокупности. Случайная величина (1.2.7) распределена так же, как (1.2.3), когда выполнена гипотеза H_0 . Рассмотрим распределение T_n . Для $n = 24$ распределение T_n для r_i , взятых из табл. 1.1.1, мы вычислили методом Монте-Карло при числе статистических испытаний $N = 50000$. (Все вычисления в книге сделаны с помощью пакета программ Sign, разработанного авторами). График полученной таким образом функции распределения $F(x)$ случайной величины T_n изображен на рис. 1.2.1.

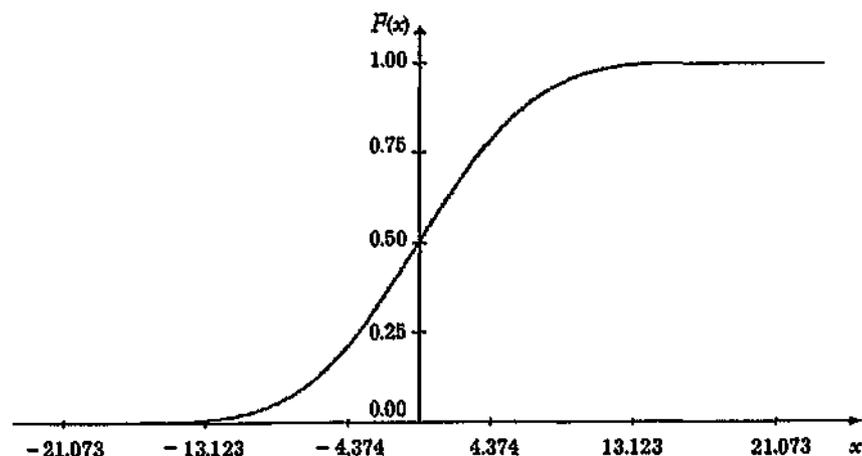


Рис. 1.2.1. График функции распределения случайной величины T_n (1.2.7), $n = 24$

Видно, что распределение случайной величины T_n (1.2.7) близко к нормальному. Это становится особенно ясным, если изобразить его

функцию распределения на нормальной вероятностной бумаге, как на рис. 1.2.2.

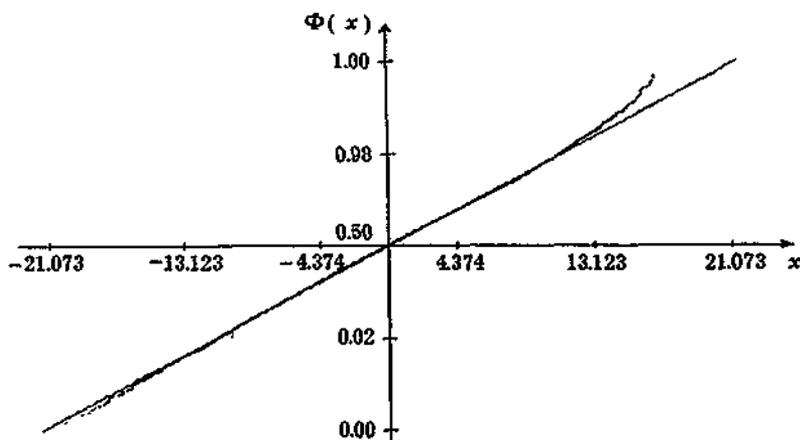


Рис. 1.2.2. Функция распределения статистики T_n в нормальном масштабе. По оси ординат отложены квантили функции распределения стандартного нормального закона, соответствующие указанным на оси вероятностям

Случайная величина T_n (1.2.7) при $n \rightarrow \infty$ распределена асимптотически как $N\left(0, \sum_{i=1}^n r_i^2\right)$, если последовательность r_1, \dots, r_n удовлетворяет условию

$$\max_{1 \leq i \leq n} r_i^2 / \sum_{i=1}^n r_i^2 \rightarrow 0. \quad (1.2.8)$$

В нашем примере (для данных из табл. 1.1.1) значение (1.2.8) приблизительно равно 0.14. Так что в данном случае $n = 24$ оказалось «достаточно большим», а 0.14 в качестве (1.2.8) «достаточно малым», чтобы нормальная аппроксимация точного распределения T_n была удовлетворительной в центральной части распределения.

Для проверки гипотезы $H_0: \theta = 0$ и для получения доверительных интервалов согласно (В.4) для неизвестного истинного значения θ (если H_0 будет отвергнута) нам необходимы квантили распределения T_n . Так как T_n имеет дискретное распределение, то уравнение для квантилей

$$P\{T_n \leq x\} = \epsilon$$

либо

а) не имеет решения,

либо

б) его решения составляют полуинтервал на числовой оси.

В случае а) ε -квантилью $T_{n,\varepsilon}$ назовем

$$T_{n,\varepsilon} = \sup \{x: \mathbf{P}\{T_n \leq x\} \leq \varepsilon\}.$$

В случае б) в качестве $T_{n,\varepsilon}$ возьмем центр упомянутого полуинтервала. Так как распределение T_n центрально симметричное, то $T_{n,1-\varepsilon} = -T_{n,\varepsilon}$. В нашем случае функция распределения T_n близка к непрерывной с достаточной для нас точностью (см. рис. 1.2.1 и 1.2.2). Поэтому мы без особых оговорок в качестве $T_{n,\varepsilon}$ будем использовать ее приближенное значение (как его дает, например, рис. 1.2.1). Для некоторых значений ε -квантили случайной переменной T_n , вычисленные методом Монте-Карло, приведены в табл. 1.2.1.

Таблица 1.2.1

ε	$T_{n,\varepsilon}$	$T_{n,1-\varepsilon}$
0.075	-7.9	7.9
0.05	-9.0	9.0
0.025	-10.6	10.6
0.0125	-12.0	12.0

Значение критериальной статистики (1.2.6) оказалось равным 18.02, что превосходит 2.5-процентный двусторонний критический уровень. Поэтому гипотеза $H_0: \theta = 0$ уверенно отвергается. Это означает, что между лучевыми скоростями туманностей и их удалениями действительно существует линейная зависимость.

Для интервального оценивания неизвестного истинного значения θ^0 согласно (В.4) надо рассмотреть функцию

$$T_n(\theta) = \sum_{i=1}^n r_i \operatorname{sign}(v_i - \theta r_i). \quad (1.2.9)$$

Доверительным множеством, содержащим θ^0 с вероятностью, не меньшей $1 - 2\alpha$, служит

$$\{\theta: |T_n(\theta)| \leq T_{n,1-\alpha}\}. \quad (1.2.10)$$

Функцию (1.2.9) можно также записать как

$$T_n(\theta) = \sum_{i=1}^n |r_i| \operatorname{sign}\left(\frac{v_i}{r_i} - \theta\right). \quad (1.2.11)$$

Из последней формулы видно, что $T_n(\theta)$ — кусочно-постоянная и невозрастающая функция. Она изменяется от $\sum_{i=1}^n |r_i|$ до $-\sum_{i=1}^n |r_i|$ и имеет скачки, равные $2|r_i|$, в точках v_i/r_i , $i = 1, \dots, n$. Из этих свойств

следует, что (1.2.10) — интервал, концами которого служат какие-то значения v_i/r_i . На рис. 1.2.3 приведен график функции $T_n(\theta)$ для значений r_i из табл. 1.1.1 и 90-процентный доверительный интервал для θ^0 .

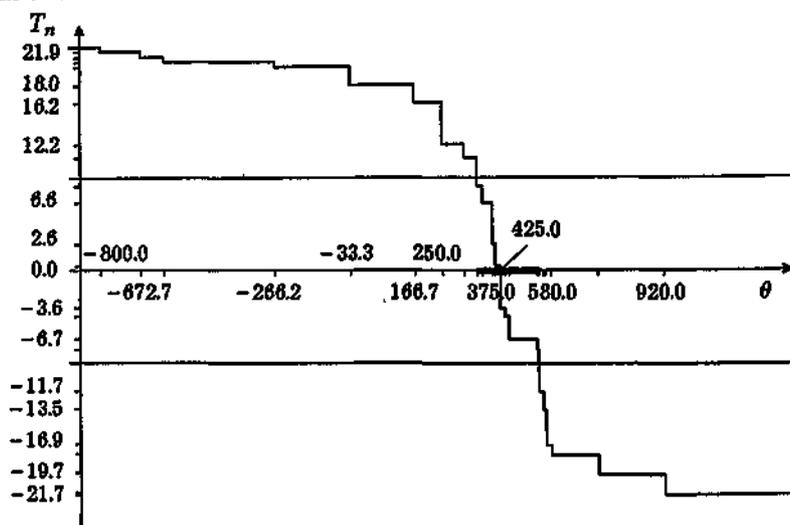


Рис. 1.2.3. График функции $T_n(\theta)$ и 90-процентный доверительный интервал (357, 540) для постоянной Хаббла

Доверительные интервалы, отвечающие иным доверительным вероятностям, точнее, их левые θ_l и правые θ_u концы, приведены в табл. 1.2.2.

Т а б л и ц а 1.2.2

Доверительная вероятность $1-2\alpha$	Левый конец доверительного интервала θ_l	Правый конец доверительного интервала θ_u
0.85	375.00	540.00
0.90	357.14	540.00
0.95	357.14	545.00
0.975	314.46	555.60

Остается сказать о точечной оценке. Согласно (В.5) ею является решение задачи

$$\left| \sum_{i=1}^n r_i \operatorname{sign}(v_i - r_i \theta) \right| \Rightarrow \min_{\theta}. \quad (1.2.12)$$

В силу отмеченных выше свойств функции $T_n(\theta)$ минимальное значение в (1.2.12) достигается на интервале. Это минимальное значение равно 0.359, и оно достигается на интервале $(\hat{\theta}_1, \hat{\theta}_2) =$

= (409.1, 425.0). Любое число из этого интервала можно считать оценкой истинного значения θ^0 .

Обратим внимание, насколько неопределенность при выборе численного значения оценки $\hat{\theta}_n$ в (1.2.12) меньше статистической неопределенности, сопутствующей оцениванию θ^0 . Представление о порядке последней дают длины доверительных интервалов для θ^0 в приведенной выше табл. 1.2.2. Именно это качественное различие позволяет употреблять выражение «точечная» оценка даже в том случае, когда решением экстремальной задачи (1.2.12) является множество. На рис. 1.2.4 представлен результат оценивания направления, вдоль которого располагаются 24 пары измерений расстояния и скорости.

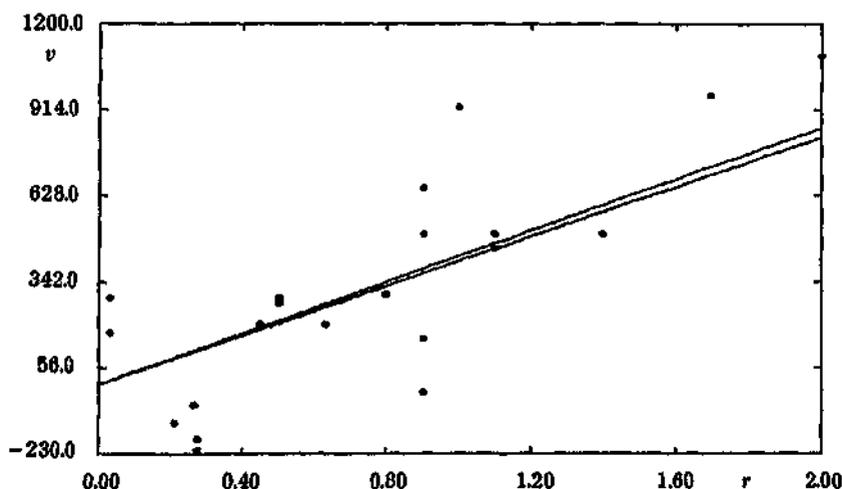


Рис. 1.2.4. Оценивание постоянной Хаббла. Графики функций $v = 409.1r$ и $v = 425r$. По горизонтальной оси отложены расстояния в Мпс, по вертикальной оси — скорости в км/с

Графики прямых $v = \hat{\theta}_1 r$ и $v = \hat{\theta}_2 r$, коэффициентами наклона которых являются границы интервала, на котором функция $|T_n(\theta)|$ принимает минимальное значение, различаются мало. К знаковым оценкам близки оценки наименьших модулей. Оценка наименьших модулей в модели (1.2.1) определяется как решение экстремальной задачи

$$\sum_{i=1}^n |v_i - \theta r_i| \Rightarrow \min_{\theta}$$

Целевая функция в последней экстремальной задаче как функция θ является выпуклой, и решение всегда существует. Легко видеть, что

его можно определить как решение уравнения

$$\sum_{i=1}^n \text{sign}(v_i - \theta r_i) \div 0. \quad (1.2.13)$$

Символ \div означает переход через нуль; решением этого уравнения называют такое значение, скажем, $\hat{\theta}_{LD}$, при котором происходит изменение знака монотонной функции $T_n(\theta)$. В нашем примере изменение знака $T_n(\theta)$ происходит при значении аргумента, равном 425.0. Поэтому $\hat{\theta}_{LD} = 425.0$.

Оценку наименьших модулей здесь можно указать не только численно, но и в общем виде, если вспомнить свойства функции $T_n(\theta)$ (1.2.9), отмеченные ранее. Из них следует, что $\hat{\theta}_{LD}$ есть медиана дискретного распределения вероятностей, когда вероятности

$$p_i = |r_i| / \sum_{i=1}^n |r_i|$$

сосредоточены в точках v_i/r_i , $i = 1, \dots, n$:

$$\hat{\theta}_{LD} = \text{med} \left\{ (v_i/r_i, p_i), \quad i = 1, \dots, n \right\}. \quad (1.2.14)$$

Мы уже упоминали выше, что данные, на которые опирался Хаббл (табл. 1.1.1), впоследствии были пересмотрены и уточнены.

В табл. 1.2.3 мы приводим их современные значения, как они даны в [109]. Расстояния получены с использованием LEDA (Lyon-Mendon Extragalactic Database, 1995). Ниже мы повторяем прежние вычисления, но с современными данными (табл. 1.2.3). Обратите внимание, насколько близкой к современным значениям постоянной Хаббла окажется знаковая оценка, которую мы сейчас получим, основываясь на столь малой выборке (всего 24 наблюдения).

Расположение 24 точек на плоскости и графики функции распределения статистики T_n из (1.2.7) изображены на рис. 1.2.5–1.2.7.

В табл. 1.2.4 приведены некоторые ϵ -квантили распределения статистики T_n , вычисленные методом Монте-Карло при числе испытаний $N = 50000$. Ступенчатая функция $T_n(\theta)$ из (1.2.11) изменяется от 150.92 до -150.92 .

На рис. 1.2.8 приведен ее график и 90-процентный доверительный интервал для постоянной Хаббла. Этот интервал есть (53.02, 85.73). Левые и правые границы доверительных интервалов, отвечающих другим доверительным вероятностям, указаны в табл. 1.2.5.

Таблица 1.2.3

Туманность	Расстояние (Мпс)	Скорость (км/с)
SMC	0.083	190
LMC	0.047	277
NGC 6822	1.60	-26
598	1.44	-204
221	1.57	-205
224	0.98	-295
5457	15.30	221
4736	8.05	297
5194	4.30	463
4449	3.68	211
4214	2.75	298
3031	5.10	-49
3627	9.70	703
4826	7.10	474
5236	7.28	503
1068	12.75	1093
5055	9.55	516
7331	15.75	835
4258	7.60	480
4151	6.28	956
4382	6.34	722
4472	6.30	983
4486	8.12	1282
4649	9.25	1114

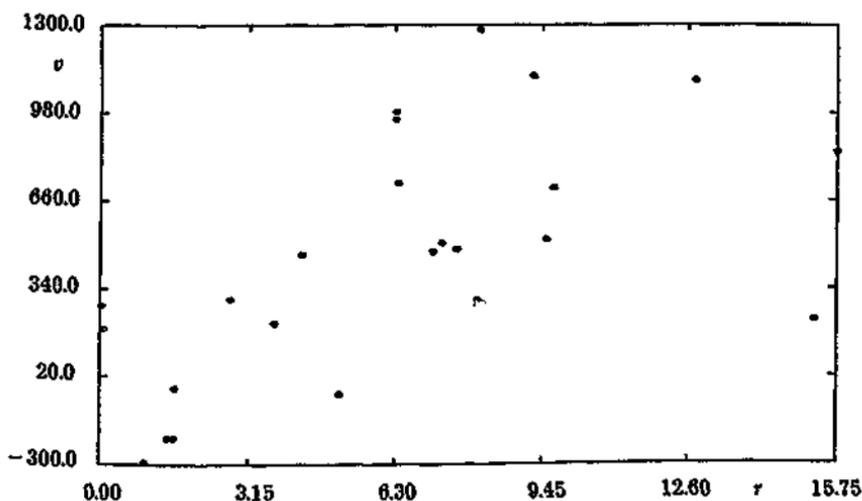


Рис. 1.2.5. Современные данные. По горизонтальной оси отложены расстояния в Мпс, по вертикальной оси — скорости в км/с

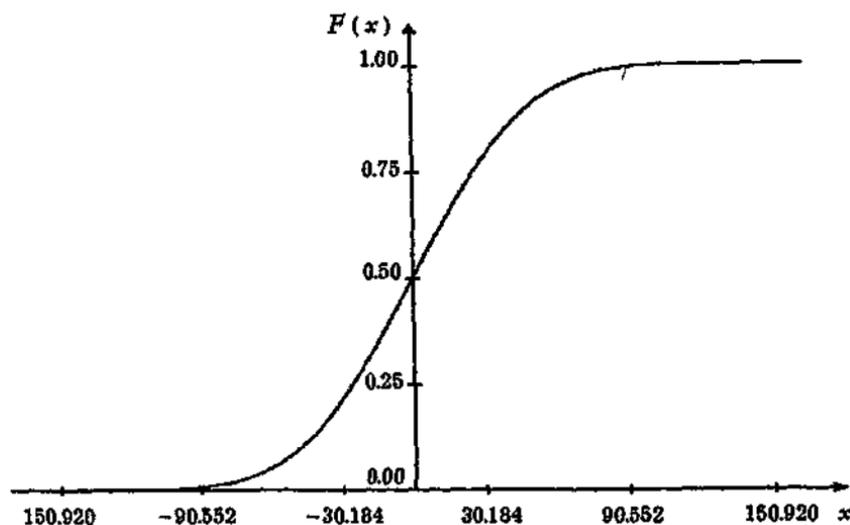
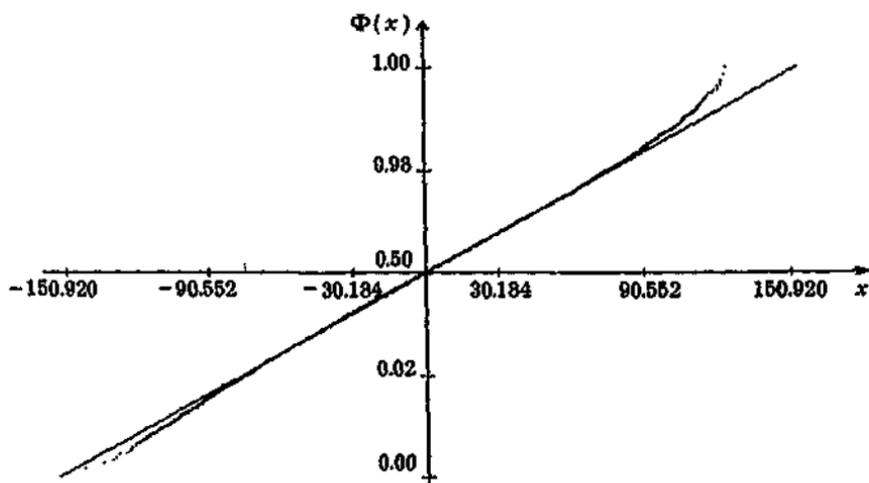
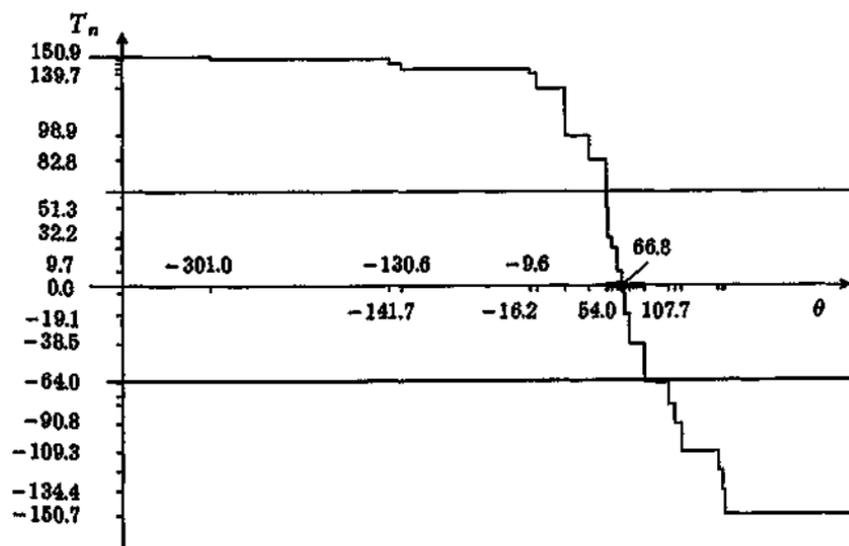
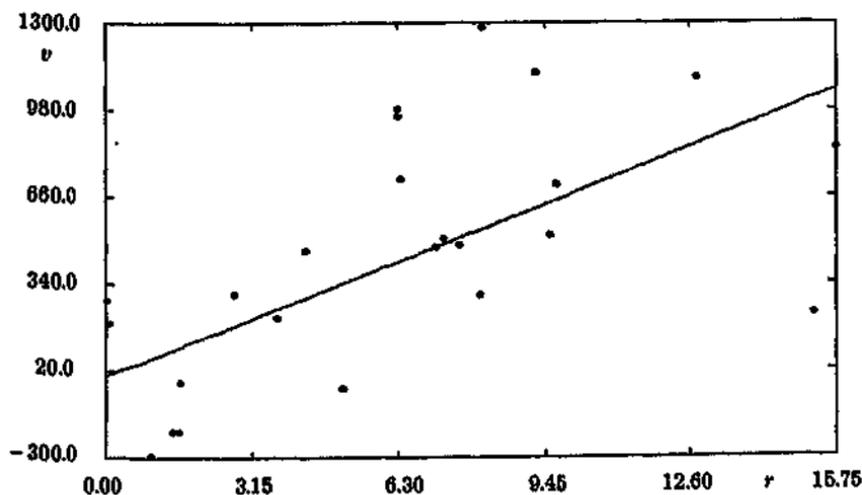
Рис. 1.2.6. График функции распределения случайной величины T_n (1.2.7), $n=24$ Рис. 1.2.7. Функция распределения T_n в нормальном масштабе. По оси ординат отложены квантили функций распределения стандартного нормального закона, соответствующие указанным вероятностям

Таблица 1.2.4

θ	$T_{n,\theta}$	$T_{n,1-\theta}$
0.075	-54	54
0.05	-62	62
0.025	-73	73
0.0125	-82	82

Таблица 1.2.5

Доверительная вероятность $1-2\alpha$	Левый конец доверительного интервала θ_l	Правый конец доверительного интервала θ_r
0.85	53.02	85.73
0.90	53.02	85.73
0.95	53.02	108.36
0.975	53.02	113.88

Рис. 1.2.8. График функции $T_n(\theta)$ и 90-процентный доверительный интервал (53.02, 85.73) для постоянной ХабблаРис. 1.2.9. Оценивание постоянной Хаббла. График функции $v = 66.76r$

Решением экстремальной задачи (1.2.12) является единственное значение 66.76. Соответствующее минимальное значение функции равно 2.58. На рис. 1.2.9 изображена прямая, наклоном которой является знаковая оценка постоянной Хаббла 66.76.

§ 1.3. Асимптотические результаты

В этом параграфе мы исследуем свойства знаковых правил при большом числе наблюдений. В этом случае распределение статистик, их критических значений, доверительных границ и т.п. становятся более простыми, чем при малых n . В типичных условиях это связано с действием центральной предельной теоремы и основанной на ней нормальной аппроксимации. По отношению к статистике (1.2.3) или случайной величине (1.2.7) теорема о нормальной аппроксимации (достаточные условия) уже была высказана. Из этой теоремы следует нормальная аппроксимация и для квантилей $T_{n,c}$. Впрочем, вопрос о том, какое число наблюдений можно считать достаточно большим, чтобы полагаться на выводы асимптотической теории, до сих пор не имеет ясного теоретического ответа. Вопрос этот для разных статистических задач решается скорее практически, на основании сопоставления теоретических результатов с вычислениями. Поэтому в наших практических рекомендациях мы всюду, где возможно, отдаем предпочтение точным методам либо методам с контролируемой точностью (методу Монте-Карло, например). Тем не менее асимптотические методы составляют важную и полезную часть статистической теории. По отношению к знаковой оценке $\hat{\theta}_n$ (1.2.12) или (1.2.13) главные асимптотические результаты — это утверждения о ее распределении при $n \rightarrow \infty$: состоятельность и асимптотическая нормальность.

1.3.1. Состоятельность. Напомним общее определение: оценка $\hat{\theta}_n$ параметра θ называется *состоятельной*, если $\hat{\theta}_n \rightarrow \theta$ (по вероятности) во всей области возможных значений θ , когда $n \rightarrow \infty$. На практике состоятельность истолковывается как приближенное равенство $\hat{\theta}_n \approx \theta^0$ при больших n , где θ^0 — это значение, принятое параметром θ в данном опыте (так называемое истинное значение θ).

Знаковая оценка $\hat{\theta}_n$ в модели (1.2.1) определена нами как решение экстремальной задачи (1.2.12). Покажем, что при некоторых естественных условиях знаковая оценка состоятельна. Положим

$$B_n^2 = \sum_{i=1}^n r_i^2$$

и рассмотрим эквивалентную задачу

$$\left| \frac{1}{B_n^2} T_n(\theta) \right| \Rightarrow \min_{\theta}$$

Как мы уже отмечали, функция $T_n(\theta)$ (1.2.9) монотонно невозрастающая. Поэтому для доказательства состоятельности $\hat{\theta}_n$ достаточно показать, что при фиксированном θ и $n \rightarrow \infty$:

$$\text{а) } \mathbf{P}\left\{\frac{1}{B_n^2}T_n(\theta) < 0\right\} \rightarrow 1, \text{ если } \theta > \theta^0;$$

$$\text{б) } \mathbf{P}\left\{\frac{1}{B_n^2}T_n(\theta) < 0\right\} \rightarrow 0, \text{ если } \theta < \theta^0;$$

$$\text{в) } \frac{1}{B_n^2}T_n(\theta^0) \rightarrow 0 \text{ по вероятности.}$$

Нетрудно показать, что при условии (1.2.8) случайная величина $T_n(\theta)$ (1.2.9) при любом фиксированном θ и $n \rightarrow \infty$ асимптотически нормальна: $N(\mathbf{E}T_n(\theta), \mathbf{D}T_n(\theta))$. Здесь

$$\mathbf{E}T_n(\theta) = \sum_{k=1}^n r_k p_k,$$

$$\mathbf{D}T_n(\theta) = \sum_{k=1}^n r_k^2(1 - p_k^2),$$

где

$$p_k := \mathbf{E} \text{sign}(v_k - \theta r_k).$$

Поэтому

$$\mathbf{P}\left\{\frac{1}{B_n^2}T_n(\theta) < 0\right\} = \mathbf{P}\{T_n(\theta) < 0\} \sim \Phi\left(-\frac{\mathbf{E}T_n(\theta)}{\sqrt{\mathbf{D}T_n(\theta)}}\right).$$

Остается убедиться, что

$$-\sum_{k=1}^n r_k p_k / \sqrt{\sum_{k=1}^n r_k^2(1 - p_k^2)} \rightarrow \begin{cases} +\infty & \text{при } \theta < \theta^0, \\ -\infty & \text{при } \theta > \theta^0. \end{cases} \quad (1.3.1)$$

Достаточно рассмотреть какой-либо один случай, например, $\theta > \theta^0$.

Для этого придется принять некоторые дополнительные предположения о плане эксперимента и о свойствах случайных ошибок. Что касается плана эксперимента, т.е. регрессионных постоянных r_1, r_2, \dots, r_n , то применительно к обсуждаемому примеру уместно отметить, что существует постоянная $r_0 > 0$ такая, что при всех $k = 1, 2, \dots$

$$r_k \geq r_0 \quad (1.3.2)$$

для некоторого $r_0 > 0$.

В асимптотическом анализе обычно рассматривают одинаково распределенные ошибки. Это предположение уместно, если условия измерений в каком-то смысле однородны. В регрессионной модели (1.2.1) естественна и другая возможность: величина случайной

ошибки ξ_k возрастает вместе с ростом расстояния r_k . Все же увеличение ошибок не должно происходить слишком быстро; иначе измерения на больших дальностях могут оказаться неинформативными. Естественно предположить, что на больших дальностях масштаб ошибки растет не быстрее, чем сама дальность. При условии (1.3.2) это предположение можно выразить в виде неравенств

$$\begin{aligned} P\left\{\frac{\xi_k}{r_k} < x\right\} &\geq G(x) \quad \text{для } x > 0, \\ P\left\{\frac{\xi_k}{r_k} < x\right\} &\leq G(x) \quad \text{для } x < 0, \end{aligned} \quad (1.3.3)$$

где $G(x)$ — некоторая монотонно возрастающая непрерывная функция, причем $G(0) = 1/2$. Для одинаково распределенных ошибок условие (1.3.3) с учетом (1.3.2) означает, что их функция распределения монотонно возрастает в окрестности нуля.

Приняв (1.3.3), получаем, что для $\theta > \theta^0$

$$p_k(\theta) \leq 1 - 2G(\theta - \theta^0) < 0,$$

а потому

$$\frac{\sum_{k=1}^n r_k p_k}{\sqrt{\sum_{k=1}^n r_k^2 (1 - p_k^2)}} \leq \frac{1 - 2G(\theta - \theta^0)}{\sqrt{1 - (2G(\theta - \theta^0) - 1)^2}} \frac{\sum_{k=1}^n r_k}{\sqrt{\sum_{k=1}^n r_k^2}}.$$

Теперь для доказательства (1.3.1) достаточно, чтобы

$$\sum_{k=1}^n r_k / \sqrt{\sum_{k=1}^n r_k^2} \rightarrow \infty \quad \text{при } n \rightarrow \infty. \quad (1.3.4)$$

В разбираемом примере естественно предполагать, что число галактик в шаре радиуса R примерно пропорционально его объему, т.е. пропорционально R^3 . Если нумеровать наблюдения в порядке возрастания расстояния, для r_n получим зависимость $r_n = O(n^{1/3})$. При таком росте r_n (1.3.4) выполняется (как и при любой другой степенной зависимости $r_n = O(n^\alpha)$, $\alpha > 0$), что легко проверить. Если же расстояния возрастают с большей скоростью, то для доказательства (1.3.1) свойства обеих последовательностей $\{r_n\}$ и $\{p_n\}$ надо рассматривать одновременно. Из (1.3.4) следует (1.3.1), откуда, в свою очередь, следуют утверждения а) и б). Наконец, утверждение в) следует из неравенства Чебышева, если учесть, что $B_n^2 \rightarrow \infty$ при $n \rightarrow \infty$ в силу (1.2.8).

Полученные результаты сформулируем в виде теоремы.

Т е о р е м а 1.3.1 (состоятельность знаковой оценки (1.2.13)).
Предположим, что в задаче (1.2.1) план эксперимента (r_1, r_2, \dots, r_n) таков, что при $n \rightarrow \infty$

$$r_k \geq r_0 > 0, \quad \frac{\max_{1 \leq k \leq n} r_k^2}{\sum_{k=1}^n r_k^2} \rightarrow 0, \quad \frac{\sum_{k=1}^n r_k}{\sqrt{\sum_{k=1}^n r_k^2}} \rightarrow \infty.$$

Относительно распределения независимых случайных ошибок ξ_k , $k = 1, 2, \dots, n$ предположим, что существует непрерывная монотонная функция $G(x)$ такая, что для $r_k \geq r_0 > 0$

$$P\{\xi_k/r_k < x\} \geq G(x) \quad \text{при } x > 0,$$

$$P\{\xi_k/r_k < x\} \leq G(x) \quad \text{при } x < 0$$

и что $G(0) = 0.5$.

При этих условиях знаковая оценка (1.2.12) состоятельна.

1.3.2. Асимптотическая нормальность. Установив, что при некоторых естественных свойствах плана эксперимента и распределений случайных ошибок знаковая оценка состоятельна, т.е. сходится к истинному значению параметра, естественно выяснить, какова скорость этой сходимости. Хорошо известно, что в случае выборок из регулярных относительно параметра распределений оценки сходятся со скоростью $n^{-1/2}$. Для задач регрессии число наблюдений n уже не отражает количества накопленной информации о неизвестных параметрах. В задаче (1.2.1) вместо n надо рассматривать переменную

$$B_n^2 = \sum_{k=1}^n r_k^2. \quad (1.3.5)$$

При этом скорость сходимости равна B_n^{-1} при некоторых естественных условиях, как будет показано ниже. Говоря, что оценка $\hat{\theta}_n$ сходится к истинному значению параметра θ^0 со скоростью B_n^{-1} , мы подразумеваем, что случайная величина

$$B_n(\hat{\theta}_n - \theta^0) \quad (1.3.6)$$

имеет невырожденное предельное распределение при $n \rightarrow \infty$. Обычно в качестве предельного выступает нормальное распределение. Ниже мы укажем достаточные для этого условия. Ради простоты мы ограничимся одинаково распределенными ошибками.

Чтобы доказать, что (1.3.6) распределено асимптотически нормально, изучим более подробно функцию T_n^1 (1.2.9) в окрестности точ-

ки $\theta = \theta^0$. Введем переменную t , положив $\theta = \theta^0 + t/B_n$. Теперь

$$T_n(\theta) = \sum_{k=1}^n r_k \operatorname{sign} \left(-t \frac{r_k}{B_n} + \xi_k \right). \quad (1.3.7)$$

Введем также переменную N_n , положив

$$N_n := \sum_{i=1}^n r_i^2 / \max_{1 \leq i \leq n} r_i^2. \quad (1.3.8)$$

При условии (1.2.8) $N_n \rightarrow \infty$ при $n \rightarrow \infty$.

Рассмотрим задачу

$$\left[\sum_{k=1}^n \frac{r_k}{B_n} \operatorname{sign} \left(\xi_k - t \frac{r_k}{B_n} \right) \right]^2 \Rightarrow \min_{t \in \mathbb{R}^1}, \quad (1.3.9)$$

которая эквивалентна (1.2.12). В дальнейшем нам понадобится теорема 1.3.2, которая дает версию стохастического дифференцирования функции $T_n(\theta)$ — функции, напомним, не только не дифференцируемой, но и не непрерывной.

Теорема 1.3.2. *Предположим, что в задаче (1.2.1) план эксперимента (r_1, r_2, \dots, r_n) удовлетворяет условию (1.2.8), а случайные ошибки $\xi_1, \xi_2, \dots, \xi_n$ независимы, одинаково распределены и для их общей функции распределения $F(x)$ выполнены условия: $F(0) = 1/2$; $F'(x)$ существует и непрерывна в некоторой окрестности нуля.*

Тогда

$$\frac{1}{B_n} \sum_{k=1}^n r_k \operatorname{sign} \left(\xi_k - t \frac{r_k}{B_n} \right) = \frac{1}{B_n} \sum_{k=1}^n r_k \operatorname{sign} \xi_k - 2t F'(0) + X_n(t) \quad (1.3.10)$$

при $n \rightarrow \infty$, где $X_n(t) \rightarrow 0$ (по вероятности) равномерно относительно t , если $|t| < N_n^\alpha$, $0 < \alpha < 1/4$, т.е. для любого $\varepsilon > 0$

$$\mathbf{P} \left\{ \sup_{|t| < N_n^\alpha} |X_n(t)| > \varepsilon \right\} \rightarrow 0. \quad (1.3.11)$$

Доказательство. Согласно (1.3.10)

$$X_n(t) = \sum_{k=1}^n \frac{r_k}{B_n} \left[\operatorname{sign} \left(\xi_k - t \frac{r_k}{B_n} \right) - \operatorname{sign} \xi_k + 2F'(0) t \frac{r_k}{B_n} \right]. \quad (1.3.12)$$

Положим

$$Y_n(t) = \sum_{k=1}^n \frac{r_k}{B_n} \left[\text{sign} \left(\xi_k - t \frac{r_k}{B_n} \right) - \text{sign} \xi_k \right]. \quad (1.3.13)$$

Отрезок $\{t: |t| \leq N_n^\alpha\}$ разобьем на отрезки одинаковой длины $h(n) = N_n^{-\beta}$, $0 < \beta < 1/4$. Множество этих отрезков обозначим Ω ; ω будет далее обозначать произвольный элемент из Ω , а τ — точку центра отрезка ω . Для доказательства (1.3.11) достаточно показать, что

$$\sum_{\omega \in \Omega} \mathbf{P} \left\{ \sup_{t \in \omega} |X_n(t)| > \varepsilon \right\} \rightarrow 0. \quad (1.3.14)$$

Число слагаемых в (1.3.14) равно $2N_n^{\alpha+\beta}$. Мы докажем (1.3.11), если для вероятностей

$$\mathbf{P} \left\{ \sup_{t \in \omega} |X_n(t)| > \varepsilon \right\}$$

сумеем получить равномерную оценку сверху вида $o(N_n^{\alpha+\beta})$.

Напомним, что *колебанием функции* $f(x)$ на множестве \mathbb{M} называют

$$W(f, \mathbb{M}) := \sup_{x \in \mathbb{M}} f(x) - \inf_{x \in \mathbb{M}} f(x)$$

и что

$$W(f+g, \mathbb{M}) \leq W(f, \mathbb{M}) + W(g, \mathbb{M}).$$

Теперь

$$\begin{aligned} \mathbf{P} \left\{ \sup_{t \in \omega} |X_n(t)| > \varepsilon \right\} &\leq \mathbf{P} \{ W(X_n, \omega) + |X_n(\tau)| > \varepsilon \} \leq \\ &\leq \mathbf{P} \{ W(X_n, \omega) > \varepsilon/2 \} + \mathbf{P} \{ |X_n(\tau)| > \varepsilon/2 \} \leq \\ &\leq \mathbf{P} \{ W(Y_n, \omega) + 2F'(0)h(n) > \varepsilon/2 \} + \mathbf{P} \{ |X_n(\tau)| > \varepsilon/2 \} \leq \\ &\leq \mathbf{P} \{ W(Y_n, \omega) > \varepsilon/3 \} + \mathbf{P} \{ |X_n(\tau)| > \varepsilon/2 \}, \quad (1.3.15) \end{aligned}$$

так как $h(n) \rightarrow 0$.

Каждую из двух последних вероятностей мы оценим сверху с помощью неравенства Чебышева, для чего нам понадобятся первые два момента $Y_n(t)$ и $X_n(t)$ и их оценки. Начнем со второй вероятности.

Пусть Δ , $\Delta \in \mathbb{R}^1$, принадлежит той упомянутой в условиях теоремы окрестности нуля, в которой $F'(x)$ существует и непрерывна. Заметим, что

$$\text{sign}(\xi_k - \Delta) - \text{sign} \xi_k = \begin{cases} 0, & \text{если } \xi_k > 0, \xi_k > \Delta, \\ 2, & \text{если } \xi_k < 0, \xi_k > \Delta, \\ -2, & \text{если } \xi_k > 0, \xi_k < \Delta, \\ 0, & \text{если } \xi_k < 0, \xi_k < \Delta, \end{cases}$$

а прочие возможности имеют вероятность, равную нулю. Поэтому

$$\begin{aligned} \mathbf{E}\{\text{sign}(\xi_k - \Delta) - \text{sign} \xi_k\} &= \\ &= 2\mathbf{P}\{0 < \xi_k < \Delta\} - 2\mathbf{P}\{\Delta < \xi_k < 0\} = -2F'(\kappa)|\Delta|, \end{aligned}$$

где κ — некоторая точка между 0 и Δ . Заметим также, что если $|t| < N_n^\alpha$, то

$$\max_{1 \leq k \leq n} \left| t \frac{r_k}{B_n} \right| \leq N_n^{-(1-\alpha)} \rightarrow 0.$$

Поэтому при $|t| < N_n^\alpha$

$$|\mathbf{E}X_n(t)| \leq 2N_n^\alpha \sum_{k=1}^n \frac{r_k^2}{B_n^2} |F'(0) - F'(\kappa_k)|,$$

где κ_k — некоторая точка между 0 и tr_k/B_n , $k = 1, \dots, n$. При этом

$$\max_{1 \leq k \leq n} |\kappa_k| \leq N_n^{-(1-\alpha)} \rightarrow 0,$$

так что $|F'(0) - F'(\kappa_k)| \leq \text{const} |\kappa_k|$ для всех $k = 1, \dots, n$. Поэтому

$$\sup_{|t| < N_n^\alpha} |\mathbf{E}X_n(t)| \leq \text{const} \cdot N_n^{-(1-2\alpha)}. \quad (1.3.16)$$

Аналогично получаем, что

$$\sup_{|t| < N_n^\alpha} \mathbf{D}X_n(t) \leq \text{const} \cdot N_n^{-(1-\alpha)}.$$

Теперь в силу (1.3.16) для достаточно больших n получаем

$$\begin{aligned} \mathbf{P}\{|X_n(\tau)| > \varepsilon/2\} &\leq \mathbf{P}\{|X_n(\tau) - \mathbf{E}X_n(\tau)| + |\mathbf{E}X_n(\tau)| > \varepsilon/2\} \leq \\ &\leq \mathbf{P}\{|X_n(\tau) - \mathbf{E}X_n(\tau)| > \varepsilon/3\} \leq (\varepsilon/3)^{-2} \mathbf{D}X_n(\tau) \leq \\ &\leq \text{const} \cdot N_n^{-(1-\alpha)}. \end{aligned} \quad (1.3.17)$$

Обратимся теперь к $\mathbf{E}W(Y_n, \omega)$ и $\mathbf{D}W(Y_n, \omega)$. Так как функции $\frac{r_k}{B_n} \text{sign}(\xi_k - t \frac{r_k}{B_n})$ монотонно не возрастают по t , то

$$W(Y_n, \omega) = \sum_{k=1}^n \frac{r_k}{B_n} \left[\text{sign}\left(\xi_k - \mu \frac{r_k}{B_n}\right) - \text{sign}\left(\xi_k - \lambda \frac{r_k}{B_n}\right) \right],$$

если $\omega = [\lambda, \mu]$. Рассуждая, как выше, находим, что

$$\mathbf{E}(Y_n, \omega) \leq \text{const} \cdot h(n),$$

$$D(Y_n, \omega) \leq \text{const} \cdot N_n^{-1} h(n)$$

для любого отрезка длины h , принадлежащего множеству $\{t: |t| < N_n^\alpha\}$.
Поэтому

$$\begin{aligned} P\left\{W(Y_n, \omega) > \frac{\varepsilon}{3}\right\} &\leq P\left\{|W(Y_n, \omega) - EW(Y_n, \omega)| + EW(Y_n, \omega) > \frac{\varepsilon}{3}\right\} \leq \\ &\leq P\left\{|W(Y_n, \omega) - EW(Y_n, \omega)| > \frac{\varepsilon}{4}\right\} \leq \\ &\leq \left(\frac{\varepsilon}{4}\right)^{-2} DW(Y_n, \omega) \leq \text{const} \cdot N_n^{-1} h(n). \quad (1.3.18) \end{aligned}$$

Обратившись к (1.3.14) и учтя, что число слагаемых там равно $2N_n^\alpha/N_n^{-\beta} = 2N_n^{\alpha+\beta}$, находим, используя (1.3.15), (1.3.17) и (1.3.18), что

$$\begin{aligned} P\left\{\sup_{|t| < N_n^\alpha} |X_n(t)| > \varepsilon\right\} &\leq \\ &\leq \text{const} \cdot N_n^{\alpha+\beta} (N_n^{-1} N_n^{-\beta} + N_n^{-1+\alpha}) \leq \text{const} \cdot N_n^{-1+2\alpha+\beta} \rightarrow 0, \end{aligned}$$

если $0 < \alpha < 1/4$, $0 < \beta < 1/4$. Теорема 1.3.2 доказана. \square

Теперь мы можем изучить асимптотическое распределение знаковой оценки $\hat{\theta}_n$ (1.3.9) в модели (1.2.1).

Теорема 1.3.3. *Предположим, что выполнены условия теоремы 1.3.2 и что $F'(0) > 0$.*

Тогда

$$\hat{\theta}_n = \theta^0 + \frac{1}{2F'(0)B_n^2} \sum_{k=1}^n r_k \text{sign } \xi_k + o_p\left(\frac{1}{B_n}\right),$$

а потому асимптотически

$$\hat{\theta}_n \sim N\left(\theta^0, \frac{1}{(2F'(0)B_n)^2}\right).$$

Доказательство. Выберем α , $0 < \alpha < 1/4$, как в теореме 1.3.2. Область изменения t в (1.3.9) разделим на три части

$$\mathbb{A} = \{t \leq -N_n^\alpha\}, \quad \mathbb{B} = \{t \geq N_n^\alpha\}, \quad \mathbb{C} = \{|t| < N_n^\alpha\}.$$

Мы покажем, что в областях \mathbb{A} и \mathbb{B} целевая функция в (1.3.9) неограниченно возрастает (по вероятности) при $n \rightarrow \infty$. В то же время в области \mathbb{C} при $t = 0$ целевая функция по вероятности ограничена. Следовательно, при достаточно больших n точка глобального минимума в (1.3.9) находится в области \mathbb{C} со сколь угодно близкой к 1 вероятностью. Поэтому при решении (1.3.9) на t можно наложить ограничение $|t| < N_n^\alpha$.

Исследуем область A . Напомним, что (1.3.10) монотонно убывает по t . Поэтому при $t \leq -N_n^\alpha$

$$\sum_{k=1}^n \frac{r_k}{B_n} \operatorname{sign} \left(\xi_k - t \frac{r_k}{B_n} \right) \geq \sum_{k=1}^n \frac{r_k}{B_n} \operatorname{sign} \left(\xi_k - N_n^\alpha \frac{r_k}{B_n} \right).$$

Обозначим правую часть через ε_n . Как легко видеть,

$$E\varepsilon_n = -2 N_n^\alpha \sum_{k=1}^n \frac{r_k^2}{B_n^2} F'(\kappa_k),$$

где κ_k — некоторая точка из $(-N_n^\alpha r_k/B_n, 0)$, причем $\max_{1 \leq k \leq n} |\kappa_k| \rightarrow 0$. Поэтому $F'(\kappa_k) \rightarrow F'(0)$ равномерно по k , причем $F'(0) > 0$ по условию. Следовательно, $E\varepsilon_n \rightarrow -\infty$. Так как дисперсия ε_n , очевидно, ограничена, то $\varepsilon_n \xrightarrow{P} -\infty$. Поэтому абсолютное значение (1.3.10), а также левая часть (1.3.9) неограниченно возрастают в области A .

Для области B справедлив аналогичный результат. А именно: минимальное значение левой части (1.3.9) стремится к $+\infty$ по вероятности.

Обратимся к области C и рассмотрим (1.3.9) с ограничением $|t| < N_n^\alpha$. По теореме 1.3.2 эту задачу можно представить в виде

$$\left(\frac{1}{B_n} \sum_{k=1}^n r_k \operatorname{sign} \xi_k - 2F'(0)t + X_n(t) \right)^2 \Rightarrow \min_{t: |t| < N_n^\alpha}, \quad (1.3.19)$$

причем $X_n(t) \xrightarrow{P} 0$ равномерно по t на указанном множестве. Благодаря этому (1.3.19) эквивалентно

$$\left(\frac{1}{B_n} \sum_{k=1}^n r_k \operatorname{sign} \xi_k - 2F'(0)t \right)^2 \Rightarrow \min_{t: |t| < N_n^\alpha} \quad (1.3.20)$$

в том смысле, что разность между решениями (1.3.19) и (1.3.20) сходится по вероятности к нулю. Если же не учитывать ограничения, то минимальное значение в (1.3.20) равно нулю и достигается при

$$\hat{t}_n = \frac{1}{2F'(0)} \sum_{k=1}^n \frac{r_k}{B_n} \operatorname{sign} \xi_k.$$

Заметим, что \hat{t}_n асимптотически распределено по закону $N\left(0, \frac{1}{(2F'(0))^2}\right)$. Поэтому

$$P\{|\hat{t}_n| < N_n^\alpha\} \rightarrow 1.$$

Следовательно, \hat{t}_n при достаточно больших n и с вероятностью, сколь угодно близкой к 1, служит также и решением (1.3.20). Таким образом, \hat{t}_n эквивалентно решению (1.3.19). Теорема доказана. \square

§ 1.4. Функция влияния

Функция влияния описывает количественный эффект, который производит отдельное наблюдение на статистическую оценку. Функция влияния как характеристика статистического правила входит в его описание наряду с такими понятиями как смещение или дисперсия. Подобно им функция влияния существует как в конечной, так и в асимптотической форме. Функциям влияния и их роли в асимптотической теории статистики посвящена превосходная книга Ф.Хампеля и соавторов [36, 73], где эти вопросы изучены для повторных выборок. В эту схему укладываются и линейные модели со случайными факторами. Для линейных моделей с неслучайными факторами, которые мы обсуждаем в этой книге, такие работы нам не известны, хотя авторы и говорят ([73, § 1.3, e, с. 47]), что для построения аналогичной теории для общих линейных моделей нужны лишь небольшие изменения.

Вместо функции влияния мы рассмотрим близкое к ней понятие функции (кривой) чувствительности (sensitivity curve), предложенное Дж.Тьюки [27, 108] (см. также [73, § 2.1, e]).

Предположим, что к наблюдениям (r_i, v_i) , $i = 1, \dots, n$, в модели $v_i = r_i\theta^0 + \xi_i$, $i = 1, \dots, n$, мы присоединили наблюдение (r, v) , где $v = r\theta^0 + \varepsilon$. Пусть оценка $\hat{\theta}_n$ получена по наблюдениям (r_i, v_i) , $i = 1, \dots, n$, а оценка $\hat{\theta}_{n+1}^*$ — по тем же наблюдениям с присоединением искусственного, неслучайного наблюдения (r, v) . Несколько изменяя определение Тьюки (с учетом того, что в нашей схеме B_n^2 замещает n), определим функцию чувствительности для точки (r, v) как

$$SC_n(r, \varepsilon) = B_n^2(\hat{\theta}_{n+1}^* - \hat{\theta}_n).$$

Мы покажем ниже (см. теорему 1.4.1), что при $n \rightarrow \infty$

$$SC_n(r, \varepsilon) = \frac{r \operatorname{sign} \varepsilon}{2F'(0)} + o_p(1). \quad (1.4.1)$$

Ясно, что предел по вероятности этой функции ограничен по переменной ε . Тем самым влияние отдельной, даже сколь угодно большой ошибки измерения отклика остается ограниченным. Впрочем, влияние ошибки в измерении абсциссы наблюдения может быть большим, если велика сама эта ошибка. Последнее обстоятельство хоть и неприятно, но не удивительно: мы изучаем задачу в классическом предположении линейного анализа, что абсциссы наблюдений измеряются вообще без ошибок.

Для сравнения приведем функцию чувствительности в той же задаче для оценки наименьших квадратов. Она равна $r\varepsilon + o_p(1)$ при

$n \rightarrow \infty$, если предполагать ошибки ξ_i независимыми, одинаково распределенными, причем $E\xi_i = 0$, $D\xi_i < \infty$. Этот вывод получается прямым вычислением. Видно, что здесь функция чувствительности неограниченно растет вместе с ошибкой ε . Это отражает хорошо известный факт, что даже одна грубая ошибка в отклике способна исказить результат оценивания, если применяется метод наименьших квадратов.

Т е о р е м а 1.4.1. *Предположим, что план эксперимента (r_1, r_2, \dots, r_n) в задаче (1.2.1) удовлетворяет условию (1.2.6). Предположим, что случайные ошибки $\xi_1, \xi_2, \dots, \xi_n$ независимы, одинаково распределены и их общая функция распределения $F(x)$ имеет непрерывную производную $F'(x) > 0$ в окрестности нуля и $F(0) = 1/2$.*

Тогда для $\varepsilon \neq 0$

$$SC_n(r, \varepsilon) \xrightarrow{P} \frac{r \operatorname{sign} \varepsilon}{2F'(0)}$$

при $n \rightarrow \infty$.

Д о к а з а т е л ь с т в о. Знаковые оценки $\hat{\theta}_n$ и $\hat{\theta}_{n+1}^*$ определяются как решения экстремальных задач

$$\left| \sum_{i=1}^n r_i \operatorname{sign}(v_i - r_i \theta) \right| \Rightarrow \min_{\theta} \quad (1.4.2)$$

$$\left| \sum_{i=1}^n r_i \operatorname{sign}(v_i - r_i \theta) + r \operatorname{sign}(v - r \theta) \right| \Rightarrow \min_{\theta} \quad (1.4.3)$$

Исследование (1.4.3) при $n \rightarrow \infty$ проводится так же, как ранее было проведено исследование оценки $\hat{\theta}_n$ (1.2.12). При этом мы сохраняем прежние обозначения. Вводим переменную $t = B_n(\theta - \theta^0)$. Решение (1.4.3) обозначим через \hat{t}_{n+1} , $\hat{t}_{n+1} = B_n(\hat{\theta}_{n+1}^* - \theta^0)$. Переписываем (1.4.3) в виде

$$\left| \frac{1}{B_n} \sum_{i=1}^n r_i \operatorname{sign} \left(\xi_i - t \frac{r_i}{B_n} \right) + \frac{r}{B_n} \operatorname{sign} \left(\varepsilon - t \frac{r}{B_n} \right) \right| \Rightarrow \min_t \quad (1.4.4)$$

Так же, как в теореме 1.3.3, показываем, что со сколь угодно близкой к 1 вероятностью решение (1.4.4) в условиях этой теоремы принадлежит области $C = \{t: |t| \leq \operatorname{const} \cdot N_n^\alpha\}$ при достаточно больших n . В этой области действует разложение (1.3.10). Кроме того, в области C при достаточно больших n

$$\operatorname{sign} \left(\varepsilon - t \frac{r}{B_n} \right) = \operatorname{sign} \varepsilon.$$

Поэтому в области S задача (1.4.4) эквивалентна задаче

$$\left| \frac{1}{B_n} \sum_{i=1}^n r_i \operatorname{sign} \xi_i + \frac{r}{B_n} \operatorname{sign} \varepsilon - 2F'(0)t + X_n(t) \right| \Rightarrow \min_{t: |t| \leq \operatorname{const} \cdot N_n^\alpha} \cdot \quad (1.4.5)$$

Те же рассуждения, что в теореме 1.3.3, показывают, что разность между решениями задач (1.4.4) и (1.4.5) по вероятности стремится к нулю и что поэтому

$$\hat{t}_{n+1} = \frac{1}{2F'(0)} \left(\frac{1}{B_n} \sum_{i=1}^n r_i \operatorname{sign} \xi_i + \frac{r}{B_n} \operatorname{sign} \varepsilon \right) + o_p(1),$$

или $\hat{t}_{n+1} = \hat{t}_n + \frac{1}{2F'(0)} \frac{r \operatorname{sign} \varepsilon}{B_n} + o_p(1)$. Отсюда следует (1.4.1). \square

ЗНАКОВЫЕ КРИТЕРИИ

§ 2.1. Общая линейная модель

Линейные модели можно встретить в нескольких областях математической статистики. К такому названию прибегают, когда математическая модель для наблюдений строится с помощью линейных операций. Несомненно, наиболее известны линейные модели регрессионного и факторного анализа, и зачастую только их и имеют в виду, когда употребляют этот термин. Однако не менее важную роль играют линейные модели и в анализе временных рядов. Это модели авторегрессии, скользящего среднего и их комбинации. Различаются эти два вида линейных моделей, например, тем, что в упомянутых первых регрессионных и факторных моделях наблюдения независимы как случайные величины, в то время как для временных рядов характерна именно взаимная зависимость наблюдаемых элементов случайных последовательностей. В первой части книги мы рассматриваем применение непараметрических (в основном знаковых) методов именно к линейным моделям независимых наблюдений. Во второй части те же идеи и методы будут развиты для линейных моделей временных рядов (для моделей авторегрессии). Мы будем заниматься только моделями со скалярным откликом. Это означает, что при каждой фиксированной комбинации значений факторов мы регистрируем и затем изучаем лишь одну зависящую от них переменную (отклик). Без сомнений, это серьезно сужает область приложений, ибо во многих случаях надо следить сразу за несколькими характеристиками и их взаимодействием. К сожалению, многооткликовый (многомерный) линейный анализ достаточно разработана лишь в гауссовском случае. Непараметрические методы для многомерного анализа пока мало изучены.

Дадим геометрическую формулировку линейной модели. Пусть X — вектор, состоящий из конечного числа наблюдений. Мы скажем, что эти наблюдения образуют линейную модель, если X можно представить в виде

$$X = l + \xi, \quad (2.1.1)$$

где ξ — случайный вектор с независимыми случайными координатами, l — неслучайный вектор. Векторы l и ξ недоступны наблюдению и неизвестны наблюдателю, однако относительно вектора l мы будем предполагать, что он заведомо принадлежит заданному линейному

подпространству \mathbb{L} , т.е. $l \in \mathbb{L}$. Разнообразие линейных моделей связано с различными формами представления X , различными способами задания \mathbb{L} и различными предположениями о свойствах случайных ошибок. В теоретических исследованиях удобно считать X n -мерным (по числу наблюдений) вектором-столбцом $X = (x_1, \dots, x_n)^T$, а подпространство \mathbb{L} задавать с помощью базиса. Пусть $r = \dim \mathbb{L}$, причем $r < n$, и векторы (столбцы) C_1, \dots, C_r образуют базис в \mathbb{L} . Пусть

$$C_\alpha = (c_{1\alpha}, c_{2\alpha}, \dots, c_{n\alpha})^T, \quad \alpha = 1, \dots, r,$$

и C есть матрица со столбцами C_α , т.е.

$$C = \|c_{i\alpha}\|, \quad i = 1, \dots, n; \quad \alpha = 1, \dots, r.$$

Пусть

$$\theta = (\theta_1, \dots, \theta_r)^T \quad (2.1.2)$$

— вектор неизвестных параметров, а

$$\xi = (\xi_1, \dots, \xi_n)^T \quad (2.1.3)$$

— вектор случайных ошибок. Эти случайные ошибки мы будем считать независимыми.

Для знакового метода основное предположение об ошибках, помимо независимости, касается их медиан. Мы будем предполагать, что все случайные ошибки имеют медианы, равные нулю. Например, в модели (2.1.1) это означает, что

$$P\{\xi_i < 0\} = P\{\xi_i > 0\} = 1/2 \quad (2.1.4)$$

для всех $i = 1, \dots, n$. Именно при этом предположении возможен знаковый анализ. Мы не требуем от ошибок одинаковой распределенности, хотя в ряде случаев вводим это свойство — для построения оптимальных правил и для их асимптотического анализа. О других, более традиционных предположениях о (2.1.3) мы скажем несколько слов ниже.

После введения таких обозначений (2.1.1) предстает в стандартной для задач регрессии форме:

$$X = C\theta + \xi. \quad (2.1.5)$$

В задачах одно- и двухфакторного анализа естественно расположить массив наблюдений в виде таблицы с одним либо двумя входами. В однофакторной задаче (задаче о нескольких выборках) единственный фактор принимает k различных значений (уровней). Эти уровни нумеруют целыми числами от 1 до k . Для каждого уровня $j = 1, \dots, k$ мы имеем n_j независимых и одинаково распределенных наблюдений x_{ij} , $i = 1, \dots, n_j$. В этой задаче всю совокупность наблюдений X можно представлять себе как таблицу, состоящую из k столбцов длины

n_1, n_2, \dots, n_k . Линейную модель здесь проще задать в координатной форме, предположив, что

$$x_{ij} = \mu + \tau_j + \xi_{ij} \quad (2.1.6)$$

для $i = 1, \dots, n_j$ и $j = 1, \dots, k$, где ξ_{ij} — независимые случайные величины, отражающие ту часть изменчивости отклика, которая не обусловлена действием выделенного фактора. Величина μ выражает некий средний уровень, от которого отклоняются наблюдаемые значения x_{ij} (на величину τ_j под действием j -го уровня фактора, на величину ξ_{ij} по случайным причинам). Чтобы модель (2.1.6) была идентифицируема, и по самому смыслу переменных τ_j на них следует наложить какую-либо линейную связь. Например, потребовать, чтобы выполнялось

$$\sum_{j=1}^k \tau_j = 0. \quad (2.1.7)$$

С геометрической точки зрения подпространство \mathbb{L} , размерность которого равна k , задано в этой задаче $k+1$ векторами — по числу параметров $\mu, \tau_1, \dots, \tau_k$. Эти $k+1$ векторов в k -мерном подпространстве уже не являются базисом. Каждому параметру соответствует вектор из этого набора. Так, параметру τ_j соответствует таблица упомянутой ранее формы, где столбец j состоит из единиц, а прочие элементы равны нулю. Параметру μ соответствует таблица из сплошных единиц — это сумма векторов (таблиц), отвечающих τ_1, \dots, τ_k . Условие (2.1.7) позволяет для каждого вектора $l \in \mathbb{L}$ дать однозначное выражение в виде линейной комбинации упомянутых порождающих векторов. В дальнейшем мы встретим и другие примеры задания \mathbb{L} .

Остается обсудить предположения, которые обычно принимают в отношении случайных ошибок, представленных в нашей модели вектором ξ . Собственно, наиболее важное предположение уже было высказано: составляющие ξ случайные величины считаются независимыми. В классическом (гауссовском) анализе они считаются нормально распределенными с нулевым средним и дисперсией σ^2 (или, кратко, по закону $N(0, \sigma^2)$). В этих условиях действует хорошо известный метод наименьших квадратов. Рекомендации гауссовской теории ошибок мы будем считать известными. Они послужат нам хорошим эталоном для сравнения. Менее ограничительна непараметрическая постановка, согласно которой случайные ошибки распределены одинаково, но этот общий закон распределения, предполагаемый непрерывным, неизвестен. При таких свойствах ошибок возможен анализ, использующий ранги наблюдений, и о нем ниже кое-что будет сказано. Для знакового метода непрерывность распределения нужна только в нуле, а одинаковая распределенность и вовсе не обязательна.

В заключение приведем сводку различных предположений относительно распределений случайных ошибок и планов эксперимента, которые встретятся в дальнейшем тексте. Подчеркнем, что мы не требуем их одновременного выполнения. Эти свойства мы будем привлекать по мере необходимости.

Условие 2.1.1. $F(0) = 0.5$.

Условие 2.1.2. $F(x)$ имеет непрерывную плотность $f(x)$ в некоторой окрестности нуля и $f(x) > 0$.

Условие 2.1.3. $f(x)$ абсолютно непрерывна в некоторой окрестности нуля и $f'(0) > 0$.

Условие 2.1.4. $F(x)$ удовлетворяет условию Липшица в нуле, т.е. существует такая константа $L > 0$, что

$$|F(u_1) - F(u_2)| < L|u_1 - u_2|.$$

Условие 2.1.5. Существуют константа $M > 0$ и некоторая окрестность нуля $\{u: |u| < d\}$ такие, что

$$|F(u_1) - F(u_2)| > M|u_1 - u_2|, \text{ если } |u_1| < d, |u_2| < d.$$

Условие 2.1.6. В некоторой окрестности нуля распределение $F(\cdot)$ имеет плотность $f(\cdot)$, удовлетворяющую условию Липшица, т.е. для некоторых $N > 0, d > 0$

$$|f(u_1) - f(u_2)| < N|u_1 - u_2|, \text{ если } |u_1| < d, |u_2| < d.$$

Условие 2.1.7. $\max_{1 \leq i \leq n} \max_{1 \leq \alpha \leq r} |c_{i\alpha}| < K < \infty$.

Условие 2.1.8. $\frac{1}{n} C^T C \rightarrow \Sigma > 0$.

§ 2.2. Локально оптимальные знаковые критерии в задаче регрессии

В этом параграфе мы рассмотрим задачу проверки гипотезы

$$H: l = 0 \tag{2.2.1}$$

в линейных моделях (2.1.1). Мы уже отмечали во введении, что с практической точки зрения такая гипотеза не часто бывает важной, но критериальные статистики для ее проверки необходимы для всего последующего — для оценивания параметров и для проверки линейных гипотез. Чтобы вывести разумные (и даже в определенном смысле наилучшие) критерии для проверки (2.2.1), примем некоторые предположения о распределении ошибок в модели (2.1.1), помимо основного предположения (2.1.4). Мы подробно рассмотрим задачу для моделей в регрессионной форме (2.1.5). Для факторных моделей получим решения, сделав необходимые изменения в результатах, относящихся к (2.1.5).

Итак, пусть наблюдения $X = (x_1, \dots, x_n)^T$ удовлетворяют модели регрессии (2.1.5) с заданной матрицей C , называемой матрицей

плана эксперимента, вектором θ неизвестных параметров (2.1.2) и вектором ξ взаимно независимых случайных ошибок (2.1.3). Будем предполагать в дальнейшем, что столбцы матрицы C линейно независимы, т.е. что матрица C имеет полный ранг. Если через c_i обозначить строку матрицы C : $c_i = (c_{i1}, \dots, c_{ir})$, то для x_i получим выражение $x_i = c_i\theta + \xi_i$, $i = 1, \dots, n$, где $c_i\theta$ есть скалярное произведение: $c_i\theta = \sum_{\alpha=1}^r c_{i\alpha}\theta_{\alpha}$. Гипотеза (2.2.1) в этой модели превращается в

$$H_0: \theta = 0. \quad (2.2.2)$$

Альтернативы к H_0 имеют вид $\theta \neq 0$. Построим статистические критерии для проверки H_0 , которые при выполнении (2.1.4) сохраняют заданный уровень значимости. Будем использовать знаковые статистические правила, т.е. правила, основанные только на знаках наблюдений x_1, \dots, x_n .

Мы построим знаковые критерии для проверки (2.2.1), оптимальные в некотором локальном смысле, разъясняемом ниже. Рассмотрим статистику знаков

$$S(X) = (\text{sign } x_1, \text{sign } x_2, \dots, \text{sign } x_n)^T. \quad (2.2.3)$$

Значениями случайного вектора $S(X)$ могут быть лишь последовательности, составленные из $+1$ и -1 . Пусть

$$s = (s_1, s_2, \dots, s_n)^T \quad (2.2.4)$$

— такая произвольная последовательность. Обозначим через \mathcal{Q} некоторую совокупность векторов типа (2.2.4), которую мы в дальнейшем будем использовать как критерий для проверки H_0 (2.2.2). Иными словами, мы предлагаем отвергнуть (2.2.2), если происходит событие

$$\{X: S(X) \in \mathcal{Q}\}. \quad (2.2.5)$$

Рассмотрим мощность критерия (2.2.5) в модели (2.1.5) как функцию параметра θ :

$$P\{S(X) \in \mathcal{Q} | \theta\}, \quad (2.2.6)$$

или, короче, $P\{\mathcal{Q} | \theta\}$. Заметим, что

$$P\{\mathcal{Q} | \theta\} = \sum_{s \in \mathcal{Q}} P\{s | \theta\},$$

где

$$P\{s | \theta\} = P\{S(X) = s | \theta\}.$$

Требования к выбору наилучшего знакового критерия (2.2.5) сформулируем в терминах, относящихся к функции мощности. Во-первых, мы будем рассматривать и сравнивать между собой знаковые критерии заданного уровня значимости, скажем $\epsilon > 0$. Поэтому

первое требование к критерию \mathbb{Q} есть

$$P\{\mathbb{Q}|\theta\}|_{\theta=0} = \varepsilon. \quad (2.2.7)$$

Поскольку при H_0 и (2.1.4) вероятность любой последовательности (2.2.4) равна 2^{-n} , в (2.2.5) могут быть реализованы лишь уровни значимости, кратные 2^{-n} . При таком уровне ε любой знаковый критерий \mathbb{Q} будет содержать $K = \varepsilon \cdot 2^n$ точек вида (2.2.4). Наилучшим среди знаковых критериев данного уровня был бы тот, функция мощности которого превосходила бы функции мощности прочих знаковых критериев этого уровня. К сожалению, для рассматриваемых нами альтернатив $H: \theta \neq 0$ такого равномерно наиболее мощного критерия не существует. Поэтому при выборе критерия мы ограничимся рассмотрением его функции мощности лишь в окрестности точки $\theta = 0$.

Начнем с однопараметрической модели

$$x_i = c_i\theta + \xi_i, \quad i = 1, \dots, n,$$

где параметр $\theta \in \mathbb{R}^1$. Такую модель мы уже рассматривали в гл. 1 как модель (1.2.1). В однопараметрической модели для проверки (2.2.2) можно построить локально наиболее мощные знаковые критерии против односторонних альтернатив $H^+: \theta > 0$ или $H^-: \theta < 0$. Затем для проверки (2.2.2) против двусторонних альтернатив можно использовать их двусторонние версии. Так статистики поступают и в нормальной теории, когда применяют двусторонние версии критерия Стьюдента; так теоретики выводят ранговые критерии.

Предположим, что случайные ошибки независимы и одинаково распределены. Относительно их общей функции распределения

$$F(u) = P\{\xi_i < u\}$$

мы предположим, что выполнены условия 2.1.1, 2.1.2.

Для определенности рассмотрим альтернативу $H^+: \theta > 0$. Для однопараметрической модели второе требование к критерию \mathbb{Q} против этой альтернативы состоит в том, что локально около $\theta = 0$ его функция мощности должна возрастать с наибольшей возможной скоростью. Иначе говоря,

$$\frac{d}{d\theta} P\{S(X) \in \mathbb{Q} | 0\}$$

должна быть максимальной при условии (2.2.7). Так как

$$\frac{d}{d\theta} P\{S(X) \in \mathbb{Q} | \theta\} = \sum_{s \in \mathbb{Q}} \frac{d}{d\theta} P\{S(X) = s | \theta\},$$

то множество \mathbb{Q} надо составить из таких последовательностей s , у

которых самые большие значения $\frac{d}{d\theta} P\{S(X) = s|0\}$. Поэтому

$$Q = \left\{ s: \frac{d}{d\theta} P\{S(X) = s|0\} \geq \text{const} \right\},$$

причем постоянную надо выбрать так, чтобы удовлетворить условию (2.2.7).

Заметим, что распределение критериальной статистики может иметь атомы, вероятности которых больше, чем 2^{-n} (в целое число раз). В этом случае выбранный уровень ε (даже кратный 2^{-n}) может оказаться недостижимым. В этой ситуации можно либо прибегнуть к рандомизации либо включить в критическую область ровно столько точек из множества

$$\left\{ s: \frac{d}{d\theta} P\{S(X) = s|0\} = \text{const} \right\},$$

сколько необходимо. Выбор этих точек не влияет на локальные свойства функции мощности критерия.

Остается вычислить $\frac{d}{d\theta} P\{S(X) = s|0\}$. Для правдоподобия $P\{S(X) = s|\theta\}$, где $S(X)$ и s заданы (2.2.3) и (2.2.4), имеем выражение

$$P\{S(X) = s|\theta\} = \prod_{i=1}^n (P\{x_i > 0\})^{\frac{1+s_i}{2}} (P\{x_i < 0\})^{\frac{1-s_i}{2}}. \quad (2.2.8)$$

Легко видеть, что при оговоренных условиях на функцию $F(u)$

$$\begin{aligned} P\{S(X) = s|\theta\} &= \prod_{i=1}^n (P\{x_i > 0\})^{\frac{1+s_i}{2}} (P\{x_i < 0\})^{\frac{1-s_i}{2}} = \\ &= 2^{-n} \prod_{i=1}^n [1 + 2f(0)c_i s_i \theta + o(\theta)] = \\ &= 2^{-n} \left[1 + 2f(0) \left(\sum_{i=1}^n c_i s_i \right) \theta + o(\theta) \right]. \end{aligned}$$

Поэтому

$$\frac{d}{d\theta} P\{S(X) = s|0\} = 2^{-(n-1)} f(0) \sum_{i=1}^n c_i s_i.$$

Следовательно, для любого уровня значимости искомым критерий имеет вид

$$Q = \left\{ s: \sum_{i=1}^n c_i s_i \geq \text{const} \right\}, \quad (2.2.9)$$

где постоянную надо выбрать так, чтобы выполнялось условие (2.2.7). Ясно, что против $H^-: \theta < 0$ локально наиболее мощный знаковый критерий имеет вид

$$Q = \left\{ s: \sum_{i=1}^n c_i s_i \leq \text{const} \right\}.$$

Статистикой этих критериев служит

$$\sum_{i=1}^n c_i \text{sign } x_i. \tag{2.2.10}$$

Для проверки гипотезы $H_0: \theta = 0$ против альтернативы $H: \theta \neq 0$ мы будем применять двустороннюю версию критерия (2.2.9) и отвергать H_0 (в пользу $H: \theta \neq 0$), если

$$\left| \sum_{i=1}^n c_i \text{sign } x_i \right| \geq \text{const}. \tag{2.2.11}$$

В теореме 2.2.1 будет показано, что при некоторых дополнительных условиях этот критерий оказывается локально оптимальным.

Отметим важное обстоятельство: если выполнено основное условие (2.1.4), то статистика (2.2.10) распределена свободно (при гипотезе H_0), даже если случайные ошибки ξ_1, \dots, ξ_n не являются одинаково распределенными. Поэтому уровни значимости критериев (2.2.9), (2.2.11) сохраняются и в этом последнем случае.

Теперь обратимся к многопараметрической модели (2.1.5), где $\theta \in \mathbb{R}^r$.

Предположим, что случайные ошибки независимы, одинаково распределены и что выполнены условия 2.1.1–2.1.3.

Условия $F(0) = 0.5$ и $f'(0) = 0$ заведомо выполнены для симметричных относительно нуля распределений, обладающих дифференцируемой плотностью. Мы не требуем такой симметрии, но сохраняем некоторые ее следствия. Будет показано, что при принятых условиях производные функции мощности

$$\begin{aligned} \frac{dP\{Q|\theta\}}{d\theta} &= \left(\frac{\partial P\{Q|\theta\}}{\partial \theta_1}, \dots, \frac{\partial P\{Q|\theta\}}{\partial \theta_r} \right), \\ \frac{d^2 P\{Q|\theta\}}{d\theta^2} &= \left\| \frac{\partial^2 P\{Q|\theta\}}{\partial \theta_i \partial \theta_j} \right\|_{i,j=1}^r \end{aligned}$$

существуют в окрестности $\theta = 0$. Как и ранее, при выборе критерия для проверки (2.2.2) будем руководствоваться его локальными свойствами, т.е. поведением его функции мощности (2.2.6) при θ , близких к гипотетическому значению $\theta = 0$. Среди всех знаковых

критериев мы будем далее рассматривать лишь так называемые локально несмещенные критерии, для которых

$$\left. \frac{dP\{Q|\theta\}}{d\theta} \right|_{\theta=0} = 0. \quad (2.2.12)$$

Для локально несмещенных критериев Q поведение функции мощности (2.2.6) в окрестности $\theta = 0$ определяет квадратичный член разложения Тейлора, т.е. квадратичная форма

$$\frac{1}{2} \theta^T \left(\frac{d^2 P\{Q|0\}}{d\theta^2} \right) \theta. \quad (2.2.13)$$

Выбор множества Q как критического для H_0 надо считать тем более удачным, чем быстрее возрастает значение $P\{Q|\theta\}$ при удалении θ от $\theta = 0$. Поскольку равномерно наиболее мощного критерия в многопараметрическом случае не существует даже в локальном смысле, нам придется заняться оптимизацией какой-либо разумно выбранной скалярной характеристики Q . Такой характеристикой для локально несмещенных критериев может быть коэффициент кривизны в точке $\theta = 0$ (в которой выполнено условие (2.2.12)) поверхности в $(r+1)$ -мерном пространстве, заданной (2.2.6). Для функций многих переменных существует несколько коэффициентов кривизны. Геометры обычно имеют дело с гауссовой кривизной. Мы же будем использовать коэффициент средней кривизны. Эта средняя кривизна пропорциональна следу уже упомянутой матрицы квадратичной формы (2.2.13). (Определения коэффициентов кривизны см. [14, Ч. I, § 8]).

Итак, перед нами экстремальная задача: выбрать совокупность Q так, чтобы

$$\text{tr} \frac{d^2 P\{Q|\theta\}}{d\theta^2} \Big|_{\theta=0} \Rightarrow \max \quad (2.2.14)$$

при условиях (2.2.7) и (2.2.12). Множество Q , которое служит решением этой задачи, будем называть *локально оптимальным знаковым критерием* уровня ε . Заметим, что в силу (2.2.6)

$$\text{tr} \frac{d^2 P\{Q|0\}}{d\theta^2} = \sum_{s \in Q} \text{tr} \frac{d^2 P\{s|0\}}{d\theta^2}.$$

При условиях 2.1.1–2.1.3 производные $\frac{dP\{s|\theta\}}{d\theta}$ и $\frac{d^2 P\{s|\theta\}}{d\theta^2}$ можно вычислить явно. Отправляясь от правдоподобия (2.2.8), легко убедиться, что при условиях 2.1.1–2.1.3

$$(P\{x_i > 0\})^{\frac{1+s_i}{2}} (P\{x_i < 0\})^{\frac{1-s_i}{2}} = \frac{1}{2} (1 + 2s_i f(0) c_i \theta + o(|\theta|^2)).$$

Отсюда следует, что

$$P\{S(X) = s | \theta\} = \left(\frac{1}{2}\right)^n \left[1 + 2f(0) \sum_{i=1}^n s_i c_i \theta + 4(f(0))^2 \times \right. \\ \left. \times \sum_{1 \leq i \neq j \leq n} (s_i c_i \theta)(s_j c_j \theta) + o(|\theta|^2) \right]. \quad (2.2.15)$$

Поэтому след $\text{tr} \frac{d^2 P\{s | \theta\}}{d\theta^2} \Big|_{\theta=0}$ пропорционален

$$\sum_{1 \leq i \neq j \leq n} s_i s_j \sum_{\alpha=1}^r c_{i\alpha} c_{j\alpha}. \quad (2.2.16)$$

Отметим, что выражение (2.2.16) одинаково для наборов s и $-s$. Поэтому s и $-s$ одновременно либо входят либо не входят в критерий максимальной средней кривизны. При этом

$$\frac{dP\{s | \theta\}}{d\theta} \Big|_{\theta=0} = - \frac{dP\{-s | \theta\}}{d\theta} \Big|_{\theta=0},$$

как следует из (2.2.15), так что для пары $\{s\} \cup \{-s\}$ получаем

$$\frac{dP\{s \cup (-s) | \theta\}}{d\theta} \Big|_{\theta=0} = 0.$$

Следовательно, критерий максимальной средней кривизны автоматически оказывается локально несмещенным. Поэтому можно сделать вывод, что оптимальный в смысле (2.2.14) знаковый критерий имеет вид

$$Q = \left\{ s: \sum_{1 \leq i \neq j \leq n} s_i s_j \sum_{\alpha=1}^r c_{i\alpha} c_{j\alpha} > \text{const} \right\}. \quad (2.2.17)$$

Добавив к правой и левой частям неравенства в (2.2.17) не зависящую от набора s величину

$$\sum_{1 \leq i \neq j \leq n} s_i s_j \sum_{\alpha=1}^r c_{i\alpha} c_{j\alpha} = \sum_{i=1}^n \sum_{\alpha=1}^r c_{i\alpha}^2,$$

получим для локально оптимального знакового критерия более удобное выражение

$$Q = \left\{ s: \sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} s_i \right)^2 > \text{const} \right\}.$$

Сформулируем полученные результаты в виде теоремы.

Теорема 2.2.1. *Локально оптимальный знаковый критерий (т.е. локально несмещенный знаковый критерий максимальной средней кривизны) в модели (2.1.5) для проверки гипотезы (2.2.2) при условиях 2.1.1–2.1.3 имеет вид*

$$Q = \left\{ X: \sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \text{sign } x_i \right)^2 > \text{const} \right\}. \quad (2.2.18)$$

При гипотезе (2.2.2) статистика этого критерия распределена свободно, если выполнено единственное условие (2.1.4).

В векторно-матричной форме знаковая статистика из (2.2.18) имеет вид

$$\sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \text{sign } x_i \right)^2 = |C^T S(X)|^2 = (S(X))^T (C C^T) (S(X)). \quad (2.2.19)$$

Как по заданному уровню значимости ε (2.2.7) выбрать в (2.2.18) критическое значение, обозначенное const, мы обсудим в § 2.3. Знаковый критерий (2.2.11) является однопараметрической формой (2.2.18).

З а м е ч а н и е. Критерий (2.2.18) оптимален и в ином смысле, впрочем, весьма близком к (2.2.14): среди локально несмещенных критериев заданного уровня значимости средняя мощность, если ее вычислять по сфере бесконечно малого радиуса, достигает максимального значения именно для критерия (2.2.18). Скажем точнее. Рассмотрим произвольный знаковый критерий для проверки (2.2.2), уровень значимости которого не превосходит ε . Пусть $\beta(\theta)$ обозначает его функцию мощности. Рассмотрим

$$\lim_{\rho \rightarrow 0} \rho^{-r-1} \int_{\theta: |\theta|=\rho} (\beta(\theta) - \beta(0)) d\theta. \quad (2.2.20)$$

Мы утверждаем, что знаковый критерий (2.2.18) уровня ε оптимален в том смысле, что для него локальная характеристика (2.2.20) достигает максимального значения при условии (2.2.7).

Вместе с (2.2.18) можно рассмотреть и другие положительно определенные квадратичные формы от $C^T S(X)$. Для них средняя мощность типа (2.2.20) достигает максимального значения, если вместо сфер взять соответствующие эллипсоиды, стягивающиеся к $\theta = 0$. В совокупности знаковых критериев квадратичные знаковые критерии являются допустимыми. Напомним, что критерий A называется допустимым в некоторой совокупности критериев, если в этой совокупности не существует критерия, такого же, как у A , или меньшего уровня значимости, мощность которого при альтернативах была

бы не ниже, а для некоторых альтернатив даже выше, чем мощность A . Если иметь в виду вычисление критических значений при большом числе наблюдений n , то следует выделить критерии вида

$$Q = \{S(X): (S(X))^T C(C^T C)^{-1} C^T S(X) > \text{const}\}. \quad (2.2.21)$$

Мы покажем в § 2.3, что их критериальная статистика

$$(S(X))^T C(C^T C)^{-1} C^T S(X) \quad (2.2.22)$$

при $n \rightarrow \infty$ и некоторых условиях на матрицу плана C асимптотически распределена по закону хи-квадрат. Поэтому для вычислений критической постоянной в (2.2.21) можно пользоваться соответствующей аппроксимацией. Заметим в заключение, что для однопараметрической регрессии все квадратичные знаковые критерии эквивалентны критерию (2.2.11).

§ 2.3. Вычисление критических значений. Асимптотическая теория

Обсудим, как при H_0 по заданному уровню значимости ϵ можно вычислить критические значения для статистик (2.2.18), (2.2.21) или для другой знаковой статистики. Введем независимые случайные величины ζ_i , $i = 1, \dots, n$, положив $\zeta_i = \text{sign } \xi_i$, $\zeta = (\zeta_1, \dots, \zeta_n)^T$. При условии (2.1.4)

$$P\{\zeta_i = 1\} = P\{\zeta_i = -1\} = 0.5. \quad (2.3.1)$$

При гипотезе (2.2.1) в модели (2.1.1) и при условии (2.1.4) случайные величины $\text{sign } x_i$ распределены так же, как ζ_i , $i = 1, \dots, n$. Поэтому при H_0 знаковая статистика (2.2.19) по распределению совпадает со случайной величиной

$$\sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \zeta_i \right)^2. \quad (2.3.2)$$

Распределение этой последней определяется матрицей плана C . Хотя указать аналитическое выражение для этого распределения не удается, численно оно может быть найдено с необходимой точностью. (О вычислении этого распределения и его квантилей см. § 2.5). Заметим, что случайная величина (2.3.2) распределена дискретно. Напомним, что мы понимаем под квантилью дискретной случайной величины.

О п р е д е л е н и е. *Квантилью уровня α дискретной случайной величины ξ , назовем следующую величину ξ_α :*

а) $\xi_\alpha = \sup\{x: P\{\xi \leq x\} \leq \alpha\}$, если решение уравнения $P\{\xi \leq x\} = \alpha$ не существует;

б) $\xi_\alpha = \frac{1}{2}(\inf\{x: P\{\xi \leq x\} = \alpha\} + \sup\{x: P\{\xi \leq x\} = \alpha\})$, если решение уравнения существует (т.е. середина интервала, на котором это решение существует).

Пусть $q_{1-\varepsilon}^n$ обозначает квантиль уровня $1-\varepsilon$ случайной величины (2.3.2); тогда

$$P\left\{\sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha}\zeta_i\right)^2 > q_{1-\varepsilon}^n\right\} \leq \varepsilon. \quad (2.3.3)$$

Аналогично определим квантиль $\tilde{q}_{1-\varepsilon}^n$ случайной величины

$$\zeta^T C (C^T C)^{-1} C^T \zeta. \quad (2.3.4)$$

Теперь можно сформулировать правила для проверки гипотезы $H_0: \theta = \theta^0$, где θ^0 задано, с помощью знаковых критериев (2.2.18) и (2.2.21).

П р а в и л о. Гипотезу $H_0: \theta = \theta^0$ в модели (2.1.5) следует отвергнуть на гарантированном уровне ε , $0 < \varepsilon < 1$, если:

а) (в случае применения критерия (2.2.18))

$$[S(X - C\theta^0)]^T C C^T [S(X - C\theta^0)] > q_{1-\varepsilon}^n; \quad (2.3.5)$$

б) (в случае применения критерия (2.2.21))

$$[S(X - C\theta^0)]^T C (C^T C)^{-1} C^T [S(X - C\theta^0)] > \tilde{q}_{1-\varepsilon}^n. \quad (2.3.6)$$

Подчеркнем, что критерии (2.3.5) и (2.3.6) действуют и в том случае, когда случайные ошибки распределены неодинаково; необходимо лишь выполнение (2.1.4).

При $n \rightarrow \infty$ для этих и других квадратичных знаковых статистик при гипотезе H_0 можно указать асимптотическое распределение при некоторых предположениях о матрице C . В этой книге далее мы будем рассматривать ограниченные и асимптотически невырожденные планы. А именно, мы примем условия 2.1.7, 2.1.8: существует не зависящее от n число K такое, что

$$\max_{1 \leq i \leq n} \max_{1 \leq \alpha \leq r} |c_{i\alpha}| < K \quad \text{для всех } n,$$

$$\frac{1}{n} C^T C \rightarrow \Sigma,$$

где матрица Σ невырожденная (и тем самым положительно определенная).

При условиях 2.1.7, 2.1.8 случайный вектор $\frac{1}{\sqrt{n}} C^T \zeta$ при $n \rightarrow \infty$ асимптотически распределен как $N(0, \Sigma)$. Из этого вытекают следствия об асимптотическом распределении статистик (2.2.19) и (2.2.22).

Теорема 2.3.1. При гипотезе (2.2.2) относительно модели (2.1.5), выполнении (2.1.4) и условий 2.1.7, 2.1.8, когда $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \operatorname{sign} x_i \right)^2 \xrightarrow{d} \sum_{\alpha=1}^r \lambda_{\alpha} \eta_{\alpha}^2, \quad (2.3.7)$$

где $\lambda_1, \dots, \lambda_r$ — собственные значения матрицы Σ , а случайные величины η_1, \dots, η_r независимы и распределены по закону $N(0, 1)$;

$$(S^T(X)C)(C^T C)^{-1}(C^T S(X)) \xrightarrow{d} \chi^2(r), \quad (2.3.8)$$

где $\chi^2(r) = \eta_1^2 + \dots + \eta_r^2$, а η_1, \dots, η_r те же, что и в (2.3.7).

Доказательство. Введем r -мерный случайный вектор Y , распределенный по нормальному закону $N(0, \Sigma)$. В силу многомерной центральной предельной теоремы вектор $\frac{1}{\sqrt{n}} C^T \zeta$ в условиях теоремы 2.3.1 при $n \rightarrow \infty$ распределен как Y . Поэтому статистики (2.3.7) и (2.3.8) асимптотически распределены как $|Y|^2$ и $Y^T \Sigma^{-1} Y$ соответственно. Пусть Q — такая ортогональная матрица, что $Q \Sigma Q^T = \Lambda$, где $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$. При этом $\lambda_1 > 0, \dots, \lambda_r > 0$, так как матрица Σ положительно определенная. Пусть $Z = QY$, $Z = (z_1, \dots, z_r)^T$. Ясно, что $Z \sim N(0, \Lambda)$, а потому z_{α} , $\alpha = 1, \dots, r$, можно представить в виде $z_{\alpha} = \sqrt{\lambda_{\alpha}} \eta_{\alpha}$, где случайные величины η_1, \dots, η_r суть независимые $N(0, 1)$. Теперь имеем

$$|Y|^2 = \sum_{\alpha=1}^r z_{\alpha}^2 = \sum_{\alpha=1}^r \lambda_{\alpha} \eta_{\alpha}^2,$$

$$Y^T \Sigma^{-1} Y = Z^T (Q^T \Sigma^{-1} Q) Z = Z^T \Lambda^{-1} Z = \sum_{\alpha=1}^r \eta_{\alpha}^2.$$

Теорема доказана. \square

Простота асимптотического распределения статистики (2.2.22) дает ей преимущество по сравнению с другими квадратичными знаковыми статистиками.

Статистике (2.2.22) можно дать и другое асимптотическое обоснование, основанное на применении принципа Роя [105].

Рассмотрим модель (2.1.5) и выберем в подпространстве $L = \{l: l = C\theta, \theta \in \mathbb{R}^r\}$ какое-либо направление $l = (l_1, \dots, l_n)^T$, $|l| = 1$. Рассмотрим одномерную альтернативу $H^{(l)}$: $C\theta = tl$, $t \in \mathbb{R}^1$, к гипотезе $H_0: \theta = 0$. Как следует из результатов § 2.2, локально оптимальный знаковый тест для проверки H_0 против $H^{(l)}$ имеет вид

$$\left| \sum_{i=1}^n l_i \operatorname{sign} x_i \right| > \operatorname{const}. \quad (2.3.9)$$

Постоянная в правой части (2.3.9) при конечных n , вообще говоря, зависит от выбора вектора l . Однако при гипотезе H_0 статистика (2.3.9) распределена асимптотически как $N(0, 1)$. Действительно, достаточным условием для этого служит условие

$$\max_{1 \leq i \leq n} |l_i|^2 \rightarrow 0, \quad (2.3.10)$$

которое выполняется, если матрица C удовлетворяет условиям 2.1.7, 2.1.8. Поэтому при больших n критические значения в (2.3.9) практически не зависят от l .

Гипотезу H_0 следует отвергнуть в пользу $H^{(l)}$, если статистика (2.3.9) принимает большое значение. Альтернатива $H: \theta \neq 0$ представляет собой объединение альтернатив $H^{(l)}$. Следуя С.Рюю [105], представляется разумным отвергать H_0 , если большое значение принимает хотя бы одна из статистик (2.3.9), когда l пробегает все возможные значения $l \in \mathbb{L}$, $\|l\| = 1$. То есть взять

$$\max_{l \in \mathbb{L}} \left(\sum_{i=1}^n l_i \operatorname{sign} x_i \right)^2 \quad (2.3.11)$$

в качестве критериальной статистики для проверки H_0 . Очевидно, что (2.3.11) равно $|\operatorname{proj}_{\mathbb{L}} S(X)|^2$, где $\operatorname{proj}_{\mathbb{L}} S(X)$ означает проекцию вектора $S(X)$ на подпространство \mathbb{L} в евклидовой метрике. Вид нормального проектора на подпространство \mathbb{L} хорошо известен из теории наименьших квадратов (см., например, [3, 48, § 7.2] или [17, § 5.2]): его матрица равна $C(C^T C)^{-1} C^T$. Поэтому

$$\operatorname{proj}_{\mathbb{L}} S(X) = C(C^T C)^{-1} C^T S(X).$$

Отсюда следует, что

$$\max_{l \in \mathbb{L}} \left(\sum_{i=1}^n l_i \operatorname{sign} x_i \right)^2 = S^T(X) C(C^T C)^{-1} C^T S(X),$$

т.е. (2.3.11) совпадает со статистикой (2.2.22).

§ 2.4. Пример. Двухфакторные таблицы

В некоторых задачах на параметр θ в моделях типа (2.1.5) могут накладываться линейные связи типа уравнений. Параметр θ в этом случае перестает быть свободным; фактически он пробегает пространство меньшего числа измерений, чем число его координат. В таких задачах при построении локально оптимальных критериев для H_0 (2.2.2) следует ориентироваться на среднюю кривизну функции мощности на упомянутом подпространстве. Обычно ее можно вычислить, не прибегая к перепараметризации. Покажем это на примере двухфакторных таблиц.

Рассмотрим двухфакторный эксперимент с одинаковым числом наблюдений в каждой ячейке. Примем аддитивную модель действия факторов. Таким образом, мы наблюдаем случайные величины x_{ijk} , которые могут быть представлены в виде модели:

$$x_{ijk} = \mu + \alpha_i + \beta_j + \xi_{ijk}, \quad (2.4.1)$$

$$i = 1, \dots, r, \quad j = 1, \dots, t, \quad k = 1, \dots, m.$$

Здесь r, t — соответственно число уровней первого и второго факторов, m — число наблюдений при каждой комбинации уровней факторов. Параметры $\mu, \alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_t$ неизвестны. Как обычно (см., например, [48, § 7.3, с] или [17, § 5.5]), считаем, что

$$\sum_{i=1}^r \alpha_i = 0, \quad \sum_{j=1}^t \beta_j = 0. \quad (2.4.2)$$

К этой модели мы еще вернемся, в частности, при обсуждении проверки линейных гипотез в гл. 4. Сейчас же заметим, что для одновременного оценивания всех неизвестных параметров нам необходим критерий для проверки гипотезы

$$H_0: \mu = 0, \quad \alpha_1 = \dots = \alpha_r = 0, \quad \beta_1 = \dots = \beta_t = 0. \quad (2.4.3)$$

Ввиду существования связей (2.4.2) при построении локально оптимального критерия надо ориентироваться на среднюю кривизну функции мощности на подпространстве параметров, заданном этими уравнениями. Чтобы сделать точность статистических выводов одинаковой для всех параметров, целесообразно перейти к их нормированным версиям, положив

$$\mu' = \mu\sqrt{rt}, \quad \alpha'_i = \alpha_i\sqrt{t}, \quad \beta'_j = \beta_j\sqrt{r}. \quad (2.4.4)$$

В этой параметризации модель (2.4.1) принимает вид

$$x_{ijk} = \frac{\mu}{\sqrt{rt}} + \frac{\alpha_i}{\sqrt{t}} + \frac{\beta_j}{\sqrt{r}} + \xi_{ijk}, \quad (2.4.5)$$

$$i = 1, \dots, r, \quad j = 1, \dots, t, \quad k = 1, \dots, m,$$

если отбросить «штрих» в обозначениях (2.4.4). Для новых параметров действуют те же связи (2.4.2). Положим по определению

$$Z_{ij} = \frac{1}{\sqrt{m}} \sum_{k=1}^m \text{sign } x_{ijk}. \quad (2.4.6)$$

Можно показать, что с точностью до неслучайного слагаемого и неслучайного сомножителя главная часть $P\{S(X)|\theta\} + P\{-S(X)|\theta\}$

согласно (2.2.15) в данном случае равна

$$\left(\mu Z_{..} + \frac{1}{\sqrt{r}} \sum_{i=1}^r Z_{i, \alpha_i} + \frac{1}{\sqrt{t}} \sum_{j=1}^t Z_{, j} \beta_j \right)^2. \quad (2.4.7)$$

Здесь замена индекса точкой означает усреднение по этому индексу:

$$Z_{i.} = \frac{1}{t} \sum_{j=1}^t Z_{ij}, \quad Z_{.j} = \frac{1}{r} \sum_{i=1}^r Z_{ij}, \quad Z_{..} = \frac{1}{rt} \sum_{i=1}^r \sum_{j=1}^t Z_{ij}. \quad (2.4.8)$$

Средняя кривизна функции (2.4.7) от переменных μ, α, β в точке

$$\mu = 0, \quad \alpha_1 = \dots = \alpha_r = 0, \quad \beta_1 = \dots = \beta_t = 0$$

равна сумме коэффициентов при квадратах переменных, т.е.

$$Z_{..}^2 + \frac{1}{r} \sum_{i=1}^r Z_{i.}^2 + \frac{1}{t} \sum_{j=1}^t Z_{.j}^2$$

Теперь вспомним, что средняя кривизна равна сумме нормальных кривизн по произвольному набору ортогональных направлений.

В $(r+t+1)$ -мерном пространстве \mathbb{R}^{r+t+1} , в котором изменяется параметр $\theta = (\mu, \alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_t)^T$, рассмотрим два одномерных (ортогональных друг другу) подпространства, M_1 и M_2 , порожденных единичными векторами

$$e = \left(0, \frac{1}{\sqrt{r}}, \dots, \frac{1}{\sqrt{r}}, 0, \dots, 0 \right)^T, \quad f = \left(0, 0, \dots, 0, \frac{1}{\sqrt{t}}, \dots, \frac{1}{\sqrt{t}} \right)^T.$$

Теперь мы можем сказать, что параметр θ в задаче (2.4.5) с учетом связей (2.4.2) изменяется в ортогональном дополнении $M_1 \oplus M_2$ до \mathbb{R}^{r+t+1} . Легко видеть, что кривизна функций (2.4.7) (в нулевой точке) на каждом из подпространств M_1, M_2 равна $Z_{..}^2$. Поэтому искомая статистика есть

$$\frac{1}{r} \sum_{i=1}^r Z_{i.}^2 + \frac{1}{t} \sum_{j=1}^t Z_{.j}^2 - Z_{..}^2 \quad (2.4.9)$$

Нетрудно убедиться, что это выражение пропорционально

$$\sum_{i=1}^r \sum_{j=1}^t (Z_{i.} + Z_{.j} - Z_{..})^2 \quad (2.4.10)$$

Поэтому локально оптимальный знаковый критерий в двухфакторной аддитивной модели в обозначениях (2.4.6) (в задаче (2.4.5) с учетом

(2.4.2)) имеет вид

$$\left\{ X: \sum_{i=1}^r \sum_{j=1}^t (Z_{i.} + Z_{.j} - Z_{..})^2 > \text{const} \right\}. \quad (2.4.11)$$

Можно показать, что при фиксированных r, t и $m \rightarrow \infty$ статистика (2.4.10) при нулевой гипотезе (2.4.3) асимптотически распределена как $\chi^2(r+t-1)$. Примечательно, что (2.4.10) по виду совпадает с числителем F -критерия для проверки H_0 (2.4.3) в нормальной модели (если не считать деления на число степеней свободы).

§ 2.5. Вычисление критических значений. Конечные объемы выборок

В этом параграфе мы обсуждаем методы вычисления критических значений знаковой статистики при H_0 , соответствующие заданному уровню значимости ε , для рассмотренных выше задач при конечных объемах выборок.

Для того чтобы воспользоваться результатами § 2.2 при построении критического множества, нужно уметь вычислять квантили распределений случайных величин (2.3.2) или (2.3.4) (см. определение 2.3.1 квантили дискретной случайной величины).

При небольших объемах выборок распределения, а значит и квантили случайных величин (2.3.2) и (2.3.4) можно вычислить перебором. Число равновероятных комбинаций равно 2^n , а соответствующие значения статистики зависят от матрицы плана эксперимента C . Мы обычно применяем этот метод при $n \leq 16$. При других умеренных n для вычисления распределения мы применяем метод статистических испытаний Монте-Карло. При больших n следует пользоваться указанной в § 2.3 асимптотикой.

Для моделирования (2.3.2) или (2.3.4) нужно получить последовательность $\{\zeta_i\}$ с указанными выше свойствами (2.3.1). Лучше от последовательности $\{\zeta_i\}$ перейти к последовательности случайных величин, принимающих значения 0 и 1 с вероятностями 0.5. Тогда моделирование случайных величин (2.3.2) и (2.3.4) можно упростить и ускорить.

В использованных методах мы можем указать точный уровень значимости знакового критерия, близкий или равный заданному.

Примеры вычисления критических значений для простейшей линейной регрессии были приведены в гл. 1 при обсуждении оценивания постоянной в законе Хаббла. В задачах линейной регрессии критические значения соответствующих статистик зависят от плана эксперимента (матрицы плана C) и должны вычисляться для каждой конкретной задачи.

Для обсуждавшейся выше задачи двухфакторного анализа план эксперимента известен, если известны количество уровней каждого фактора в двухфакторной таблице и число наблюдений в каждой клетке этой таблицы.

Таблица 2.5.1

Квантили случайной величины ζ (2.5.1)

	$m=1$		$m=2$		$m=3$		$m=4$		$m=\infty$	
	q	p	q	p	q	p	q	p	q	p
$r=2, t=2$	0.000	0.125	4.000	0.758	5.333	0.880	5.000	0.874	6.251	0.900
	3.000	0.625	5.500	0.945	6.333	0.909	6.000	0.915		
	4.000	1.000	6.000	0.977	6.667	0.961	6.750	0.956		
$t=3$	2.000	0.281	6.000	0.854	6.000	0.818	6.667	0.864	7.779	0.900
	4.667	0.844	7.333	0.924	6.889	0.901	7.667	0.918		
	6.000	1.000	8.000	0.959	8.667	0.951	9.000	0.954		
$t=4$	6.000	0.617	7.250	0.868	7.500	0.856	7.625	0.857	9.238	0.900
	6.500	0.930	9.250	0.945	8.833	0.907	9.000	0.908		
	8.000	1.000	10.000	0.968	10.167	0.963	10.625	0.957		
$t=5$	6.000	0.557	9.200	0.887	9.467	0.880	9.400	0.858	10.645	0.900
	7.600	0.791	10.000	0.911	10.533	0.921	10.400	0.903		
	8.400	0.967	11.200	0.966	11.867	0.951	12.000	0.951		
$r=3, t=3$	6.333	0.691	7.778	0.859	7.741	0.865	7.889	0.853	9.238	0.900
	7.222	0.973	9.111	0.924	9.222	0.913	8.889	0.902		
	9.000	1.000	9.778	0.953	10.407	0.960	10.556	0.951		
$t=4$	8.667	0.880	9.000	0.855	9.333	0.854	9.333	0.852	10.645	0.900
	9.333	0.924	10.333	0.922	10.222	0.904	10.500	0.908		
	10.000	0.995	11.667	0.957	12.000	0.952	12.167	0.954		
$t=5$	10.200	0.856	10.400	0.852	10.511	0.851	10.600	0.851	12.017	0.900
	10.733	0.911	11.467	0.903	11.756	0.908	11.800	0.901		
	11.267	0.955	12.933	0.953	13.356	0.951	13.533	0.950		
$r=4, t=4$	10.000	0.879	10.500	0.854	10.583	0.857	10.500	0.851	12.017	0.900
	11.000	0.912	11.500	0.903	11.667	0.901	11.750	0.901		
	11.750	0.956	13.000	0.953	13.333	0.953	13.500	0.951		

Рассмотрим статистику двухфакторного анализа (2.4.10) для проверки нулевой гипотезы (2.4.3). Распределение такой статистики совпадает с распределением случайной величины

$$\zeta = \frac{4}{m} \sum_{i=1}^r \sum_{j=1}^t \left(B_{ij} + B_{.j} - B_{..} - \frac{m}{2} \right)^2, \quad (2.5.1)$$

где через B_{ij} , $i = 1, \dots, r$; $j = 1, \dots, t$, обозначены независимые случайные величины, распределенные по биномиальному закону с параметрами $(m, 0.5)$. Точкой, как и ранее, обозначено усреднение по соответствующему индексу. Распределение дискретной случайной величины ζ можно определить точно, перебирая все возможные для нее значения и вычисляя соответствующие вероятности. Отметим, что

количество различных значений, принимаемых этой случайной величиной, невелико. Например, при $r = t = 2$, $m = 1$ оно равно 3, при $r = 2$, $t = 4$, $m = 4$ равно 81 и т.д. В табл. 2.5.1 приведены некоторые квантильные значения q и соответствующие им точные вероятности $p := \mathbf{P}\{\zeta \leq q\}$ для некоторых комбинаций параметров r , t , m и для некоторых верхних «хвостовых» вероятностей p . В столбце с $m = \infty$ для сравнения указаны 0.90- и 0.95-квантили предельного при $m \rightarrow \infty$ для ζ распределения $\chi^2(r+t-1)$, о котором упоминалось в § 2.3 и § 2.4.

Глава 3

ЗНАКОВЫЕ ОЦЕНКИ

§ 3.1. Знаковые оценки и их вычисление

Во введении уже было сказано, как через тестовые статистики можно определить точечные (В.5) и доверительные (В.4) оценки параметров. Имея знаковые критерии гл. 2 для модели (2.1.5) и их статистики (2.2.19), (2.2.22), мы можем с их помощью ввести и знаковые оценки в схеме (2.1.5). Например,

$$\hat{\theta}_n = \arg \min_{\theta \in \mathbb{R}^r} (S(X - C\theta))^T C C^T S(X - C\theta), \quad (3.1.1)$$

либо

$$\hat{\theta}_n = \arg \min_{\theta \in \mathbb{R}^r} (S(X - C\theta))^T C (C^T C)^{-1} C^T S(X - C\theta), \quad (3.1.2)$$

либо им подобные, основанные на других квадратичных знаковых критериях для проверки $H_0: \theta = 0$.

Соответствующие доверительные множества для неизвестного параметра в модели (2.1.5) с коэффициентом доверия $1 - \varepsilon$ имеют вид

$$\left\{ (S(X - C\theta))^T C C^T S(X - C\theta) \leq q_{1-\varepsilon} \right\}, \quad (3.1.3)$$

$$\left\{ (S(X - C\theta))^T C (C^T C)^{-1} C^T S(X - C\theta) \leq \tilde{q}_{1-\varepsilon} \right\}, \quad (3.1.4)$$

где квантиль $q_{1-\varepsilon}$ была определена в (2.3.3), а $\tilde{q}_{1-\varepsilon}$ — квантиль случайной величины (2.3.4).

В § 3.1 описаны вычислительные алгоритмы, основанные на теоретических результатах гл. 2, и показана работа этих алгоритмов на модельных примерах. Вычислительные алгоритмы реализованы программно и оформлены в виде диалоговой системы, названной Sign. Эта система работает на IBM PC и совместимых с ними компьютерах. Приведенные в книге примеры выполнены в этой системе.

Обратимся к вычислению знаковых оценок. Начнем с оценивания параметров линейной регрессии в модели (2.1.5). Здесь и далее мы ограничимся вычислением оценок вида (3.1.1). Мы предлагаем два

алгоритма: переборный и итерационный. Запишем задачу (3.1.1) в координатной форме:

$$\sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \operatorname{sign} \left(x_i - \sum_{\beta=1}^r c_{i\beta} \theta_{\beta} \right) \right)^2 \Rightarrow \min_{\theta \in \mathbb{R}^r}. \quad (3.1.5)$$

Решение такой задачи всегда существует, поскольку целевая функция является кусочно постоянной: при изменении векторного параметра θ во всем пространстве она принимает конечное число значений. Ясно, что решение может быть неединственным и в общем случае является некоторым множеством r -мерного параметрического пространства. Пример из гл. 1 (определение постоянной Хаббла, § 1.2) хорошо иллюстрирует этот факт. В этом примере параметрическое пространство — числовая прямая. Обработывая данные из статьи Э.Хаббла [80], мы получили в качестве знаковой оценки интервал (θ_1, θ_2) . В то же время для другого набора наблюдений (современные данные) знаковая оценка свелась к единственному значению.

Мы реализовали два алгоритма вычисления знаковой оценки (3.1.1) при $\theta \in \mathbb{R}^r$:

1) при $r = 2$ реализован алгоритм перебора всевозможных значений функции из (3.1.5) (в этом случае мы строим и доверительные множества заданного уровня доверия);

2) для произвольной размерности r построен итерационный алгоритм определения решения задачи (3.1.5).

Опишем эти алгоритмы.

1) Переборный алгоритм для $r = 2$.

Каждое слагаемое в сумме, входящей в (3.1.5), является функцией двух параметров: θ_1 и θ_2 . Изобразим на плоскости с координатами (θ_1, θ_2) прямые $x_i - c_{i1}\theta_1 - c_{i2}\theta_2 = 0$, $i = 1, \dots, n$. Точки пересечения этих прямых назовем *узловыми*. Таким образом, плоскость разбивается на многоугольники (конечные или бесконечные), в каждом из которых функция из (3.1.5) принимает постоянное значение. Предположим, что матрица плана C и вектор наблюдений X таковы, что любые две прямые пересекаются и что в узловых точках пересекаются только две прямые. (Нетрудно обобщить алгоритм на случай произвольного числа пересечений в узловой точке). Пусть M_{ij} — точка пересечения i -й и j -й прямых (пусть ее координаты $(\theta_1^{ij}, \theta_2^{ij})$). Эти прямые делят плоскость на четыре части, в каждой из которых знаки разностей $x_i - c_{i1}\theta_1 - c_{i2}\theta_2$ и $x_j - c_{j1}\theta_1 - c_{j2}\theta_2$ могут быть определены. Для других функций $x_k - c_{k1}\theta_1 - c_{k2}\theta_2$ ($k \neq i, j$) знак в каждой смежной с узловой точкой M_{ij} области определяется знаком разности $x_k - c_{k1}\theta_1^{ij} - c_{k2}\theta_2^{ij}$. Поэтому в самой узловой точке, в прилегающих к ней смежных областях и на прямолинейных отрезках (где или $x_i - c_{i1}\theta_1 - c_{i2}\theta_2 = 0$, или $x_j - c_{j1}\theta_1 - c_{j2}\theta_2 = 0$) легко определить значение целевой функции из (3.1.5). Перебирая все узловые точки, нетрудно определить минимальное значение целевой функции из (3.1.5) и множество на плоскости, где этот минимум достигается.

Множеством минимума целевой функции (3.1.5) может быть многоугольник, интервал, точка или их объединение. Ниже мы приведем примеры, в которых реализуются некоторые из этих возможностей. Переборный алгоритм при умеренных значениях n позволяет также строить для неизвестных истинных значений (θ_1, θ_2) доверительные множества, что мы также покажем ниже.

2) Итерационный алгоритм.

Как уже отмечалось, множество минимума в (3.1.5) можно найти перебором всех значений целевой функции и при $r > 2$. Однако при увеличении числа наблюдений объем таких вычислений быстро возрастает. Поэтому для решения (3.1.5) мы предлагаем итерационный алгоритм. Он сводится к последовательной минимизации каждого из слагаемых в (3.1.5). Теоретически его сходимость не доказана, но во всех многочисленных случаях его применений мы получали правильные результаты.

Вычислительный алгоритм состоит в последовательном вычислении r -мерных векторных значений $\theta(k) = (\theta_1(k), \dots, \theta_r(k))^T$ для $k = 1, 2, \dots$

Шаг 1. Выбираем начальное приближение $\theta(0) = (\theta_1(0), \dots, \theta_r(0))^T$.

Шаг 2. Шаг итерации состоит в переходе от $\theta(k-1) = (\theta_1(k-1), \dots, \theta_r(k-1))^T$ к $\theta(k) = (\theta_1(k), \dots, \theta_r(k))^T$, $k = 1, 2, \dots$. Значение $\theta_\alpha(k)$, $\alpha = 1, \dots, r$, выбираем по правилу

$$\theta_\alpha(k) =$$

$$\arg \min_{\theta_\alpha \in \mathbb{R}^1} \left(\sum_{i=1}^n c_{i\alpha} \operatorname{sign} \left(x_i - c_{i\alpha} \theta_\alpha - \sum_{\beta < \alpha} c_{i\beta} \theta_\beta(k) - \sum_{\beta > \alpha} c_{i\beta} \theta_\beta(k-1) \right) \right)^2,$$

$$\alpha = 1, \dots, r.$$

Шаг 3. Процесс останавливаем на шаге k , если при заданной вычислительной точности ε

$$\sqrt{\sum_{\alpha=1}^r (\theta_\alpha(k) - \theta_\alpha(k-1))^2} < \varepsilon,$$

либо при обращении функции из (3.1.5) в нуль.

З а м е ч а н и е. Мы применяем итерационную процедуру типа Зейделя, которая в данном случае состоит в использовании на k -м шаге итерации при вычислении α -й координаты $\theta_\alpha(k)$ всех уже вычисленных на этом k -м шаге значений координат $\theta_1(k), \dots, \theta_{\alpha-1}(k)$. В отличие от покоординатной процедуры, использование процедуры типа Зейделя позволяет закончить итерационный процесс за меньшее число шагов.

З а м е ч а н и е. Тестирование итерационной процедуры на модельных данных показало, что выбор начального приближения влияет

на количество шагов итерации. Во многих примерах мы использовали в качестве первого шага итерационной процедуры оценку наименьших квадратов и получали за небольшое число итераций знаковую оценку.

Пример 1. Моделируем регрессионную зависимость

$$x_i = \theta_1 + c_i \theta_2 + \xi_i, \quad i = 1, \dots, n, \quad (3.1.6)$$

Здесь $\theta_1 = 50$, $\theta_2 = -10$, а случайные величины ξ_i имеют одинаковое засоренное нормальное распределение с функцией распределения Тьюки $F(x) = (1-\delta)\Phi(x/\sigma_0) + \delta\Phi(x/\sigma_1)$, где $\delta = 0.15$, $\sigma_0 = 2$, $\sigma_1 = 30$, $\Phi(x)$ — функция стандартного нормального распределения, $n = 50$. На рис. 3.1.1 изображены результаты оценивания параметров модели знаковым методом и методом наименьших квадратов.

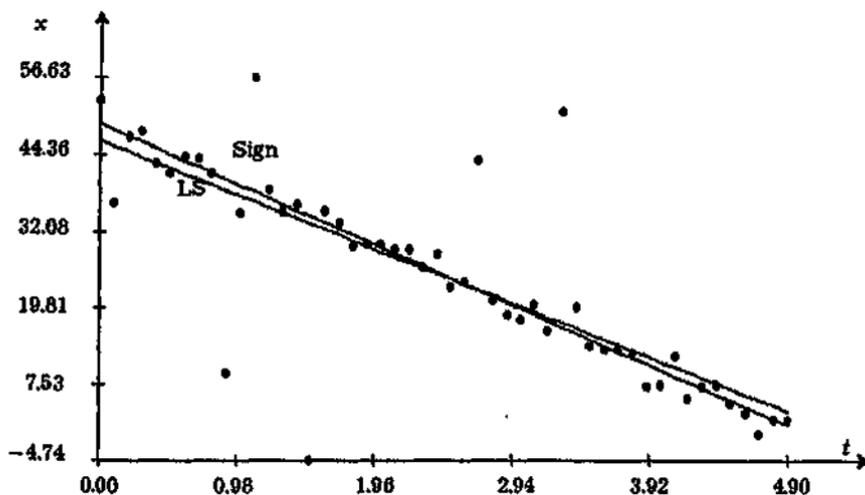


Рис. 3.1.1. Графики функции $x(t) = \hat{\theta}_1 + \hat{\theta}_2 t$, вычисленной методом наименьших квадратов (LS) и знаковым методом (Sign)

На рис. 3.1.2 изображено множество минимумов функции (3.1.5), полученное перебором. Это множество является объединением двух интервалов, концы которых находятся в точках M_1 , M_2 , M_3 соответственно: $M_1 = (49.092, -9.886)$, $M_2 = (49.085, -9.852)$, $M_3 = (49.928, -10.028)$ (они изображены отрезками прямых, пересекающихся в точке M_2). Итерационная знаковая процедура дала значение $(49.09, -9.852)$, что соответствует точке M_2 в пределах вычислительной точности. Было выполнено семь итераций. В качестве начальной итерации была выбрана оценка наименьших квадратов. Оценкой наименьших модулей является точка M_2 . Оценивание по методу наименьших квадратов дало значение $(46.37, -8.827)$, отстоящее от истинного значения $(50, -10)$ значительно дальше, чем значение, полученное знаковым методом.

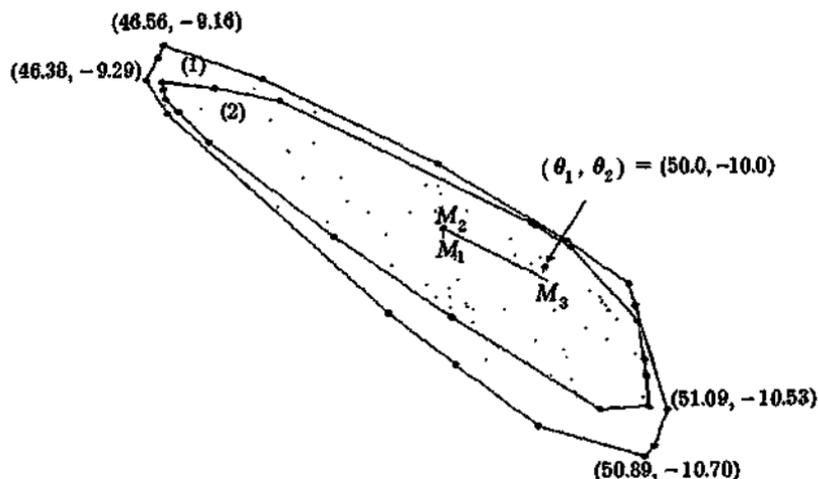


Рис. 3.1.2. Доверительные множества уровней доверия $\gamma = 0.95$ (1) и $\gamma = 0.90$ (2) для примера 1

Нетрудно построить и доверительные множества. Для этого методом Монте-Карло мы определили распределение случайной величины

$$\sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \zeta_i \right)^2$$

($P\{\zeta_i = 1\} = P\{\zeta_i = -1\} = 0.5$) и ее квантили уровней $\gamma = 0.95$ и $\gamma = 0.90$ соответственно. Точечные множества на плоскости, в которых критериальная статистика из (3.1.3) не превосходит величины q_γ , найдены перебором. Эти множества оказались невыпуклыми. На рис. 3.1.2 изображены их выпуклые оболочки. Мелкими точками отмечены точки пересечения прямых, заданных уравнениями $x_i = \theta_1 + c_i \theta_2$, $i = 1, \dots, n$. Отметим, что оценка наименьших квадратов не попала в построенные доверительные множества, в то время как истинную точку $(\theta_1, \theta_2) = (50.0, -10.0)$ они «накрыли».

Пример 2. Моделируем зависимость $x_i = \theta_1 + \theta_2 \sin(t_i) + \xi_i$, $t_i \in [0.41, 12.37]$, $i = 1, \dots, 30$, $\theta_1 = 5$, $\theta_2 = 20$. Распределение ошибок засоренное гауссовское (модель Тьюки) с параметрами $\delta = 0.15$, $\sigma_1 = 3$, $\sigma_2 = 30$. Здесь множеством минимума функции из (3.1.5) является многоугольник, вершины которого — точки плоскости (θ_1, θ_2) с координатами (4.380, 20.895), (4.709, 21.227), (5.272, 20.661), (5.115, 20.464), (4.158, 19.469). На рис. 3.1.3 изображены экспериментальные данные, истинная кривая и две кривые, полученные методом наименьших квадратов (кривая LS) и итерационной знаковой процедурой (Sign).

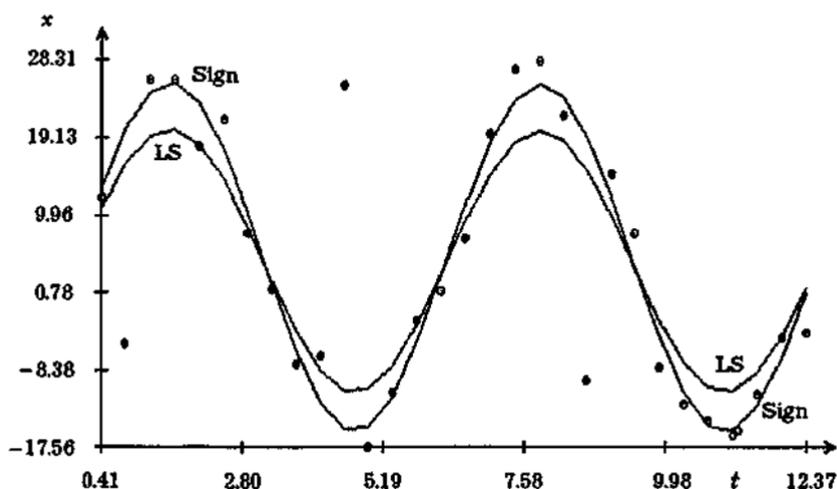


Рис. 3.1.3. Графики функции $x(t) = \hat{\theta}_1 + \hat{\theta}_2 \sin(t)$, вычисленной методом наименьших квадратов и знаковым методом (истинная кривая практически не отличается от кривой, полученной знаковым методом)

Оценка наименьших квадратов (LS) равна (4.40, 15.54). Знаковая оценка, полученная итерационной процедурой (SI) (число итераций равно семи, начальная итерация — оценка наименьших квадратов), равна (4.795, 20.63). Оценка наименьших модулей (LAD) есть (4.16, 19.47). На рис. 3.1.4 изображены область минимума и выпуклые оболочки доверительных множеств с уровнями доверия, не меньшими, чем 0.90 и 0.95 соответственно.

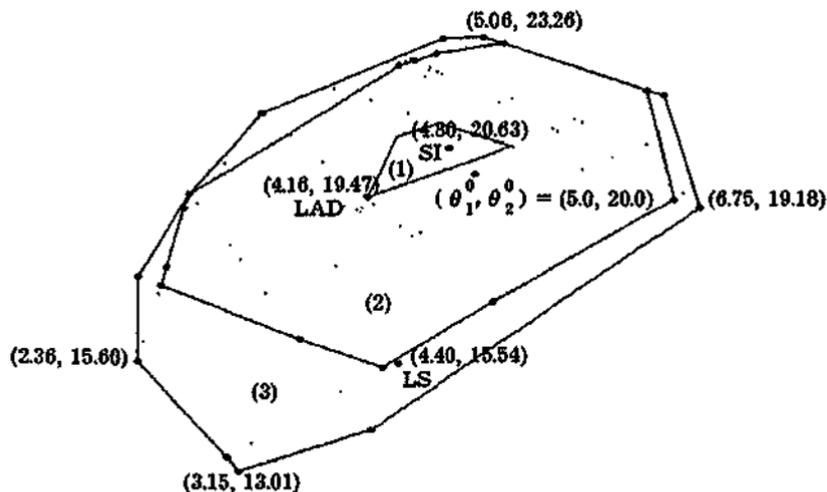


Рис. 3.1.4. Область минимума (1) и выпуклые оболочки доверительных множеств уровней доверия не меньше $\gamma = 0.90$ (2) и $\gamma = 0.95$ (3) для примера 2

Пример 3. В этом примере минимум функции из (3.1.5) достигается в одной точке. Мы моделировали зависимость вида $x_i = \theta_1 + \theta_2 J_0(t_i) + \xi_i$, $t_i \in [0.0, 9.9]$, $i = 1, \dots, 100$, $\theta_1 = 15$, $\theta_2 = 30$, $J_0(t)$ — функция Бесселя нулевого порядка. Распределение ошибок засоренное гауссовское (модель Тьюки) с параметрами $\delta = 0.15$, $\sigma_1 = 1.5$, $\sigma_2 = 15$.

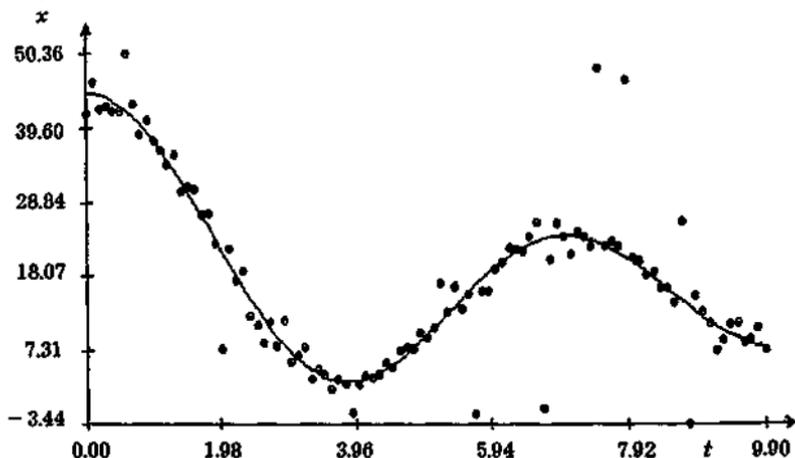


Рис. 3.1.5. График функции $x(t) = \hat{\theta}_1 + \hat{\theta}_2 J_0(t)$, вычисленной знаковым методом

На рис. 3.1.5 изображены экспериментальные данные вместе с кривой, полученной итерационной знаковой процедурой, для которой $(\hat{\theta}_1, \hat{\theta}_2) = (14.63, 29.95)$ (число итераций равно 10, начальное значение совпадает с оценкой наименьших квадратов $(\theta_1^*, \theta_2^*) = (14.6, 30.32)$). Полученное значение совпадает с точкой минимума функции и с оценкой наименьших модулей (LAD).

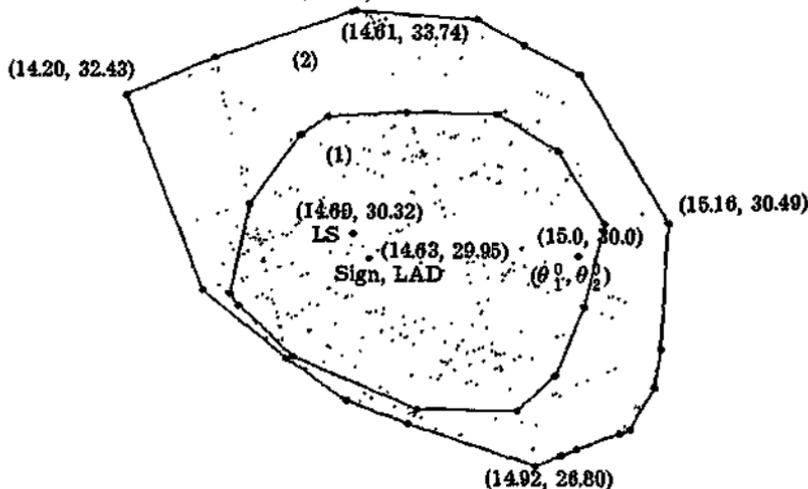


Рис. 3.1.6. Выпуклые оболочки доверительных множеств уровней доверия не меньше $\gamma = 0.90$ (1) и $\gamma = 0.95$ (2) для примера 3

На рис. 3.1.6 изображены выпуклые оболочки доверительных множеств с уровнями доверия, не меньшими, чем 0.90 и 0.95 соответственно. В примере 3 все найденные оценки, в том числе и оценка наименьших квадратов, попали в указанные доверительные множества. Истинное значение параметра $(\theta_1, \theta_2) = (15.0, 30.0)$ также принадлежит этим множествам.

Пример 4. Рассмотрим уже упоминавшуюся во введении линейную модель

$$x_i = \theta_1 + \theta_2 t_i + \theta_3 \sin(\omega t_i) + \xi_i, \quad i = 1, \dots, n, \quad (3.1.7)$$

с известной частотой ω и неизвестными, подлежащими оцениванию параметрами $\theta_1, \theta_2, \theta_3$. (Такие модели могут быть полезны, например, при обнаружении периодического сигнала на фоне шума. Частота сигнала, как правило, известна, а определяемым параметром является его амплитуда. Конечно, в реальной ситуации, помехи ξ_i могут оказаться коррелированными, и их следует описывать каким-либо случайным процессом. Однако в первом приближении можно рассматривать и модель с независимыми ошибками. Условие гауссовости ошибок, как правило, нарушается. Кроме того, данные могут быть засорены импульсными помехами, не имеющими ничего общего с описываемым явлением.) Вычислим знаковые оценки в модели (3.1.7), по-разному задавая распределение случайных ошибок.

Пусть распределение ξ_i несимметрично, но имеет нулевую медиану: при $x < 0$ это гауссовское распределение, при $x > 0$ — распределение Коши (распределение с тяжелыми хвостами):

$$F(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-y^2/(2\sigma^2)} dy, & x \leq 0, \\ \frac{1}{2} + \frac{1}{\pi} \arctan \frac{\sqrt{\pi/2} x}{\sigma}, & x > 0. \end{cases}$$

Здесь $\sigma = 3$. На рис. 3.1.7 приведены наблюдения ($n = 101$), подчиняющиеся модели (3.1.7) при $\omega = 2\pi/7.5$, $t \in [0.3, 20.3]$. Две кривые, полученные знаковым методом и методом наименьших квадратов, помечены словами Sign и LS, модельная кривая — словом Model. Рисунок выполнен в масштабе, в котором не изображен единичный выброс при $t = 18.50$, равный 1277.68. Различие между модельной детерминированной составляющей и кривой Sign незначительно. Модельные значения параметров и их оценки двумя методами приведены в табл. 3.1.1. Стабилизация знаковых оценок наступила после восьми шагов итерации.

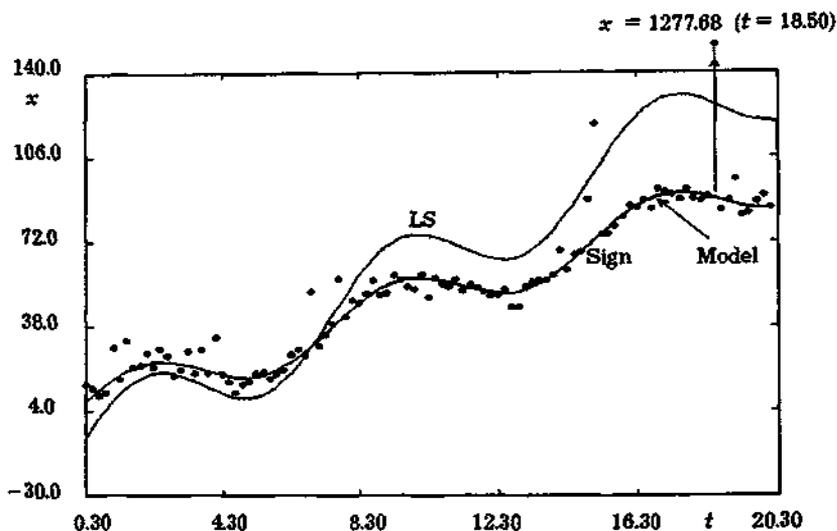


Рис. 3.1.7. Графики функции $x(t) = \hat{\theta}_1 + \hat{\theta}_2 t + \hat{\theta}_3 \sin(\omega t)$, вычисленной методом наименьших квадратов и знаковым методом

Т а б л и ц а 3.1.1

Параметр	Model	LS	Sign
θ_1	3.50	-13.67	3.46
θ_2	4.50	7.41	4.52
θ_3	10.00	16.66	10.12

Пусть распределение ξ_i ($i = 1, \dots, 101$) несимметрично и имеет ненулевую медиану: это засоренное гауссовское распределение с 10-процентным засорением, имеющим положительный сдвиг μ . Функция распределения такого распределения есть

$$F(x) = \frac{1-\varepsilon}{\sqrt{2\pi}\sigma_1} \int_{-\infty}^x e^{-y^2/(2\sigma_1^2)} dy + \frac{\varepsilon}{\sqrt{2\pi}\sigma_2} \int_{-\infty}^x e^{-(y-\mu)^2/(2\sigma_2^2)} dy.$$

Мы положили $\sigma_1 = 3.0$, $\sigma_2 = 10.0$, $\mu = 50$, $\varepsilon = 0.1$.

На рис. 3.1.8 точками показаны наблюдения, подчиняющиеся модели, в которой случайные ошибки следуют указанному выше распределению. Указаны также результаты оценивания знаковым методом и методом наименьших квадратов. Модельные значения параметров и их оценки двумя методами приведены в табл. 3.1.2. Стабилизация знаковых оценок произошла за 31 шаг итерации.

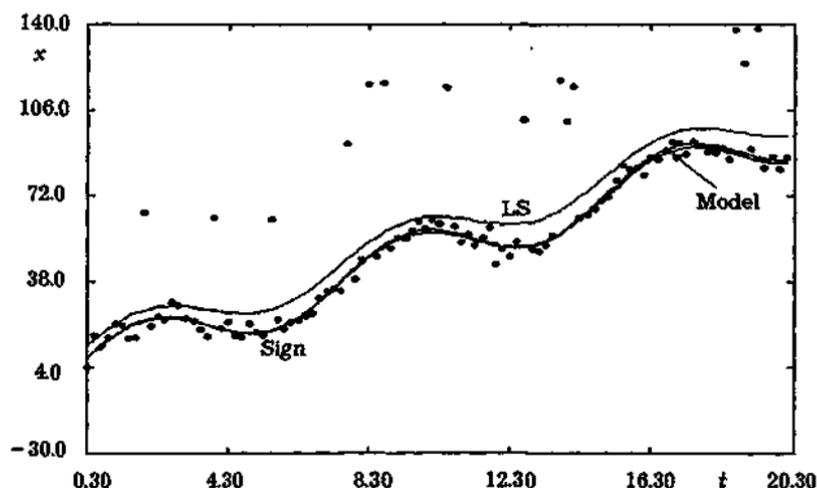


Рис. 3.1.8. Графики функции $x(t) = \hat{\theta}_1 + \hat{\theta}_2 t + \hat{\theta}_3 \sin(\omega t)$, вычисленной методом наименьших квадратов и знаковым методом

Таблица 3.1.2

Параметр	Model	LS	Sign
θ_1	3.50	9.18	3.42
θ_2	4.50	4.68	4.57
θ_3	10.00	8.37	10.40

Как видно из рис. 3.1.7–3.1.8 и табл. 3.1.1–3.1.2, оценки, полученные знаковым методом, значительно ближе к истинным значениям параметров, чем оценки, полученные методом наименьших квадратов.

Пример 5. Рассмотрим модель двухфакторного анализа. В гл. 2, § 2.4 был описан локально оптимальный знаковый критерий для проверки нулевой гипотезы в схеме двухфакторного анализа. Применяя упомянутый критерий и метод знакового оценивания, получим метод оценивания параметров μ , α_i , β_j , $i = 1, \dots, r$, $j = 1, \dots, t$, в схеме (2.4.1)–(2.4.2). Вид критериальной статистики задан выражениями (2.4.9) и (2.4.10), где Z_{ij} , $Z_{i.}$, $Z_{.j}$, $Z_{..}$ определены в (2.4.6) и (2.4.8). Поэтому знаковые оценки неизвестных параметров являются решением экстремальной задачи

$$\min_{\mu, \alpha, \beta} \tilde{T}(\mu, \alpha, \beta) = \min_{\mu, \alpha, \beta} \sum_{i=1}^r \sum_{j=1}^t (\tilde{Z}_{i.} + \tilde{Z}_{.j} - \tilde{Z}_{..})^2$$

при ограничениях

$$\sum_{i=1}^r \alpha_i = 0, \quad \sum_{j=1}^t \beta_j = 0,$$

где $\tilde{Z}_{ij} = \frac{1}{\sqrt{m}} \sum_{k=1}^m \text{sign}(x_{ijk} - \mu - \alpha_i - \beta_j)$, а $\tilde{Z}_{i.}$, $\tilde{Z}_{.j}$ — усреднения по j и i соответственно, $\tilde{Z}_{..}$ — усреднение \tilde{Z}_{ij} по i и j . Функцию $\tilde{T}(\mu, \alpha, \beta)$ можно представить в виде

$$\tilde{T}(\mu, \alpha, \beta) = t \sum_{i=1}^r \tilde{Z}_{i.}^2 + r \sum_{j=1}^t \tilde{Z}_{.j}^2 - rt \tilde{Z}_{..}^2.$$

Оценки параметров находятся с помощью описанного выше итерационного алгоритма. В этом примере была принята следующая схема вычислений.

Шаг 1. При фиксированных β_1, \dots, β_t находим значения $\mu, \alpha_1, \dots, \alpha_r$ как решение задачи

$$(\mu, \alpha) = \arg \min_{\mu, \alpha} \sum_{i=1}^r \tilde{Z}_{i.}^2$$

при условии $\sum_{i=1}^r \alpha_i = 0$.

Шаг 2. После подстановки найденных значений $\alpha_1, \dots, \alpha_r$ в функции $\tilde{Z}_{.j}^2$ находим оценки $\mu, \beta_1, \dots, \beta_t$ как решение задачи

$$(\mu, \beta) = \arg \min_{\mu, \beta} \sum_{j=1}^t \tilde{Z}_{.j}^2$$

при условии $\sum_{j=1}^t \beta_j = 0$.

Шаг 3. Если расстояние в евклидовой метрике между оценками двух соседних итераций превосходит заданное малое число ε , то возвращаемся к шагу 1, в противном случае процесс заканчивается.

Приведем численные результаты оценивания параметров в схеме двухфакторного анализа. Были смоделированы данные в схеме (2.4.1)–(2.4.2) с ошибками ξ_{ijk} , имеющими распределение Тьюки с плотностью распределения:

$$f(x) = \frac{1 - \varepsilon}{\sqrt{2\pi} \sigma_1} e^{-x^2/(2\sigma_1^2)} + \frac{\varepsilon}{\sqrt{2\pi} \sigma_2} e^{-(x-\mu)^2/(2\sigma_2^2)},$$

где $\varepsilon = 0.15$, $\sigma_1 = 0.5$, $\sigma_2 = 5$, $\mu = 7.5$. Здесь число уровней первого фактора $r = 5$; число уровней второго фактора $t = 7$; общее число наблюдений при каждой комбинации уровней факторов $m = 12$.

Результаты оценивания параметров знаковым методом и методом наименьших квадратов приведены в табл. 3.1.3.

Таблица 3.1.3

Параметр	Model	LS	Sign
α_1	0.50	1.22	0.65
α_2	2.00	1.78	1.93
α_3	7.80	7.70	7.82
α_4	-5.30	-5.67	-5.42
α_5	-5.00	-5.03	-4.99
β_1	-6.40	-6.40	-6.36
β_2	7.30	6.78	7.11
β_3	13.50	13.49	13.59
β_4	-7.00	-6.78	-6.93
β_5	5.40	5.54	5.50
β_6	-2.60	-2.25	-2.69
β_7	-10.20	-10.39	-10.24
μ	1.00	2.20	1.09

Гистограммы остатков для двух методов оценивания приведены на рис. 3.1.9 и 3.1.10.

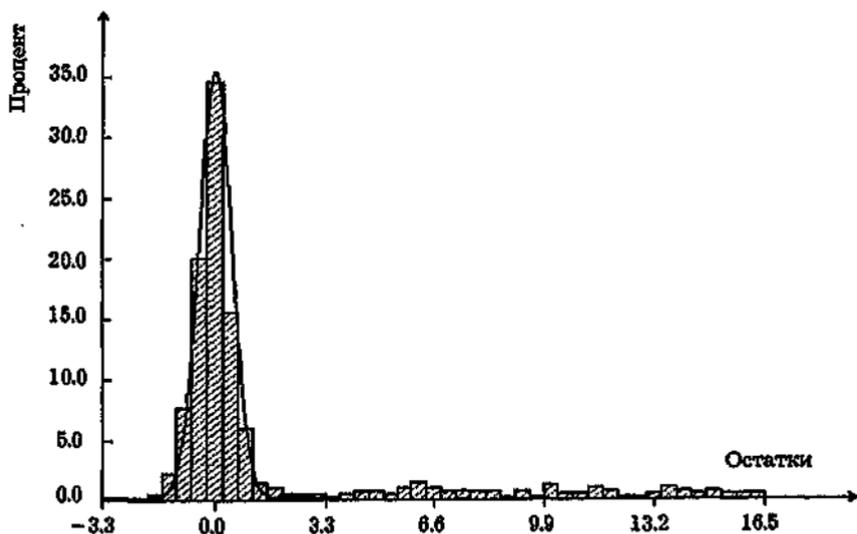


Рис. 3.1.9. Гистограмма остатков и гауссовская плотность основной части распределения при $\sigma_1 = 0.5$ в знаковом методе для примера 5

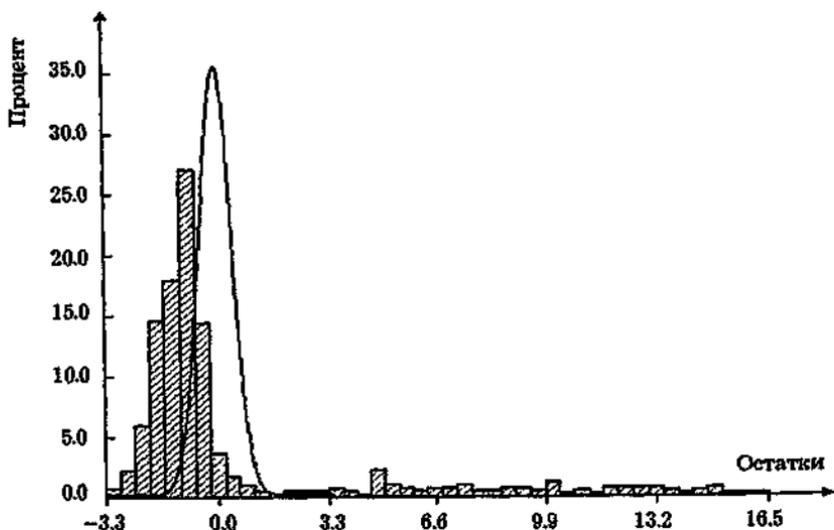


Рис. 3.1.10. Гистограмма остатков и гауссовская плотность основной части распределения при $\sigma_1 = 0.5$ в методе наименьших квадратов для примера 5

Рассмотренные примеры показывают, что знаковый метод оценивания можно успешно применять для линейных моделей с разными типами случайных ошибок: распределенных несимметрично, с тяжелыми хвостами, с отличными от нуля медианами, с большой долей выбросов и т.д. Примеры подтверждают высокую робастность знаковых оценок.

§ 3.2. Знаковое оценивание. Асимптотическая теория

3.2.1. О роли асимптотической теории. Эта теория исследует свойства знаковых выводов при большом числе наблюдений. В этом случае формулы для распределений статистик, их критических значений, для доверительных границ и т.д. становятся более простыми, чем для малых n . Нам уже пришлось говорить об асимптотическом распределении критериальных статистик в гл. 2 при обсуждении проверки гипотез. В отношении оценивания асимптотическая теория состоит из двух основных утверждений: о состоятельности и об асимптотической нормальности знаковых оценок. Напомним определения (в той форме, в какой они здесь необходимы). Мы всюду будем предполагать, что $n \rightarrow \infty$.

Оценка $\hat{\theta}_n$ неизвестного параметра θ называется *состоятельной*, если $\hat{\theta}_n \xrightarrow{P} \theta$ ($\hat{\theta}_n$ сходится к θ по вероятности).

Оценка $\hat{\theta}_n$ неизвестного параметра θ называется *асимптотически нормальной*, если распределение случайной величины $\sqrt{n}(\hat{\theta}_n - \theta)$ сходится к нормальному (с нулевым средним и некоторой ненулевой матрицей ковариаций).

Состоятельность знаковых оценок (в схеме регрессии) будет доказана в п. 3.2.2, а их асимптотическая нормальность — в п. 3.2.3.

На практике состоятельность истолковывают как приближенное равенство $\hat{\theta}_n \approx \theta$ при достаточно большом числе наблюдений n . Здесь есть две неясности: какое число наблюдений достаточно велико и какова точность объявленного приближения? Мы уже отмечали, что первый вопрос не имеет удовлетворительного теоретического решения. При практическом приложении предельных теорем мы обычно руководствуемся смесью теоретических, экспериментальных, эвристических и догматических представлений — в разной пропорции, в зависимости от разработанности предмета. Поэтому в этой книге мы всюду, когда дело касается вычислений, отдаем предпочтение точным методам либо методам с контролируемой точностью (например, методу Монте-Карло).

Вопрос о точности приближения $\hat{\theta}_n$ к θ , т.е. о величине разности $\hat{\theta}_n - \theta$, решается с помощью теоремы об асимптотической нормальности. Зная асимптотическую матрицу ковариаций случайной величины $\sqrt{n}(\hat{\theta}_n - \theta)$, можно указать для θ (а при желании и для линейных функций от θ) доверительные множества, которые его содержат. Асимптотическую матрицу ковариаций для знаковых оценок мы приводим в п. 3.2.4. Правда, доверительные вероятности при этом можно указать лишь приближенно. Считается, что для достаточно больших n точность приближения становится приемлемой (см. сказанное об этом выше). Несмотря на указанные слабости, асимптотические результаты составляют традиционную и важную часть статистической теории.

Представим асимптотические свойства знаковых оценок для схемы линейной регрессии. В отличие от схемы повторной выборки, где единственным ведущим параметром является число наблюдений, в схеме регрессии знания только n недостаточно. К количеству наблюдений должны присоединиться другие характеристики матрицы плана эксперимента. Это значительно усложняет дело. Ради упрощения мы рассмотрим только ограниченные планы (т.е. планы, удовлетворяющие условию 2.1.7) и только одинаково распределенные случайные ошибки. К этим основным предположениям, по мере необходимости, мы будем присоединять и некоторые другие.

3.2.2. Состоятельность знаковых оценок. Покажем, что знаковые оценки параметра θ в модели (2.1.5) состоятельны при некоторых условиях на матрицу плана C и распределение ошибок.

Теорема 3.2.1. Предположим что в модели (2.1.5) случайные ошибки (2.1.3) независимы, одинаково распределены и их общая функция распределения удовлетворяет условиям 2.1.1, 2.1.4, 2.1.5. Относительно матрицы плана C примем условия 2.1.7, 2.1.8.

Тогда знаковые оценки (3.1.1) и (3.1.2) состоятельны.

Доказательство. Истинное значение неизвестного параметра обозначим через θ^0 , оставив за θ смысл переменной величины (как это понимается в математическом анализе).

Состоятельность всех знаковых квадратических оценок доказывается одинаково. Поэтому рассмотрим какую-либо одну из них, например оценку (3.1.1). Сейчас удобно несколько изменить выражение целевой функции (3.1.1), внедрив множитель $1/n$. Затем полезно сделать замену переменных, положив

$$\theta = \theta^0 + t.$$

Тогда $\hat{\theta}_n = \theta^0 + \hat{t}_n$, где

$$\hat{t}_n = \arg \min_{t \in \mathbb{R}^r} \frac{1}{n} \sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) \right)^2. \quad (3.2.1)$$

Остается показать, что $\hat{t}_n \xrightarrow{P} 0$.

Доказательство состоятельности состоит из двух частей.

1. Начинаем с утверждения, что

$$\sup_{|t| \leq T} \left| \frac{1}{n} \sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) - \mathbf{E} \left(\frac{1}{n} \sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) \right) \right| \xrightarrow{P} 0 \quad (3.2.2)$$

для $\alpha = 1, \dots, r$ и любого $T > 0$. Доказательство этого факта вынесено в п. 3.2.5 в виде теоремы 3.2.3 о равномерном законе больших чисел. В силу упомянутой теоремы

$$\arg \min_{|t| \leq T} \left| \frac{1}{n} \sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) \right|^2 \quad (3.2.3)$$

при $n \rightarrow \infty$ приближается (по вероятности) с

$$\arg \min_{|t| \leq T} \left| \frac{1}{n} \mathbf{E} \sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) \right|^2 \quad (3.2.4)$$

при любом значении $T > 0$. С другой стороны,

$$\mathbf{E} \frac{1}{n} \sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) = \frac{1}{n} \sum_{i=1}^n c_{i\alpha} (1 - 2F(c_i t)), \quad (3.2.5)$$

и своего минимального значения, равного нулю, функция из (3.2.4) достигает при $t = 0$ (поскольку $F(0) = 0.5$). Более того, для достаточно больших n эта точка $t = 0$ минимума для (3.2.4) единственная, что мы сейчас покажем.

Предположим противное: существует $t \neq 0$, для которого функция (3.2.5) равна нулю при каждом $\alpha = 1, \dots, r$. Умножив правую

часть выражения (3.2.5) на t_α и просуммировав по α , получим

$$\frac{1}{n} \sum_{i=1}^n (c_i t) (1 - 2 F(c_i t)) = 0. \quad (3.2.6)$$

Легко видеть, что при всяком $u \in \mathbb{R}^1$

$$u(1 - 2 F(u)) \leq 0, \quad (3.2.7)$$

а равенство в (3.2.7) при выполнении условия 2.1.5 достигается только при $u = 0$. Поэтому (3.2.6) может быть равно нулю только в том случае, когда $c_i t = 0$, $i = 1, \dots, n$. В этом случае $\left(\sum_{i=1}^n c_i\right) t = 0$ и,

следовательно, $t^T \left(\frac{1}{n} C^T C\right) t = 0$ для данного $t \neq 0$. При достаточно больших n это противоречит условию 2.1.8.

Таким образом, (3.2.3) сходится к 0 по вероятности при $n \rightarrow \infty$ и любом $T > 0$. Отсюда следует, что для любого $T > 0$

$$\arg \min_{|t| \leq T} \frac{1}{n} \sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) \right) \xrightarrow{\mathbf{P}} 0.$$

2. Покажем, что существуют такие $a > 0$ и $T > 0$, что с вероятностью, сколь угодно близкой к 1,

$$\sum_{\alpha=1}^r \left(\frac{1}{n} \sum_{i=1}^n c_{i\alpha} \text{sign}(\xi_i - c_i t) \right)^2 > a \quad (3.2.8)$$

для достаточно больших n , если $|t| > T$. Тем самым будет доказано, что минимум выражения (3.2.1) при достаточно больших n с вероятностью, сколь угодно близкой к 1, лежит внутри некоторого компакта $|t| \leq T$. В соединении с результатом первой части неравенство (3.2.8) доказывает состоятельность знаковой оценки.

Докажем (3.2.8). Рассмотрим вектор $\varphi(t) = (\varphi_1(t), \dots, \varphi_r(t))$, где

$$\varphi_\alpha(t) = \frac{1}{n} \sum_{i=1}^n c_{i\alpha} \text{sign}(-\xi_i + c_i t).$$

Введем единичный вектор $\tau = t/|t|$ и покажем, что существуют $a > 0$ и $T > 0$ такие, что для достаточно больших n с вероятностью, сколь угодно близкой к 1, для $|t| \geq T$

$$|\varphi(t)\tau|^2 > a,$$

т.е. покажем, что

$$\frac{1}{|t|^2} \left(\frac{1}{n} \sum_{i=1}^n (c_i t) \operatorname{sign}(c_i t - \xi_i) \right)^2 > \alpha. \quad (3.2.9)$$

Из этого неравенства и неравенства Коши—Шварца будет следовать (3.2.8). Заметим, что

$$(c_i t) \operatorname{sign}(c_i t - \xi_i) = |c_i t|, \text{ если } |c_i t| > |\xi_i|.$$

Воспользуемся неравенством $|A + B| > |A| - |B|$ и запишем цепочку соотношений

$$\begin{aligned} \left| \frac{1}{n} \sum_{i=1}^n (c_i t) \operatorname{sign}(c_i t - \xi_i) \right| &> \frac{1}{n} \sum_{\{i: |c_i t| > |\xi_i|\}} |c_i t| - \frac{1}{n} \sum_{\{i: |c_i t| \leq |\xi_i|\}} |c_i t| = \\ &= \frac{1}{n} \sum_{i=1}^n |c_i t| - \frac{2}{n} \sum_{\{i: |c_i t| \leq |\xi_i|\}} |c_i t|. \end{aligned} \quad (3.2.10)$$

Оценим сверху второе слагаемое в правой части (3.2.10). Заметим, что для любого $R > 0$

$$|c_i t| I(|c_i t| \leq |\xi_i|) \leq R I(|\xi_i| \leq R) + |c_i t| I(|\xi_i| > R).$$

Выберем R так, чтобы

$$\gamma = 1 - 2\mathbf{P}\{|\xi_i| > R\} > 0.$$

В результате для второго слагаемого в правой части (3.2.10) получаем оценку

$$\frac{2}{n} \sum_{\{i: |c_i t| \leq |\xi_i|\}} |c_i t| \leq \frac{2}{n} \sum_{i=1}^n \left\{ R I(|\xi_i| \leq R) + |c_i t| I(|\xi_i| > R) \right\}. \quad (3.2.11)$$

В силу условия 2.1.7 и по закону больших чисел правая часть (3.2.11) при $n \rightarrow \infty$ есть

$$2R \mathbf{P}\{|\xi_i| \leq R\} + 2\mathbf{P}\{|\xi_i| > R\} \frac{1}{n} \sum_{i=1}^n |c_i t| + o_p(1).$$

Воспользовавшись этим, продолжим (3.2.10) и получим, что для любого $\varepsilon > 0$ с вероятностью, сколь угодно близкой к 1, для достаточно

больших n

$$\left| \frac{1}{n} \sum_{i=1}^n (c_i t) \operatorname{sign}(c_i t - \xi_i) \right| \geq \\ \geq -2R \mathbf{P}\{|\xi_i| \leq R\} + (1 - 2\mathbf{P}\{|\xi_i| > R\}) \frac{1}{n} \sum_{i=1}^n |c_i t| - \varepsilon. \quad (3.2.12)$$

При условиях 2.1.7, 2.1.8 мы имеем

$$\frac{1}{n} \sum_{i=1}^n |c_i t|^2 \leq K \left(\frac{1}{n} \sum_{i=1}^n |c_i t| \right) |t|, \quad (3.2.13)$$

а также

$$\frac{1}{n} \sum_{i=1}^n |c_i t|^2 = \frac{1}{n} \sum_{i=1}^n (c_i t)^T (c_i t) = t^T \left(\frac{1}{n} C^T C \right) t \rightarrow t^T \Sigma t \geq k |t|^2,$$

где k — некоторое положительное число. Следовательно, при достаточно больших n

$$\frac{1}{n} \sum_{i=1}^n |c_i t|^2 \geq \frac{k}{2} |t|^2. \quad (3.2.14)$$

Из (3.2.13) и (3.2.14) следует, что для достаточно больших n

$$\frac{1}{n} \sum_{i=1}^n |c_i t| \geq \frac{k}{2K} |t|. \quad (3.2.15)$$

Вернувшись к (3.2.12), мы получаем с учетом (3.2.15), что при достаточно больших n со сколь угодно близкой к 1 вероятностью

$$\left| \frac{1}{n} \sum_{i=1}^n (c_i t) \operatorname{sign}(c_i t - \xi_i) \right| > \tilde{k} |t| - 2R \mathbf{P}\{|\xi_i| \leq R\} - \varepsilon, \quad (3.2.16)$$

где $\tilde{k} = \frac{k}{2K} (1 - 2\mathbf{P}\{|\xi_i| > R\})$. Вернувшись к (3.2.9), отсюда получаем для достаточно больших $|t|$ и достаточно больших n , что

$$|\varphi(t)\tau|^2 > \frac{(\tilde{k}|t| - 2R \mathbf{P}\{|\xi_i| \leq R\})^2}{|t|^2} \geq a > 0$$

со сколь угодно близкой к 1 вероятностью, что и требовалось доказать. Состоятельность знаковой оценки (3.1.1) доказана. \square

3.2.3. Асимптотическая нормальность знаковых оценок. Покажем, что при некоторых предположениях о распределении случайных ошибок (несколько более сильных, чем в теореме 3.2.1) знаковые оценки (3.1.1), (3.1.2) или основанные на других квадратичных знаковых статистиках асимптотически нормальны. Для всех таких знаковых оценок исследование проводится одинаково, и все они имеют одинаковые асимптотические распределения. Поэтому выберем какую-либо одну из них. Например, оценку (3.1.1). Будет доказана следующая теорема.

Теорема 3.2.2. *Предположим, что в модели (2.1.5) случайные ошибки (2.1.3) независимы, одинаково распределены и их общая функция распределения удовлетворяет условиям 2.1.1, 2.1.2, 2.1.6. Относительно матрицы плана C примем условия 2.1.7, 2.1.8.*

Тогда знаковая оценка (3.1.1) представима в виде

$$\hat{\theta}_n = \theta^0 + \frac{1}{2f(0)} (C^T C)^{-1} C^T \zeta + o_p\left(\frac{1}{\sqrt{n}}\right), \quad (3.2.17)$$

где $\zeta = (\zeta_1, \dots, \zeta_n)^T$, $\zeta_i = \text{sign } \xi_i$, так что случайные величины ζ_i , $i = 1, \dots, n$, независимы и $P\{\zeta_i = 1\} = P\{\zeta_i = -1\} = 0.5$. Отсюда следует, что знаковая оценка $\hat{\theta}_n$ асимптотически нормальна с параметрами θ^0 и $\frac{1}{(2f(0))^2} \Sigma^{-1}$. В качестве асимптотической

матрицы ковариаций можно взять также $\frac{1}{(2f(0))^2} \left(\frac{1}{n} C^T C\right)^{-1}$.

Доказательство. Мы уже доказали (теорема 3.2.1), что в этих условиях $\hat{\theta}_n \xrightarrow{P} \theta^0$. Изучим более подробно характер этой сходимости. Сделаем замену переменных, положив $\tau = \sqrt{n}(\theta - \theta^0)$. При такой замене $\hat{\theta}_n = \theta^0 + \frac{1}{\sqrt{n}} \hat{\tau}_n$, где

$$\hat{\tau}_n = \arg \min_{\tau \in \mathbb{R}^r} \sum_{\alpha=1}^r \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n c_{i\alpha} \text{sign} \left(\xi_i - \frac{c_i \tau}{\sqrt{n}} \right) \right)^2. \quad (3.2.18)$$

Доказательство асимптотической нормальности $\hat{\tau}_n$ состоит из двух частей. Сначала (пункт 1°) покажем, что для любого $0 < \gamma < 1/4$

$$\inf_{\{\tau: |\tau| \geq n^\gamma\}} \sum_{\alpha=1}^r \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n c_{i\alpha} \text{sign} \left(\xi_i - \frac{c_i \tau}{\sqrt{n}} \right) \right)^2 \xrightarrow{P} \infty. \quad (3.2.19)$$

Так как при $\tau = 0$ целевая функция в (3.2.19) в силу центральной предельной теоремы ограничена по вероятности, то

$$P\{|\hat{\tau}_n| < n^\gamma\} \rightarrow 1.$$

Поэтому $\tilde{\tau}_n$ по вероятности эквивалентна

$$\tilde{\tau}_n = \arg \min_{\{\tau: |\tau| < n^\gamma\}} \sum_{\alpha=1}^r \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n c_{i\alpha} \operatorname{sign} \left(\xi_i - \frac{c_{i\alpha} \tau}{\sqrt{n}} \right) \right)^2. \quad (3.2.20)$$

Далее (пункт 2°) мы исследуем $\tilde{\tau}_n$. Исследование опирается на теорему 3.2.4 о равномерной линейности, (см. п. 3.2.5), из которой следует, что

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n c_i \operatorname{sign} \left(\xi_i - \frac{c_i \tau}{\sqrt{n}} \right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n c_i \left(\operatorname{sign} \xi_i - 2 f(0) \frac{c_i \tau}{\sqrt{n}} \right) + \omega_n(\tau), \quad (3.2.21)$$

где для любого $0 < \gamma < 1/4$

$$\sup_{|\tau| < n^\gamma} |\omega_n(\tau)| \xrightarrow{\mathbf{P}} 0.$$

Эта линейнизация целевой функции в (3.2.20) позволяет изучить асимптотические свойства $\tilde{\tau}_n$ и завершить доказательство теоремы.

1°. Пусть $e \in \mathbb{R}^r$ — произвольный вектор единичной сферы S^{r-1} .

Положим $\tau = |\tau|e$ и умножим скалярно на e вектор

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n c_i \operatorname{sign} \left(\xi_i - \frac{c_i \tau}{\sqrt{n}} \right).$$

Получим, положив $a_i = c_i e$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i |\tau|}{\sqrt{n}} \right). \quad (3.2.22)$$

По неравенству Коши—Шварца

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i |\tau|}{\sqrt{n}} \right) \right)^2 \leq \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n c_i \operatorname{sign} \left(\xi_i - \frac{c_i \tau}{\sqrt{n}} \right) \right|^2. \quad (3.2.23)$$

Поэтому для (3.2.19) достаточно показать, что

$$\sup_{\{\tau: |\tau| \geq n^\gamma\}} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i |\tau|}{\sqrt{n}} \right) \right)^2 \xrightarrow{\mathbf{P}} \infty. \quad (3.2.24)$$

Заметим, что относительно $|\tau|$ функция (3.2.22) монотонно не воз-

растает. Поэтому для $|\tau| \geq n^\gamma$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i |\tau|}{\sqrt{n}} \right) \leq \frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i n^\gamma}{\sqrt{n}} \right). \quad (3.2.25)$$

Легко видеть, что

$$v_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i n^\gamma}{\sqrt{n}} \right) \xrightarrow{P} -\infty. \quad (3.2.26)$$

Действительно, Dv_n ограничена, а

$$E v_n = -\frac{2n^\gamma}{n} \sum_{i=1}^n a_i^2 f(\kappa_i) \rightarrow -\infty.$$

(Здесь κ_i — некоторая точка между 0 и $a_i n^{-(1/2-\gamma)}$, $i = 1, \dots, n$. Так как $f(\kappa_i) \rightarrow f(0) > 0$ равномерно по $i = 1, \dots, n$ и для всякого $e \in \mathbb{S}^{r-1}$ в силу условия 2.1.8

$$\frac{1}{n} \sum_{i=1}^n a_i^2 \rightarrow e^T \Sigma e \geq \text{const} > 0,$$

то $E v_n \rightarrow -\infty$.) Из (3.2.25), (3.2.26) заключаем, что

$$\inf_{\{\tau: |\tau| \geq n^\gamma\}} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign} \left(\xi_i - \frac{a_i |\tau|}{\sqrt{n}} \right) \right)^2 \geq v_n^2.$$

Отсюда следует (3.2.24) и (3.2.19).

2°. Используя (3.2.21), заменим (3.2.20) на

$$\min_{\{\tau: |\tau| < n^\gamma\}} \sum_{\alpha=1}^r \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n c_{i\alpha} \left(\operatorname{sign} \xi_i - 2f(0) \frac{c_i \tau}{\sqrt{n}} \right) + \omega_n(\tau) \right)^2, \quad (3.2.27)$$

где добавка $\omega_n(\tau)$ равномерно мала (по вероятности) на множестве $\{\tau: |\tau| < n^\gamma\}$. Поэтому при $n \rightarrow \infty$ решение задачи (3.2.27) асимптотически эквивалентно решению задачи

$$\min_{\{\tau: |\tau| < n^\gamma\}} \sum_{\alpha=1}^r \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n c_{i\alpha} \left(\operatorname{sign} \xi_i - 2f(0) \frac{c_i \tau}{\sqrt{n}} \right) \right)^2. \quad (3.2.28)$$

Если отвлечься от ограничения $|\tau| < n^\gamma$, то решение (3.2.28) есть

решение системы линейных уравнений

$$\sum_{i=1}^n c_{i\alpha} \left(\text{sign } \xi_i - 2f(0) \frac{c_i \tau}{\sqrt{n}} \right) = 0, \quad \alpha = 1, \dots, r. \quad (3.2.29)$$

Чтобы его записать, введем n -мерный вектор ζ с координатами $\zeta_i = \text{sign } \xi_i$, $i = 1, \dots, n$. Теперь (3.2.29) можно записать в векторно-матричной форме

$$\frac{1}{\sqrt{n}} C^T \zeta = 2f(0) \left(\frac{1}{n} C^T C \right) \tau, \quad (3.2.30)$$

откуда его решение, скажем, τ_n^* , есть

$$\tau_n^* = \frac{1}{2f(0)} \left(\frac{1}{n} C^T C \right)^{-1} \left(\frac{1}{\sqrt{n}} C^T \zeta \right). \quad (3.2.31)$$

Мы уже отмечали в § 2.3, что при условиях 2.1.1, 2.1.7, 2.1.8 случайный вектор $\frac{1}{\sqrt{n}} C \zeta^T$ асимптотически распределен как $N(0, \Sigma)$.

При этих же условиях величина τ_n^* асимптотически распределена как

$$N\left(0, \frac{1}{(2f(0))^2} \Sigma^{-1}\right), \text{ или } N\left(0, \frac{1}{(2f(0))^2} \left(\frac{1}{n} C^T C\right)^{-1}\right).$$

Нас, однако, интересует решение задачи (3.2.28), которая отличается от (3.2.29) присутствием ограничения $|\tau| < n^\gamma$. Покажем, что это различие несущественно. Поскольку τ_n^* имеет невырожденное асимптотическое распределение, мы можем выбрать с помощью этого последнего постоянную A так, чтобы для достаточно больших n со сколь угодно близкой к 1 вероятностью выполнялось неравенство $|\tau_n^*| \leq A$. В этом случае τ_n^* становится и решением (3.2.28). А поскольку минимум в (3.2.28) достигается во внутренней точке, τ_n^* при этом совпадает с $\tilde{\tau}_n$. Это завершает доказательство теоремы 3.2.2. \square

З а м е ч а н и е. В линейной модели (2.1.1) можно рассматривать различные параметризации. Например, вектор $l \in \mathbb{L}$ может быть представлен как

$$l = D\tau,$$

где D матрица размера $n \times r$, τ — вектор-столбец из \mathbb{R}^r . Существует некоторая невырожденная матрица, скажем, G (размера $r \times r$), для которой

$$\theta = G\tau, \quad D = CG.$$

Имея две параметрические записи исходной линейной модели (2.1.1), мы можем получить знаковые оценки как для θ , так и для τ . Обозначим их через $\hat{\theta}_n$ и $\hat{\tau}_n$ соответственно. Из теоремы 3.2.2 следует,

что

$$\hat{\theta}_n = G \hat{\tau}_n + o_p\left(\frac{1}{\sqrt{n}}\right).$$

3.2.4. Асимптотическая ковариация знаковых оценок. Для построения асимптотических доверительных множеств для θ^0 , как о том говорилось в п. 3.2.1, надо знать асимптотическую матрицу ковариаций знаковой оценки $\hat{\theta}_n$. Чтобы воспользоваться для этого результатами теоремы 3.2.2, достаточно указать состоятельную оценку для $f(0)$. Мы предложим такую оценку, опираясь на (3.2.21). Пусть $\hat{\theta}_n$ — знаковая оценка θ^0 в модели (2.1.5) либо иная \sqrt{n} -состоятельная оценка θ^0 ,

$$\hat{\theta}_n = \theta^0 + \frac{1}{\sqrt{n}} \hat{\tau}_n.$$

Выберем произвольный вектор $e \in \mathbb{S}^{r-1}$. Положим $a_i = c_i e$ и для $h \in \mathbb{R}^1$ введем статистику

$$G_n(h) := \frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \operatorname{sign}\left(x_i - e_i \hat{\theta}_n + \frac{a_i h}{\sqrt{n}}\right).$$

Умножив (3.2.1) на e , получим, что

$$G_n(h) = \frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \left(\operatorname{sign} \xi_i - 2f(0) \frac{c_i \hat{\tau}_n}{\sqrt{n}} + 2f(0) \frac{a_i h}{\sqrt{n}} \right) + o_p(1),$$

причем $o_p(1) \xrightarrow{P} 0$ равномерно по h , если $|h| < n^\gamma$, $0 < \gamma < 1/4$. Очевидно,

$$\frac{1}{2}(G_n(h) - G_n(-h)) = 2f(0)h \left(\frac{1}{n} \sum_{i=1}^n a_i^2 \right) + o_p(1).$$

Отсюда для $(2f(0))^{-1}$ получаем состоятельную оценку

$$\frac{e^T (C^T C) e}{\frac{1}{2}(G_n(h) - G_n(-h))} h \xrightarrow{P} (2f(0))^{-1},$$

если положить $h = h(n) = n^\gamma$, γ произвольное, $0 < \gamma < 1/4$. В качестве e мы рекомендуем взять собственный вектор матрицы $C^T C$, которому соответствует наибольшее собственное значение.

3.2.5. Равномерный закон больших чисел. В этом и последующих пунктах мы рассматриваем схему серий, когда элементы матрицы плана $C = \|c_{i\alpha}\|$, $\alpha = 1, \dots, r$, $i = 1, \dots, n$, и элементы последовательности d_i , $i = 1, \dots, n$, зависят также и от n . Мы, однако, не

станем отмечать эту зависимость, чтобы не усложнять последующих формул.

Теорема 3.2.3. *Случайная функция*

$$U_n(t) = \frac{1}{n} \sum_{i=1}^n d_i (\text{sign}(\xi_i + c_i t) - \mathbf{E} \text{sign}(\xi_i + c_i t))$$

переменного вектора $t = (t_1, \dots, t_r)^T$ сходится к нулю по вероятности при $n \rightarrow \infty$ равномерно относительно t на любом компакте $|t| \leq C$, где C — произвольная постоянная, т.е. для любого $\varepsilon > 0$

$$\mathbf{P}\left\{ \sup_{|t| \leq C} |U_n(t)| > \varepsilon \right\} \rightarrow 0 \quad \text{при } n \rightarrow \infty, \quad (3.2.32)$$

при следующих условиях:

i) последовательности $\{d_i, i = 1, 2, \dots, n\}$, $\{c_{i\alpha}, i = 1, \dots, n\}$, $\alpha = 1, \dots, r$, ограничены, т.е. существует постоянная $K > 0$ такая, что

$$\max_{1 \leq i \leq n} (|d_i|, |c_{i1}|, \dots, |c_{ir}|) \leq K;$$

ii) случайные величины ξ_1, ξ_2, \dots независимы, одинаково распределены, и их функция распределения удовлетворяет условию 2.1.4, т.е. существует такая постоянная $L > 0$, что

$$|F(u_1) - F(u_2)| < L |u_1 - u_2|,$$

где $F(u) = \mathbf{P}\{\xi_i < u\}$, $i = 1, 2, \dots$

Доказательство. Пусть $y_i(t) = \text{sign}(\xi_i + c_i t) - \mathbf{E} \text{sign}(\xi_i + c_i t)$. Заметим, что $\mathbf{E} \text{sign}(\xi_i + c_i t) = 1 - 2F(-c_i t)$. Легко видеть, что $U_n(t) \rightarrow 0$ по вероятности при любом фиксированном t . Действительно,

$$\mathbf{E}(U_n(t))^2 = \mathbf{D}U_n(t) = \frac{1}{n^2} \sum_{i=1}^n d_i^2 \mathbf{D} \text{sign}(\xi_i + c_i t) \leq \frac{1}{n^2} \sum_{i=1}^n d_i^2,$$

так как

$$\mathbf{D} \text{sign}(\xi_i + c_i t) = \mathbf{E}(\text{sign}(\xi_i + c_i t))^2 - (\mathbf{E} \text{sign}(\xi_i + c_i t))^2 \leq 1.$$

В силу условия теоремы $\frac{1}{n} \sum_{i=1}^n d_i^2 \leq K^2$, и потому $\mathbf{E}(U_n(t))^2 \rightarrow 0$.

Отсюда следует сходимость $U_n(t)$ к нулю по вероятности.

Теперь докажем (3.2.32). Рассмотрим разбиение пространства \mathbb{R}^r на конгруэнтные кубы с помощью кубической решетки. Пусть V — произвольный замкнутый куб из этого разбиения, h — длина его диагонали, T — конечная совокупность таких кубов, которые покрывают компакт $\{t: |t| \leq C\}$.

Рассмотрим произвольный куб V , $V \in T$. Заметим, что линейная функция $c_i t$ переменного t достигает максимального и минимального значений на кубе V в некоторых диагонально расположенных его вершинах. Обозначим эти вершины через μ_i и λ_i соответственно. Поскольку функции $\text{sign } x$ и $1 - 2F(-x)$ монотонно неубывают по $x \in \mathbb{R}^1$, можно утверждать, что при $t \in V$

$$\text{sign}(\xi_i + c_i \lambda_i) \leq \text{sign}(\xi_i + c_i t) \leq \text{sign}(\xi_i + c_i \mu_i),$$

$$1 - 2F(-c_i \lambda_i) \leq 1 - 2F(-c_i t) \leq 1 - 2F(-c_i \mu_i).$$

Пусть

$$d_i^+ = \begin{cases} d_i, & \text{если } d_i > 0 \\ 0, & \text{если } d_i \leq 0, \end{cases} \quad d_i^- = \begin{cases} 0, & \text{если } d_i > 0 \\ -d_i, & \text{если } d_i \leq 0. \end{cases}$$

Теперь при $t \in V$ для $U_n(t)$ можно указать оценки сверху и снизу:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n d_i y_i(t) &\leq \frac{1}{n} \sum_{i=1}^n d_i^+ (\text{sign}(\xi_i + c_i \mu_i) - (1 - 2F(-c_i \mu_i))) - \\ &\quad - \frac{1}{n} \sum_{i=1}^n d_i^- (\text{sign}(\xi_i + c_i \lambda_i) - (1 - 2F(-c_i \lambda_i))) + \\ &\quad + \frac{1}{n} \sum_{i=1}^n d_i^+ ((1 - 2F(-c_i \mu_i)) - (1 - 2F(-c_i \lambda_i))) - \\ &\quad - \frac{1}{n} \sum_{i=1}^n d_i^- ((1 - 2F(-c_i \lambda_i)) - (1 - 2F(-c_i \mu_i))) = \\ &= A_1 - A_2 + A_3 - A_4. \end{aligned}$$

Оценка снизу выглядит аналогично:

$$\frac{1}{n} \sum_{i=1}^n d_i y_i(t) \geq A'_1 - A'_2 + A'_3 - A'_4,$$

где

$$\begin{aligned} A'_1 &= \frac{1}{n} \sum_{i=1}^n d_i^+ (\text{sign}(\xi_i + c_i \lambda_i) - (1 - 2F(-c_i \lambda_i))), \\ A'_2 &= \frac{1}{n} \sum_{i=1}^n d_i^- (\text{sign}(\xi_i + c_i \mu_i) - (1 - 2F(-c_i \mu_i))), \end{aligned}$$

$$A_3' = \frac{1}{n} \sum_{i=1}^n d_i^+ ((1 - 2F(-c_i \lambda_i)) - (1 - 2F(-c_i \mu_i))),$$

$$A_4' = \frac{1}{n} \sum_{i=1}^n d_i^- ((1 - 2F(-c_i \mu_i)) - (1 - 2F(-c_i \lambda_i))).$$

Рассмотрим статистическую сумму A_1 . Обозначим через $\nu_1, \dots, \nu_k, \dots$, где $k = 1, \dots, 2^{r+1}$, вершины куба V . Положим

$$a_{ik} = \begin{cases} 1, & \text{если } \mu_i = \nu_k, \\ 0 & \text{в противном случае.} \end{cases}$$

Теперь

$$A_1 = \sum_{i=1}^{2^{r+1}} \frac{1}{n} \sum_{k=1}^n a_{ik} d_i^+ y_i(\nu_k).$$

В силу доказанной выше сходимости в каждой точке, для каждого k

$$\sum_{i=1}^n a_{ik} d_i^+ y_i(\nu_k) \xrightarrow{P} 0.$$

Поэтому $A_1 \xrightarrow{P} 0$. Аналогично можно показать, что $A_2, A_1', A_2' \xrightarrow{P} 0$. Обозначим через Ω_n событие, состоящее в том, что для каждого $V \in T$ величины A_1, A_2, A_1', A_2' по абсолютному значению не превосходят δ . В силу доказанного $P\{\Omega_n\} \rightarrow 1$ для произвольных фиксированных $h > 0, \delta > 0$.

Теперь обратимся к суммам A_3, A_4, A_3', A_4' . Заметим, что в силу условий теоремы

$$|F(-c_i \mu_i) - F(-c_i \lambda_i)| \leq Lh$$

для любого i и для любого куба $V \in T$. Поэтому $|A_3|, |A_4|, |A_3'|, |A_4'| \leq 2LKh$ для любого куба $V \in T$. Выберем h столь малым, что $2LKh < \delta$.

Вернемся к (3.2.32). Заметим, что

$$\begin{aligned} P\left\{ \sup_{|t| \leq C} |U_n(t)| < \varepsilon \right\} &= P\left\{ \max_{V \in T} \sup_{t \in V} |U_n(t)| < \varepsilon \right\} \geq \\ &\geq P\left\{ \max_{V \in T} \sup_{t \in V} |U_n(t)| < \varepsilon, \Omega_n \right\}. \end{aligned}$$

Выберем $\delta < \varepsilon/4$. Если произошло событие Ω_n , то для каждого $V \in T$

$$\sup_{t \in V} |U_n(t)| < 4\delta,$$

т.е. событие $\left\{ \sup_{|t| \leq C} |U_n(t)| < \varepsilon \right\}$ становится следствием события Ω_n , а потому

$$\left\{ \sup_{|t| \leq C} |U_n(t)| < \varepsilon \right\} = P\{\Omega_n\} \rightarrow 1,$$

что и требовалось доказать. \square

3.2.6. Теорема о равномерной линейности.

Теорема 3.2.4. *Случайная функция*

$$X_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i \left(\text{sign} \left(\xi_i + \frac{c_i t}{\sqrt{n}} \right) - \text{sign} \xi_i - 2 f(0) \frac{c_i t}{\sqrt{n}} \right) \quad (3.2.33)$$

вектора $t = (t_1, \dots, t_r)^T$ сходится к нулю по вероятности при $n \rightarrow \infty$ равномерно относительно t , изменяющегося на множестве $\{|t| \leq \text{const} \cdot n^\gamma\}$, для произвольного $0 < \gamma < 1/4$, т.е. для любого $\varepsilon > 0$

$$P\left\{ \sup_{|t| \leq \text{const} \cdot n^\gamma} |X_n(t)| > \varepsilon \right\} \rightarrow 0$$

при следующих условиях:

i) последовательности $\{d_i, i = 1, 2, \dots\}$, $\{c_{i\alpha}, i = 1, 2, \dots\}$, где $\alpha = 1, \dots, r$, ограничены, т.е. существует постоянная $K > 0$ такая, что при всех n

$$\max_{1 \leq i \leq n} (|d_i|, |c_{i1}|, \dots, |c_{ir}|) < K;$$

ii) случайные величины $\xi_i, i = 1, 2, \dots$, независимы и одинаково распределены, причем их функция распределения удовлетворяет условиям 2.1.1, 2.1.2, 2.1.6.

Эту теорему мы часто будем применять в такой форме:

При перечисленных выше условиях i), ii)

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n d_i \text{sign} \left(\xi_i + \frac{c_i t}{\sqrt{n}} \right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i \left(\text{sign} \xi_i + 2 f(0) \frac{c_i t}{\sqrt{n}} \right) + o_p(1), \quad (3.2.34)$$

где $o_p(1)$ сходится к нулю по вероятности равномерно относительно t , $|t| < \text{const} \cdot n^\gamma$, для произвольного $0 < \gamma < 1/4$.

Доказательство. Доказательство теоремы состоит из двух этапов. Сначала (пункт 1°) мы доказываем сходимость к нулю в каждой точке; доказательство равномерной сходимости составляет пункт 2°.

1°. Возьмем в \mathbb{R}^r куб с центром в начале координат, ребро которого имеет длину Cn^γ , $0 < \gamma < 1/4$. Разобьем этот куб на более мелкие конгруэнтные кубы с помощью кубической решетки. Длину

диагонали этого разбиения $h = h(n) \rightarrow 0$ будем выбирать далее в зависимости от n . Пусть $V = V(\tau)$ обозначает такой произвольный малый куб, τ — его центр. Пусть $T = T(n)$ обозначает множество таких центров. Для доказательства равномерной сходимости достаточно показать, что

$$\mathbf{P}\left\{\sup_{\tau \in T(n)} \sup_{t \in V(\tau)} |X_n(t)| > \varepsilon\right\} \rightarrow 0 \quad \text{при } n \rightarrow \infty.$$

Ясно, что

$$\begin{aligned} \mathbf{P}\left\{\sup_{\tau \in T(n)} \sup_{t \in V(\tau)} |X_n(t)| > \varepsilon\right\} &\leq \sum_{\tau \in T(n)} \mathbf{P}\left\{\sup_{t \in V(\tau)} |X_n(t)| > \varepsilon\right\} \leq \\ &\leq \sum_{\tau \in T(n)} \mathbf{P}\left\{\sup_{t \in V(\tau)} X_n(t) > \varepsilon\right\} + \\ &\quad + \sum_{\tau \in T(n)} \mathbf{P}\left\{\inf_{t \in V(\tau)} X_n(t) < -\varepsilon\right\}. \quad (3.2.35) \end{aligned}$$

Число слагаемых в каждой из двух сумм есть $O(n^{\gamma r} h^{-r})$.

Оценим сверху $\mathbf{P}\left\{\sup_{t \in V(\tau)} X_n(t) > \varepsilon\right\}$ и $\mathbf{P}\left\{\inf_{t \in V(\tau)} X_n(t) < -\varepsilon\right\}$

и покажем, что правая часть (3.2.35) стремится к нулю. Для этого достаточно получить для этих вероятностей равномерную оценку сверху вида $o((n^{\gamma r} h^{-r})^{-1})$. Поскольку исследование каждой из упомянутых вероятностей проходит одинаково, рассмотрим только первую из них.

Заметим, что линейная форма (от переменных t_1, \dots, t_r) $c_i t = \sum_{\alpha=1}^r c_{i\alpha} t_\alpha$ достигает на кубе $V(\tau)$ своих максимального и минимального значений на некоторой паре диагонально расположенных вершин. Обозначим их через μ_i и λ_i . Поскольку $\text{sign } x$ — монотонно неубывающая функция, для $t \in V(\tau)$ справедливо неравенство

$$\text{sign}\left(\xi_i + \frac{c_i \lambda_i}{\sqrt{n}}\right) \leq \text{sign}\left(\xi_i + \frac{c_i t}{\sqrt{n}}\right) \leq \text{sign}\left(\xi_i + \frac{c_i \mu_i}{\sqrt{n}}\right).$$

Представим d_i в виде $d_i = d_i^+ - d_i^-$, где

$$d_i^+ = \begin{cases} d_i, & \text{если } d_i > 0, \\ 0, & \text{если } d_i \leq 0, \end{cases} \quad d_i^- = \begin{cases} 0, & \text{если } d_i \geq 0, \\ -d_i, & \text{если } d_i < 0. \end{cases}$$

Теперь для $X_n(t)$, $t \in V(\tau)$, получаем оценки сверху и снизу. Оценки

сверху:

$$\begin{aligned}
 X_n(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i \left(\text{sign} \left(\xi_i + \frac{c_i t}{\sqrt{n}} \right) - \text{sign} \xi_i - 2 f(0) \frac{c_i t}{\sqrt{n}} \right) \leq \\
 &\leq \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i^+ \left(\text{sign} \left(\xi_i + \frac{c_i \mu_i}{\sqrt{n}} \right) - \text{sign} \xi_i - 2 f(0) \frac{c_i \lambda_i}{\sqrt{n}} \right) - \\
 &- \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i^- \left(\text{sign} \left(\xi_i + \frac{c_i \lambda_i}{\sqrt{n}} \right) - \text{sign} \xi_i - 2 f(0) \frac{c_i \mu_i}{\sqrt{n}} \right) \leq \\
 &\leq \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i^+ \left(\text{sign} \left(\xi_i + \frac{c_i \mu_i}{\sqrt{n}} \right) - \text{sign} \left(\xi_i + \frac{c_i \tau}{\sqrt{n}} \right) \right) + \\
 &+ \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i^- \left(\text{sign} \left(\xi_i + \frac{c_i \tau}{\sqrt{n}} \right) - \text{sign} \left(\xi_i + \frac{c_i \lambda_i}{\sqrt{n}} \right) \right) + \\
 &+ \frac{1}{\sqrt{n}} \left| \sum_{i=1}^n d_i \left(\text{sign} \left(\xi_i + \frac{c_i \tau}{\sqrt{n}} \right) - \text{sign} \xi_i - 2 f(0) \frac{c_i \tau}{\sqrt{n}} \right) \right| + \\
 &+ \frac{2 f(0)}{n} \left| \sum_{i=1}^n c_i d_i (\tau - t) \right| \quad (3.2.36)
 \end{aligned}$$

Обозначим слагаемые в правой части через A_1, A_2, A_3 и A_4 соответственно. Ясно, что

$$\begin{aligned}
 \mathbf{P} \left\{ \sup_{t \in V(\tau)} X_n(t) > \varepsilon \right\} &\leq \\
 &\leq \mathbf{P} \left\{ A_1 > \frac{\varepsilon}{4} \right\} + \mathbf{P} \left\{ A_2 > \frac{\varepsilon}{4} \right\} + \mathbf{P} \left\{ A_3 > \frac{\varepsilon}{4} \right\} + \mathbf{P} \left\{ A_4 > \frac{\varepsilon}{4} \right\}. \quad (3.2.37)
 \end{aligned}$$

Заметим, что в силу условия i) $A_4 \leq \frac{2 f(0)}{n} K^2 h$, и потому при $h = h(n) \rightarrow 0$ это слагаемое стремится к нулю. Таким образом, при достаточно малых h последняя вероятность в (3.2.37) равна 0.

2°. Оценим сверху остальные вероятности. Начнем с A_1 и A_2 . Так как их структура одинакова, мы рассмотрим лишь одно из них. Для другого доказательство проходит аналогично.

По неравенству Чебышева для целого $m > 0$

$$\mathbf{P} \left\{ A_1 > \frac{\varepsilon}{4} \right\} < \left(\frac{\varepsilon}{4} \right)^{-m} \mathbf{E} A_1^m. \quad (3.2.38)$$

Оценим $\mathbf{E} A_1^m$. Положим

$$\omega_i = \omega_i(\tau) := \text{sign} \left(\xi_i + \frac{c_i \mu_i}{\sqrt{n}} \right) - \text{sign} \left(\xi_i + \frac{c_i \tau}{\sqrt{n}} \right)$$

и отметим, что случайные величины $\omega_1, \dots, \omega_n$ независимы. Далее,

$$EA_1^m = n^{-m/2} \sum_{i_1} \dots \sum_{i_m} d_{i_1}^+ \dots d_{i_m}^+ E\omega_{i_1} \dots \omega_{i_m}. \quad (3.2.39)$$

Для $E\omega_i$ имеем

$$\begin{aligned} E\omega_i &= P\left\{\xi_i + \frac{c_i \mu_i}{\sqrt{n}} > 0 > \xi_i + \frac{c_i \tau}{\sqrt{n}}\right\} = \\ &= P\left\{-\frac{c_i \mu_i}{\sqrt{n}} < \xi_i < -\frac{c_i \tau}{\sqrt{n}}\right\} = f(\kappa_i) \frac{c_i(\mu_i - \tau)}{\sqrt{n}}, \end{aligned}$$

где κ_i — некоторая точка, промежуточная между $-c_i \mu_i / \sqrt{n}$ и $-c_i \tau / \sqrt{n}$. Заметим, что $\kappa_i \rightarrow 0$ равномерно по i и всем кубам разбиения. Поэтому все $f(\kappa_i)$ могут быть ограничены сверху одной постоянной. Поэтому

$$|E\omega_i| \leq \text{const} \cdot \frac{h}{\sqrt{n}}. \quad (3.2.40)$$

Назовем *уровнем* произведения $\omega_{i_1} \dots \omega_{i_m}$ число различных значений среди индексов i_1, \dots, i_m .

Представим (3.2.39) в виде $EA_1^m = \Sigma_{(m)} + \dots + \Sigma_{(1)}$, где $\Sigma_{(p)}$, $1 \leq p \leq m$, содержит все слагаемые уровня p . Заметим, что $\left(\frac{\omega_i}{2}\right)^p = \frac{\omega_i}{2}$ для любого целого $p \geq 1$, и потому $E\omega_i^p = 2^{p-1} E\omega_i$. В силу (3.2.40) для $E\omega_{i_1} \dots \omega_{i_p}$ из $\Sigma_{(p)}$ имеем оценку

$$|E\omega_{i_1} \dots \omega_{i_p}| \leq \text{const} \cdot \left(\frac{h}{\sqrt{n}}\right)^p.$$

Учитывая, что количество слагаемых, составляющих $\Sigma_{(p)}$, есть $O(n^p)$, получаем, что

$$|\Sigma_{(p)}| \leq \text{const} \cdot \left(\frac{1}{\sqrt{n}}\right)^m n^p \left(\frac{h}{\sqrt{n}}\right)^p. \quad (3.2.41)$$

Теперь выберем $h = h(n)$ так, чтобы $h(n) \rightarrow 0$ и $\sqrt{n}h(n) \rightarrow \infty$ при $n \rightarrow \infty$. Например, положим $h(n) = n^{-1/4}$. Тогда $\Sigma_{(p)} \leq \text{const} \cdot \Sigma_{(m)}$, и потому в дальнейшем достаточно исследовать сумму $\Sigma_{(m)}$, которая является главным членом в EA_1^m . Из (3.2.41) следует, что $EA_1^m \leq \text{const} \cdot h^m$.

Вернемся к (3.2.35) и вспомним, что количество слагаемых вида $P\left\{\sup_{t \in V(\tau)} X_n(t) > \varepsilon\right\}$, есть величина $O(n^{\gamma r} h^{-r})$. При выборе $h = n^{-1/4}$ эта величина есть $O\left(\left(\frac{n^\gamma}{\sqrt{n}}\right)^r n^{\frac{3}{4}r}\right) = O\left(n^{\frac{3}{4}r}\right)$. Мы получим,

что общая оценка для членов типа EA_1^m будет стремиться к нулю при $n \rightarrow \infty$, если выбрать $m \geq 3r + 1$.

Вероятность $P\{A_2 > \varepsilon/4\}$ оценивается аналогично.

Обратимся к $P\{A_3 > \varepsilon/4\}$. Положим

$$w_i = w_i = \text{sign} \left(\xi_i + \frac{c_i \tau}{\sqrt{n}} \right) - \text{sign} \xi_i - 2 f(0) \frac{c_i \tau}{\sqrt{n}}.$$

Оценим Ew_i и $E|w_i|^p$, $p > 0$. Имеем

$$\begin{aligned} Ew_i &= E \text{sign} \left(\xi_i + \frac{c_i \tau}{\sqrt{n}} \right) - 2 f(0) \frac{c_i \tau}{\sqrt{n}} = \\ &= P \left\{ \xi_i + \frac{c_i \tau}{\sqrt{n}} > 0 \right\} - P \left\{ \xi_i + \frac{c_i \tau}{\sqrt{n}} < 0 \right\} - 2 f(0) \frac{c_i \tau}{\sqrt{n}} = \\ &= 1 - 2 F \left(- \frac{c_i \tau}{\sqrt{n}} \right) - 2 f(0) \frac{c_i \tau}{\sqrt{n}} = 2 \left(F(0) - F \left(- \frac{c_i \tau}{\sqrt{n}} \right) \right) - 2 f(0) \frac{c_i \tau}{\sqrt{n}}. \end{aligned}$$

В силу условия i) и условия $|\tau| < \text{const} \cdot n^\gamma$ аргумент $-c_i \tau / \sqrt{n}$ стремится к нулю, и потому при достаточно больших n попадает в окрестность нуля, упомянутую в условии 2.1.2. Поэтому

$$F(0) - F \left(- \frac{c_i \tau}{\sqrt{n}} \right) = f(\kappa_i) \frac{c_i \tau}{\sqrt{n}},$$

где κ_i — некоторая промежуточная точка между 0 и $-c_i \tau / \sqrt{n}$. При этом $|\kappa_i| \leq \left| \frac{c_i \tau}{\sqrt{n}} \right| \leq \text{const} \cdot n^{-\varepsilon}$, где $\varepsilon = \frac{1}{2} - \gamma > 0$. Продолжая выкладки, можем записать, что

$$Ew_i = 2 \left(f(\kappa_i) - f(0) \right) \frac{c_i \tau}{\sqrt{n}}.$$

В силу условия 2.1.6

$$|Ew_i| \leq 2L |\kappa_i| \left| \frac{c_i \tau}{\sqrt{n}} \right| \leq \text{const} \cdot n^{-2\varepsilon}. \quad (3.2.42)$$

Обратимся к Ew_i^2 . Имеем

$$\begin{aligned} Ew_i^2 &= \left(2 - 2f(0) \frac{c_i \tau}{\sqrt{n}} \right)^2 P \left\{ \xi_i + \frac{c_i \tau}{\sqrt{n}} > 0, \xi_i < 0 \right\} + \\ &+ \left(-2 - 2f(0) \frac{c_i \tau}{\sqrt{n}} \right)^2 P \left\{ \xi_i + \frac{c_i \tau}{\sqrt{n}} < 0, \xi_i > 0 \right\} + \end{aligned}$$

$$\begin{aligned}
 & + \left(-2f(0) \frac{c_i t}{\sqrt{n}} \right)^2 \left(1 - \mathbf{P} \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} > 0, \xi_i < 0 \right\} - \right. \\
 & \quad \left. - \mathbf{P} \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} < 0, \xi_i > 0 \right\} \right) \leq \\
 & \leq \left(2 + 2f(0) \left| \frac{c_i t}{\sqrt{n}} \right| \right)^2 \left(\mathbf{P} \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} > 0, \xi_i < 0 \right\} + \right. \\
 & \quad \left. + \mathbf{P} \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} < 0, \xi_i > 0 \right\} \right) + \left(2f(0) \right)^2 \left| \frac{c_i t}{\sqrt{n}} \right|^2. \quad (3.2.43)
 \end{aligned}$$

Заметим, что из двух событий

$$\left\{ \xi_i + \frac{c_i t}{\sqrt{n}} > 0, \xi_i < 0 \right\}, \quad \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} < 0, \xi_i > 0 \right\}$$

может осуществиться только одно, поэтому

$$\mathbf{P} \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} > 0, \xi_i < 0 \right\} + \mathbf{P} \left\{ \xi_i + \frac{c_i t}{\sqrt{n}} < 0, \xi_i > 0 \right\} = f(\kappa_i) \left| \frac{c_i t}{\sqrt{n}} \right|, \quad (3.2.44)$$

где κ_i было введено выше. Действительно, если $-c_i t / \sqrt{n} > 0$, то положительна вторая из этих вероятностей. При этом в силу условия 2.1.2

$$\mathbf{P} \left\{ 0 < \xi_i < -\frac{c_i t}{\sqrt{n}} \right\} = F \left(-\frac{c_i t}{\sqrt{n}} \right) - F(0) = -f(\kappa_i) \frac{c_i t}{\sqrt{n}}.$$

Если же $-c_i t / \sqrt{n} < 0$, то положительна первая из этих вероятностей. При этом она равна

$$\mathbf{P} \left\{ -\frac{c_i t}{\sqrt{n}} < \xi_i < 0 \right\} = F(0) - F \left(-\frac{c_i t}{\sqrt{n}} \right) = f(\kappa_i) \left| \frac{c_i t}{\sqrt{n}} \right|.$$

С учетом (3.2.44), условия i) и условия 2.1.6 при выборе $|t| \leq \text{const} \cdot n^\gamma$ из (3.2.43) следует, что

$$\mathbf{E} w_i^2 \leq \text{const} \cdot n^{-2\epsilon}.$$

Поскольку случайные величины w_i ограничены, мы получаем отсюда, что

$$\mathbf{E} |w_i|^p \leq \text{const} \cdot n^{-2\epsilon} \quad (3.2.45)$$

для любого $p \geq 1$.

Напомним, что

$$\mathbf{E} A_3^m = n^{-m/2} \sum_{i_1} \dots \sum_{i_m} d_{i_1} \dots d_{i_m} \mathbf{E} w_{i_1} \dots w_{i_m} \quad (3.2.46)$$

мы представляем в виде

$$EA_3^m = n^{-m/2} \{ \Sigma_{(m)} + \Sigma_{(m-1)} + \dots + \Sigma_{(p)} + \dots + \Sigma_{(1)} \}, \quad (3.2.47)$$

где в $\Sigma_{(p)}$ входят те слагаемые из (3.2.46), в которых ровно p различных значений индексов i_1, \dots, i_m . Слагаемые в $\Sigma_{(p)}$ допускают равномерную оценку сверху величиной $\text{const} \cdot n^{-2\epsilon p}$. Число слагаемых в $\Sigma_{(p)}$ есть $O(n^p)$. Поэтому

$$|\Sigma_{(p)}| \leq \text{const} \cdot n^{-m/n} n^p n^{-2\epsilon p}. \quad (3.2.48)$$

Ясно, что главным членом (в силу оценки (3.2.48)) в (3.2.47) является $\Sigma_{(m)}$, и потому

$$EA_3^m \leq \text{const} \cdot n^{m(1-4\epsilon)/2}. \quad (3.2.49)$$

При этом $1-4\epsilon < 0$, ибо $1-4\epsilon = 4\gamma - 1$ и $\gamma < 1/4$ по условию теоремы.

Ранее, при обсуждении (3.2.35), мы отмечали, что нам нужны такие оценки, которые стремятся к нулю и после умножения на $n^{\gamma r} h^{-r}$. При сделанном выборе $h = n^{-1/4}$ это $n^{\gamma r + r/4}$. Очевидно, что за счет выбора достаточно большого m можно добиться того, чтобы $n^{\gamma r + r/4} \rightarrow 0$. А именно, надо взять m таким, чтобы было выполнено неравенство

$$m \frac{4\gamma - 1}{2} + r \left(\gamma + \frac{1}{4} \right) < 0$$

или $m > r \frac{4\gamma + 1}{2(1-4\gamma)}$. Доказательство теоремы завершено. \square

3.2.7. Асимптотическая мощность знаковых критериев.

Вернемся к знаковым критериям гл. 2 для проверки гипотез (2.2.1) или (2.2.2) в модели (2.1.5). Теорема о равномерной линейности позволяет судить об их мощности при большом числе наблюдений, если принять предположения о матрице плана и распределении ошибок, сделанные в формулировке теоремы 3.2.4. Рассмотрим альтернативы, близкие к гипотезе $H_0: \theta = 0$, а именно: пусть

$$\theta = \tau / \sqrt{n}, \quad (3.2.50)$$

где $\tau = (\tau_1, \dots, \tau_r)^T$, $|\tau| < \text{const}$. Рассмотрим для таких θ знаковые критерии (2.2.18) и (2.2.21). Статистика критерия (2.2.18) в форме (2.2.19) представляет собой скалярный квадрат вектора

$$\frac{1}{\sqrt{n}} C^T S(X) = \frac{1}{\sqrt{n}} C^T S \left(\xi + \frac{1}{\sqrt{n}} C \tau \right). \quad (3.2.51)$$

По теореме 3.2.4 о равномерной линейности

$$\frac{1}{\sqrt{n}} C^T S \left(\xi + \frac{1}{\sqrt{n}} C \tau \right) = \frac{1}{\sqrt{n}} C^T \zeta + \frac{1}{n} (C^T C) \tau + o_p(1), \quad (3.2.52)$$

где $\zeta = (\zeta_1, \dots, \zeta_n)^T$ и $\zeta_i = \text{sign } \xi_i$.

Следовательно, при близких альтернативах статистика критерия (2.2.18) асимптотически распределена как квадрат нормального вектора $N(\Sigma\tau, \Sigma)$. Статистика критерия (2.2.21) с учетом (3.2.51) и (3.2.52) асимптотически эквивалентна

$$n \left(\frac{1}{\sqrt{n}} C^T \zeta + \frac{1}{n} (C^T C) \tau \right)^T (C^T C)^{-1} \left(\frac{1}{\sqrt{n}} C^T \zeta + \frac{1}{n} (C^T C) \tau \right).$$

Аналогично (2.3.8) асимптотическое распределение этой случайной величины есть хи-квадрат с r степенями свободы и параметром нецентральности

$$\frac{1}{n} \tau^T (C^T C) \tau = \left| \frac{1}{\sqrt{n}} C \tau \right|^2$$

или $\tau^T \Sigma \tau$.

3.2.8. Функция чувствительности. В гл. 1 мы уже обсуждали роль функций влияния и функций чувствительности для описания устойчивости статистических процедур по отношению к выбросам (грубым ошибкам). Используя асимптотические результаты этой главы, мы вычисляем функцию чувствительности (по Тьюки) для знаковых оценок.

Предположим, что в схеме (2.1.5) к наблюдениям x_i , где $x_i = c_i \theta^0 + \xi_i$, $i = 1, \dots, n$, присоединена еще одна точка с координатами (b, y) , где $b = (b_1, \dots, b_n)^T$, $y = b^T \theta^0 + \eta$ и b и η фиксированы. Пусть $\hat{\theta}_n$ обозначает знаковую оценку (3.1.1), а $\hat{\theta}_{n+1}^*$ — знаковую оценку, полученную тем же способом и по тем же наблюдениям, но с присоединением к ним точки (b, y) . Теперь чувствительность знаковой оценки (3.1.1) как функции от (b, η) равна

$$SC_n(b, \eta) = n(\hat{\theta}_{n+1}^* - \hat{\theta}_n). \quad (3.2.53)$$

Предположим, что выполнены условия теоремы 3.2.2. Покажем, что в этом случае при $n \rightarrow \infty$

$$SC_n(b, \eta) = \frac{\text{sign } \eta}{2f(0)} \left(\frac{1}{n} C^T C \right)^{-1} b + o_p(1). \quad (3.2.54)$$

Сделаем замену переменных $\tau = \sqrt{n}(\theta - \theta^0)$. Пусть

$$\hat{\tau}_n = \sqrt{n}(\hat{\theta}_n - \theta^0), \quad \hat{\tau}_{n+1}^* = \sqrt{n}(\hat{\theta}_{n+1}^* - \theta^0).$$

Рассуждая так же, как при доказательстве теоремы 3.2.2, можно показать, что $\hat{\tau}_{n+1}^*$ по вероятности асимптотически эквивалентна решению системы

$$\sum_{i=1}^n c_{i\alpha} \left(\text{sign } \xi_i - \frac{2f(0)}{\sqrt{n}} c_i \tau \right) + b_\alpha \text{sign}(\eta - \frac{1}{\sqrt{n}} b \tau) = 0, \quad (3.2.55)$$

$\alpha = 1, \dots, r$. Отсюда с учетом (3.2.29)–(3.2.31) получаем, что

$$\hat{\tau}_{n+1}^* = \hat{\tau}_n + \frac{1}{\sqrt{n}} \frac{1}{2f(0)} \left(\frac{1}{n} C^T C \right)^{-1} b \operatorname{sign} \eta + o_p(1). \quad (3.2.56)$$

При переходе от (3.2.55) к (3.2.56) мы учли, что при $\eta \neq 0$ $\operatorname{sign}(\eta - \frac{1}{\sqrt{n}} b^T \hat{\tau}_{n+1}^*) \xrightarrow{P} \operatorname{sign} \eta$, так как случайный вектор $\hat{\tau}_{n+1}^*$ ограничен по вероятности. Подставляя (3.2.56) в (3.2.53), получим требуемое (3.2.54). \square

§ 3.3. Сравнение оценок

3.3.1. Как сравнивают оценки. Многие методы могут применяться для оценивания параметров линейных моделей. Мы обсудим здесь лишь некоторые наиболее известные из них. Мы ограничимся теми методами, которые (подобно знаковым) можно применять и тогда, когда распределение случайных ошибок неизвестно. (В отличие от метода наибольшего правдоподобия, например.) Свойства этих методов нас будут интересовать преимущественно в сравнении со свойствами знаковых оценок.

Говоря о статистическом оценивании параметров, имеют в виду две задачи. Первая — вычисление (на основании имеющихся наблюдений) приближенных значений для этих параметров. Это задача точечного оценивания. Вторая — построение таких областей (множеств) в пространстве параметров, которые с гарантированной вероятностью содержат в себе («накрывают») истинное значение параметра. Вероятность такого накрытия называют *доверительной вероятностью*, само случайное множество — *доверительным множеством*, а весь подход в целом — *доверительным оцениванием*. Когда доверительная вероятность близка к единице, истинное значение практически достоверно принадлежит доверительному множеству.

Говоря вообще, точечные оценки и доверительные множества для неизвестных параметров получают разными путями. Поэтому напрямую они не связаны друг с другом. Если же точечное и доверительное оценивание осуществляются в рамках общего подхода (например, такого, как знаковый), то доверительные множества формируются вокруг точечных оценок. Размеры и формы доверительных множеств в этом случае позволяют судить о том, какая точность достигается при точечном оценивании: чем уже доверительные множества, тем выше точность оценивания. Эта точность, естественно, зависит в первую очередь от свойств статистических ошибок, но также и от выбранного метода оценивания. Возможность дополнить точечную оценку доверительными множествами очень важна: оценка как приближенное значение мало полезна на практике, если остается неясной достигаемая при этом точность. Если же мы можем последовательно построить доверительные множества с несколькими различными доверительными вероятностями, то их размеры и формы дадут нам

наглядное представление об этой точности. Подчеркнем, что мы говорим о доверительных множествах, доверительных вероятности для которых известны точно (при выполнении модельных предположений). Асимптотические доверительные множества (когда объем наблюдений стремится к бесконечности) построить для параметров линейной модели относительно несложно. Для этого достаточно иметь какую-либо асимптотически нормальную оценку. Широкий класс таких оценок составляют, например, M -оценки Хубера. (см., например, [36, 73; 42, 81], и M -оценки включают в себя, в частности, оценки наименьших квадратов и оценки наименьших модулей. Наши знаковые оценки из § 3.1 тоже являются M -оценками.

Точное доверительное оценивание при конечных объемах данных является более трудным делом. Для линейных моделей точные доверительные множества хорошо известны, когда ошибки распределены по нормальному закону. В непараметрической постановке мы показали в § 3.1, как это можно делать с помощью знаковых тестов. По своим математическим основаниям и возможностям ближе других к знаковым методам стоят методы ранговые и знако-ранговые (см., например, [38, 76; 41, 79; 103]). Ранговые и знако-ранговые методы тоже позволяют получать для неизвестных параметров доверительные множества, хотя с вычислительной точки зрения ранговые методы сложнее, чем знаковые. Для ранговых методов теоретически корректные результаты получаются только если случайные ошибки непрерывно и одинаково распределены. Для знаковых методов это требование не необходимо. Для знако-ранговых методов к сказанному о ранговых методах надо добавить еще одно: предположение о симметрии в распределении ошибок. Понятно, что чем больше предположений, тем больше и шансов на то, что некоторые из них на деле окажутся невыполненными. В таком случае все выводы, сделанные в этих предположениях, оказываются сомнительными. Поэтому чем меньше предположений о свойствах ошибок мы принимаем, тем надежнее наши результаты.

Области применимости упомянутых методов доверительного оценивания не совпадают. Уже поэтому нельзя сказать, что один из них превосходит другие. Но даже тогда, когда применимы какие-то два способа доверительного оценивания, их результаты оказываются столь различными, что сравнить их не удастся. Положение полностью аналогично сравнению статистических тестов: по своей мощности один из них редко превосходит другой равномерно.

Точечное оценивание как вычисление приближенных значений для неизвестных параметров и их функций представляет собой объект давних усилий. Как статистический метод оно восходит ко временам Лапласа и Гаусса (см., например, [107]). Поэтому не приходится удивляться, что разнообразие методов точечного оценивания велико. Далее мы будем сравнивать точечные знаковые оценки с некоторыми наиболее известными: оценками наименьших квадратов и наименьших модулей и ранговыми оценками.

Чтобы понять, какой из методов оценивания дает лучшие результаты, надо сопоставить получаемые оценки как случайные величины. Более точной будет та из оценок, распределение которой окажется более сконцентрированным около истинного значения. К сожалению, с

этих позиций распределения оценок обычно оказываются несопоставимыми. Поэтому не удивительно, что при конечном числе наблюдений наилучшего способа оценивания не существует.

В асимптотической постановке сравнение методов оценивания все же оказывается возможным. Познакомиться с асимптотической теорией оптимальности можно по многим источникам. Мы рекомендуем книгу [19], отражающую современные взгляды на методы асимптотического сравнения точечных оценок и [19, гл. 6] или [16, гл. 2]. Впрочем, наши сравнения знаковых и других оценок не потребуют сложной теории, поскольку мы ограничимся лишь вычислениями асимптотической эффективности знаковых оценок относительно некоторых других.

Для всех разумных методов оценивания типично, что оценки параметров асимптотически нормальны при увеличении числа наблюдений. Точнее, при увеличении объема информации, так как помимо числа наблюдений многое зависит от плана эксперимента. Чтобы упростить свою задачу, мы будем предполагать, что планы экспериментов удовлетворяют условиям 2.1.7, 2.1.8. При таких предположениях число наблюдений n становится единственной ведущей переменной. Мы рассматриваем сейчас поведение оценок параметров линейной модели при $n \rightarrow \infty$.

Ниже мы дадим некоторые сведения о ранговых оценках (п. 3.3.2) и об оценках наименьших модулей и наименьших квадратов (п. 3.3.3). Знако-ранговые оценки по свойствам близки к ранговым. По этой причине и потому, что для их корректного применения необходима симметрия в распределении ошибок (что весьма ограничительно), мы о них говорить не будем.

3.3.2. Ранговое оценивание. Напомним, что *рангами* элементов произвольной конечной совокупности называют те номера, которые получают элементы этой совокупности после ее упорядочения по какому-либо определенному правилу. Числовую совокупность можно, к примеру, упорядочить по возрастанию, от меньших чисел к большим. При таком упорядочении ранг 1 получает наименьшее из чисел, ранг 2 — наименьшее из оставшихся (после удаления элемента с рангом 1) и т.д. Именно такие ранги мы будем иметь в виду в последующем изложении, хотя бывают полезны и другие упорядочения (т.е. другие ранги). Математическая теория предполагает, что среди чисел нет совпадающих. На практике совпадения обыкновенны. Обычно всем совпадающим (одинаковым) числам приписывают общий средний ранг (см., например, [94] или [41, 79]).

Ранги совокупности u_1, u_2, \dots, u_n мы будем обозначать как $R(u_1), R(u_2), \dots, R(u_n)$. Заметим, что ранги не изменяются при переходе от совокупности u_1, u_2, \dots, u_n к $u_1 + c, u_2 + c, \dots, u_n + c$, где c — произвольное число, или к ku_1, ku_2, \dots, ku_n , где $k > 0$. Эти свойства рангов обуславливают возможности ранговых методов в анализе линейных моделей, о чем мы еще будем говорить определеннее позже.

Ранговое оценивание и другие ранговые правила в случае линейных моделей основываются на рангах видимых остатков. Прежде чем

определить их формально, полезно выбрать форму для линейной модели (2.1.1) или (2.1.5), наилучше приспособленную для ранговых методов. Мы примем линейную модель в виде

$$x_i = \theta_0 + \sum_{\alpha=1}^r c_{i\alpha} \theta_\alpha + \varepsilon_i, \quad i = 1, \dots, n, \quad (3.3.1)$$

считая, что $\sum_{i=1}^n c_{i\alpha} = 0$ для каждого $\alpha = 1, \dots, r$. При асимптотическом анализе (когда $n \rightarrow \infty$) для матрицы $\|c_{i\alpha}\|$ из (3.3.1) мы сохраняем условия 2.1.7, 2.1.8. Упомянутые видимые остатки в модели (3.3.1)

$$x_i - \theta_0 - \sum_{\alpha=1}^r c_{i\alpha} \theta_\alpha, \quad i = 1, \dots, n,$$

считаются функциями переменных $\theta_0, \theta_1, \dots, \theta_r$. Ясно, что их ранги не зависят от θ_0 и равны $R\left(x_i - \sum_{\alpha=1}^r c_{i\alpha} \theta_\alpha\right)$, $i = 1, \dots, n$. Поэтому параметр θ_0 (параметр пересечения, как иногда говорят) не может быть оценен с помощью рангов. Так отзывается потеря информации, которая происходит от замены наблюдений их рангами.

Анализ модели (3.3.1) производится с помощью линейных ранговых статистик

$$\sum_{i=1}^n c_{i\alpha} a_n \left(R\left(x_i - \theta_0 - \sum_{\alpha=1}^r c_{i\alpha} \theta_\alpha\right) \right), \quad \alpha = 1, \dots, r,$$

где $a_n(1), a_n(2), \dots, a_n(n)$ — достаточно произвольная последовательность чисел. Числа $a_n(1), \dots, a_n(n)$ называют метками. Обычно в качестве меток $a_n(k)$, $k = 1, \dots, n$, рассматривают либо

$$a_n(k) = E\varphi(v_n^{(k)}), \quad k = 1, \dots, n,$$

либо

$$a_n(k) = \varphi\left(\frac{k}{n+1}\right), \quad k = 1, \dots, n,$$

где $\varphi(\cdot)$ — неубывающая функция, заданная на отрезке $[0, 1]$. Через $v_n^{(1)}, v_n^{(2)}, \dots, v_n^{(n)}$ обозначены порядковые статистики, полученные по выборке объема n из равномерного распределения на отрезке $[0, 1]$.

Метки $a_n(k)$ следовало бы выбирать, сообразуясь с распределением случайных ошибок: для каждого распределения есть свои, наиболее подходящие метки. На практике это неосуществимо, поскольку распределение ошибок неизвестно. (А будь оно известно, не пришлось бы прибегать к непараметрическим методам.) Часто используют так называемые нормальные метки (они оптимальны, когда ошибки распределены нормально), еще чаще уилкоксоновские (когда

$a_n(k) = k/(n+1)$). Вычисления с ними особенно просты. Для уилкоксонских меток составлены таблицы распределений многих ранговых статистик (см., например, [94] или [41, 79]).

Ранговые оценки $\theta_1^*, \dots, \theta_r^*$ параметров $\theta_1, \dots, \theta_r$ определяют как решение, например, такой экстремальной задачи

$$\sum_{\alpha=1}^r \left(\sum_{i=1}^n c_{i\alpha} a_n(R(x_i - c_i\theta)) \right)^2 \Rightarrow \min_{\theta_1, \dots, \theta_r}. \quad (3.3.2)$$

Отметим попутно, что статистика из (3.3.2) может служить и для построения доверительных множеств для истинных значений $\theta^0 = (\theta_1^0, \dots, \theta_r^0)^T$ — так же, как это мы делали выше для знаковых статистик.

Мы уже говорили, что ранговые методы применимы лишь в случае одинаково распределенных ошибок. Относительно этого общего распределения, функцию которого обозначим через $F(\cdot)$, и относительно меток $a_n(i)$ далее предполагается, что:

1) существует абсолютно непрерывная плотность $f(x) = F'(x)$ и ее производная $f'(x)$;

2) конечна информация (по Фишеру)

$$\int_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)} \right)^2 f(x) dx;$$

3) метки $a_n(i)$ имеют вид либо

$$a_n(i) = \mathbf{E}\varphi(U_n^{(i)}),$$

либо

$$a_n(i) = \varphi\left(\frac{i}{n+1}\right),$$

где $U_n^{(1)} < U_n^{(2)} < \dots < U_n^{(n)}$ обозначает вариационный ряд, построенный по выборке объема n из равномерного на $[0, 1]$ распределения;

4) функция меток $\varphi(\cdot)$ невозрастающая, интегрируемая с квадратом на $[0, 1]$, не являющаяся постоянной.

Теперь мы можем сформулировать результаты об асимптотической нормальности ранговых оценок в модели (3.3.1) при $n \rightarrow \infty$, как они даны в [75].

При сделанных выше предположениях о плане эксперимента, о метках и о распределении случайных ошибок вектор θ^ ранговых оценок:*

1) *распределен асимптотически по закону*

$$N\left(\theta^0, \frac{A^2}{\gamma^2} (C^T C)^{-1}\right),$$

где θ^0 — истинное значение параметра,

$$A^2 = \int_0^1 (\varphi(u) - \bar{\varphi})^2 du, \quad \bar{\varphi} = \int_0^1 \varphi(u) du,$$

$$\gamma = \int_{-\infty}^{\infty} \varphi(F(x))f'(x) dx;$$

2) все решения задачи (3.3.2) асимптотически эквивалентны в том смысле, что для любых решений θ' и θ''

$$\sqrt{n}(\theta' - \theta'') \xrightarrow{P} 0 \quad \text{при } n \rightarrow \infty.$$

3.3.3. Оценки наименьших квадратов и наименьших модулей. По-видимому, это наиболее известные и издавна применяемые методы оценивания, особенно метод наименьших квадратов. Этот метод нередко приписывают Гауссу, который его действительно успешно применял. Сам Гаусс говорил о нем как об уже известном.

Оценки наименьших квадратов в линейной модели определяются как решение экстремальной задачи

$$\sum_{i=1}^n (x_i - c; \theta)^2 \Rightarrow \min_{\theta}. \quad (3.3.3)$$

Для решения, которое обозначим $\hat{\theta}$, есть явная формула:

$$\hat{\theta} = (C^T C)^{-1} C^T X. \quad (3.3.4)$$

К предположениям о независимости и одинаковой распределенности ошибок обычно добавляют условия

$$E\xi_i = 0, \quad \sigma^2 := E\xi_i^2 < \infty.$$

Тогда при условии 2.1.8 о матрице плана оценки (3.3.4) состоятельны и асимптотически нормальны $N(\theta, \frac{\sigma^2}{n} \Sigma^{-1})$. Этот факт столь давно и хорошо известен, что трудно указать его автора. О свойствах оценок наименьших квадратов при более общих условиях на матрицу плана см., например, [63] и [2, 45, гл. 2].

Метод наименьших квадратов самым тесным образом связан с гауссовскими линейными моделями, т.е. с линейными моделями с нормально распределенными ошибками. В этом случае оценки наименьших квадратов служат и оценками наибольшего правдоподобия. Для гауссовских моделей метод наименьших квадратов в определенном смысле наилучший. Он дает наилучшие оценки для параметров,

позволяет получать для них доверительные множества, проверять линейные гипотезы и т.д. Нет необходимости говорить здесь об этом методе подробнее, так как его изложение в той или иной форме входит в учебники по математической статистике и отражено в нескольких монографиях.

Метод наименьших модулей известен столь же давно, как и метод наименьших квадратов. В частности, Лаплас был его сторонником. Оценки наименьших модулей определяются как решение экстремальной задачи

$$\sum_{i=1}^n |x_i - c_i \theta| \Rightarrow \min_{\theta}. \quad (3.3.5)$$

Как будет видно, оценки наименьших модулей близки к знаковым оценкам, если говорить о линейной модели. Оценки наименьших модулей оказываются оценками наибольшего правдоподобия, если ошибки подчиняются двустороннему показательному закону. (Иногда такое распределение называют распределением Лапласа.) Для этого масштабно-сдвигового семейства распределений стандартная плотность вероятности есть $p(x) = 0.5 e^{-|x|}$.

Явной формулы, выражающей оценку наименьших модулей (наподобие (3.3.4)), не существует. Вообще, в вычислительном отношении задача (3.3.5) труднее, чем (3.3.3). Минимум в (3.3.5) находят либо методами линейного программирования, либо итерационно, опираясь на выпуклость целевой функции (см., например, [51] или [23]).

С асимптотической точки зрения знаковые оценки и оценки наименьших модулей эквивалентны. Несмотря на старинное происхождение метода, асимптотические свойства оценок наименьших модулей были изучены относительно недавно. Приведем теорему Г.Бассета и Р.Кенкера [47]; см. также [102].

Т е о р е м а (об асимптотической нормальности оценок наименьших модулей). *Предположим, что в модели (2.1.5) случайные ошибки независимы и одинаково распределены, что их функция распределения $F(x)$ имеет равную нулю медиану, функция плотности $f(\cdot)$ непрерывна и положительна в нуле. Относительно матрицы C примем условие 2.1.8.*

Тогда при $n \rightarrow \infty$ оценка наименьших модулей $\hat{\theta}_n$ (3.3.5) асимптотически нормальна:

$$N\left(0, \frac{1}{(2f(0))^2} (C^T C)^{-1}\right),$$

или

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\left(0, \frac{1}{(2f(0))^2} \Sigma^{-1}\right).$$

Мы видим, что, исключая условие 2.1.7, имеющее технический характер, условия и утверждения этой теоремы совпадают с теоремой 3.2.2 (хотя последняя говорит об асимптотическом поведении знаковых оценок несколько больше).

3.3.4. Асимптотическая эффективность знаковых оценок.

Взглянув на асимптотические распределения перечисленных оценок, мы видим, что все они (при определенных условиях) гауссовские, что их асимптотические векторы средних одинаковы и равны вектору искомым параметров и что их асимптотические матрицы ковариаций пропорциональны. Для знаковых оценок асимптотическую матрицу ковариаций можно рассматривать как J^{-1} , где J — фишеровская матрица информации о векторе θ , которая содержится в случайных величинах $\text{sign}(x_i - c_i\theta)$, $i = 1, \dots, n$. Это утверждение легко доказать с помощью формулы (2.2.15). При этом матрица $C^T C$ отражает влияние плана эксперимента, а множитель $(2f(0))^2$ — влияние метода оценивания.

Такую же структуру имеют асимптотические матрицы ковариаций и многих других оценок. Они пропорциональны матрице $(C^T C)^{-1}$, или $\frac{1}{n} \Sigma^{-1}$ в условиях 2.1.8. Ее сопровождают скалярные множители, разные для разных способов оценивания. Такая пропорциональность асимптотических матриц ковариаций позволяет (для больших выборок) сравнивать способы оценивания по их точности — в тех случаях, когда они применимы одновременно. Действительно, из двух нормальных законов с общим вектором средних тот более сконцентрирован около общего центра, у которого меньше дисперсия (матрица ковариаций). Если эти матрицы пропорциональны, их можно сравнить и выяснить, которая меньше. Отношение упомянутых скалярных множителей для каких-либо двух методов оценивания называют *асимптотической эффективностью* одного метода относительно другого (см., например, [19, § 6.2] или [3, 48, § 4.4.B]).

Сравнение методов оценивания по их асимптотической эффективности традиционно обсуждается в литературе. Результаты сравнения зависят от законов распределения ошибок. При этом сравнивать методы можно лишь для таких распределений, когда для обоих методов верны соответствующие асимптотические результаты. Так как упомянутые выше методы оценивания различаются в первую очередь своими областями корректного применения, их сопоставление по асимптотической эффективности имеет ограниченное значение.

Так как знаковые оценки и оценки наименьших модулей асимптотически эквивалентны, то все, что известно об асимптотической эффективности оценок наименьших модулей, оказывается верным и для знаковых оценок. Можно сопоставить, например, знаковый метод и метод наименьших квадратов, когда ошибки распределены по нормальному закону $N(0, \sigma^2)$. Мы получим, что асимптотическая эффективность знакового метода (и метода наименьших модулей) по отношению к методу наименьших квадратов в этих условиях составляет $2/\pi$. (Это значит, что оба метода дадут одинаковые по точности

заклучения, если дисперсия ошибки при применении метода наименьших квадратов будет в $\pi/2$ раз больше, чем дисперсия ошибки при знаковом методе.)

Пусть теперь распределение ошибок двустороннее показательное с дисперсией σ^2 . Плотность распределения ошибок в таком случае равна

$$\frac{1}{\sigma\sqrt{2}} \exp\left\{-\frac{\sqrt{2}|x|}{\sigma}\right\}, \quad x \in \mathbb{R}^1.$$

Множитель при $(C^T C)^{-1}$ в асимптотической матрице ковариаций знаковых оценок при этом равен $\left(\frac{1}{2f(0)}\right)^2 = \frac{\sigma^2}{2}$. Сопоставляя эту матрицу с матрицей ковариаций для оценки наименьших квадратов из п. 3.3.3, мы получаем, что для такого распределения ошибок знаковый метод асимптотически вдвое более эффективен, чем метод наименьших квадратов.

Легко видеть, что асимптотическая эффективность знакового метода по отношению к методу наименьших квадратов может быть вообще сколь угодно велика. Сравним их, например, для ошибок, распределенных по Тьюки. Это удобная модель для нормальных данных, засоренных некоторой долей выбросов, т.е. грубых ошибок. Предположим, что стандартное нормальное распределение и нормальное распределение со средним 0 и дисперсией τ^2 смешаны в пропорции $1 - \epsilon$ к ϵ , $\epsilon > 0$. Плотность ошибок при этом равна

$$(1 - \epsilon) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} + \epsilon \frac{1}{\sqrt{2\pi\tau}} e^{-x^2/(2\tau^2)},$$

дисперсия ошибок $\sigma^2 = (1 - \epsilon) + \epsilon\tau^2$, $(2f(0))^2 = \frac{2}{\pi} \left(1 - \epsilon + \frac{\epsilon}{\tau}\right)$. Здесь асимптотическая эффективность знакового метода относительно метода наименьших квадратов составляет

$$\left(\frac{2f(0)}{\sigma}\right)^2 = \frac{2}{\pi} \frac{1 - \epsilon + \epsilon/\tau}{1 - \epsilon + \epsilon\tau^2}.$$

С ростом τ она неограниченно возрастает.

ПРОВЕРКА ЛИНЕЙНЫХ ГИПОТЕЗ

§ 4.1. Знаковые критерии для линейных гипотез

Линейной гипотезой в модели (2.1.1) называют гипотезу

$$H: l \in L_1, \quad (4.1.1)$$

где L_1 — заданное линейное подпространство, причем $L_1 \subset L$. В § 4.3 мы рассмотрим несколько наиболее часто встречающихся линейных гипотез, а сейчас опишем общий метод для их проверки. Фактически этот метод повторяет принятую в гауссовской теории схему проверки линейных моделей.

Рассмотрим дополнение L_1 до всего подпространства L и обозначим его через L_2 . Ради некоторых полезных свойств функции мощности того критерия, который сейчас создается, выберем L_2 ортогональным к L_1 . Теперь $L = L_1 \oplus L_2$, $L_1 \perp L_2$. Пусть $r_1 = \dim L_1$, $r_2 = \dim L_2$, $r = \dim L$. Пусть векторы $(a_{1\alpha}, a_{2\alpha}, \dots, a_{n\alpha})^T$, $\alpha = 1, \dots, r_1$, образуют базис L_1 ; векторы $(b_{1\beta}, b_{2\beta}, \dots, b_{n\beta})^T$, $\beta = 1, \dots, r_2$, образуют базис L_2 . Составим из перечисленных векторов как из столбцов матрицы A и B размеров $n \times r_1$ и $n \times r_2$ соответственно, положив

$$A = \|a_{i\alpha}\|, \quad i = 1, \dots, n, \quad \alpha = 1, \dots, r_1,$$

$$B = \|b_{i\beta}\|, \quad i = 1, \dots, n, \quad \beta = 1, \dots, r_2.$$

В силу ортогональности L_1 и L_2 справедливо соотношение

$$A^T B = 0. \quad (4.1.2)$$

С помощью матриц A и B подпространства L_1 и L_2 можно представить в виде

$$L_1 = \{l: l = A\sigma, \sigma \in \mathbb{R}^{r_1}\}, \quad L_2 = \{l: l = B\tau, \tau \in \mathbb{R}^{r_2}\}.$$

В этих обозначениях модель (2.1.1) принимает вид

$$X = A\sigma + B\tau + \xi, \quad (4.1.3)$$

Условие 4.2.2.

$$\frac{1}{n}A^T A \rightarrow \Sigma_1 > 0, \quad \frac{1}{n}B^T B \rightarrow \Sigma_2 > 0 \quad \text{при } n \rightarrow \infty.$$

Относительно случайных ошибок примем предположения теоремы 3.2.4 (о равномерной линейности), т.е. условия 2.1.1, 2.1.2, 2.1.6. При перечисленных условиях с учетом (4.1.2) из теоремы 3.2.4 для модели (4.1.3) следует (мы запишем вывод в векторной форме), что

$$\begin{aligned} \frac{1}{\sqrt{n}}A^T S(\xi + \frac{1}{\sqrt{n}}A\sigma + \frac{1}{\sqrt{n}}B\tau) &= \\ &= \frac{1}{\sqrt{n}}A^T S(\xi) + 2f(0)\frac{1}{n}(A^T A)\sigma + \omega_n, \quad (4.2.2) \\ \sup\{|\omega_n|: |\sigma| < Cn^\gamma, |\tau| < Cn^\gamma\} &= o_p(1). \end{aligned}$$

Опираясь на (4.2.2), мы докажем следующую теорему о свойствах знаковых оценок при близких альтернативах.

Теорема 4.2.1. *Предположим, что в модели (4.1.3) случайные ошибки независимы, одинаково распределены и их общая функция распределения удовлетворяет условиям 2.1.1, 2.1.2, 2.1.6. Относительно матриц A и B предположим, что они удовлетворяют (4.1.2) и условиям 4.2.1, 4.2.2.*

Тогда знаковая оценка (4.1.5) при близких альтернативах (4.2.1) представима в виде

$$\hat{\sigma}_n = \sigma + \frac{1}{\sqrt{n}} \frac{1}{2f(0)} \left(\frac{1}{n} A^T A \right)^{-1} \left(\frac{1}{\sqrt{n}} A^T \zeta \right) + o_p(1),$$

где $\zeta = (\zeta_1, \dots, \zeta_n)^T$, $\zeta_i = \text{sign } \xi_i$ для $i = 1, \dots, n$. Эта оценка асимптотически нормальна с параметрами $\left(0, \frac{1}{n} \frac{1}{(2f(0))^2} \left(\frac{1}{n} A^T A \right)^{-1} \right)$, или $\left(0, \frac{1}{n} \frac{1}{(2f(0))^2} \Sigma_1^{-1} \right)$.

Доказательство. Оно в основных чертах повторяет доказательство теоремы 3.2.2.

В задаче (4.1.5) сначала рассмотрим область изменения $\{\sigma: |\sigma| \geq n^{-1/2+\gamma}, \sigma \in \mathbb{R}^{r_1}\}$. Как и при доказательстве теоремы 3.2.2, сначала покажем, что для близких альтернатив

$$\inf_{\{\sigma: |\sigma| \geq n^{-1/2+\gamma}\}} \frac{1}{n} \sum_{\alpha=1}^{r_1} \left(\sum_{i=1}^n a_{i\alpha} \text{sign}(x_i - a_i \sigma) \right)^2 \xrightarrow{P} \infty. \quad (4.2.3)$$

Представим $\sigma \in \mathbb{R}^{r_1}$ в виде $\sigma = |\sigma|e$, где $e \in \mathbb{S}^{r_1-1} = \{e \in \mathbb{R}^{r_1}: |e| =$

а линейная гипотеза (4.1.1) выражается как

$$H: \tau = 0. \quad (4.1.4)$$

Применим следующее естественное правило для проверки (4.1.4). Оно будет состоять из трех шагов.

Шаг 1. Считая гипотезу $H: \tau = 0$ верной, найдем знаковую оценку параметра σ :

$$\hat{\sigma}_n = \arg \min_{\sigma \in \mathbb{R}^{r_1}} \sum_{\alpha=1}^{r_1} \left(\sum_{i=1}^n a_{i\alpha} \operatorname{sign}(x_i - a_i \sigma) \right)^2, \quad (4.1.5)$$

где $a_i = (a_{i1}, \dots, a_{ir_1})$.

Шаг 2. От наблюдений x_i перейдем к

$$\hat{x}_i = x_i - \sum_{\alpha=1}^{r_1} a_{i\alpha} \hat{\sigma}_{n\alpha}, \quad i = 1, \dots, n. \quad (4.1.6)$$

Вместо $\hat{\sigma}_n$ в (4.1.6) можно использовать любую знаковую оценку σ , что не отразится на асимптотических свойствах правила в целом (см. § 4.2).

Шаг 3. К вектору $\widehat{X} = (\hat{x}_1, \dots, \hat{x}_n)^T$ применим знаковый критерий (2.2.18) для проверки (4.1.4), приняв модель

$$X = B\tau + \xi. \quad (4.1.7)$$

Иначе говоря, гипотеза (4.1.1) отвергается на уровне значимости (приблизительно) ε , если

$$\sum_{\beta=1}^{r_2} \left(\sum_{i=1}^n b_{i\beta} \operatorname{sign} \hat{x}_i \right)^2 > q_{1-\varepsilon}, \quad (4.1.8)$$

где $q_{1-\varepsilon}$ есть $(1 - \varepsilon)$ -квантиль случайной величины

$$\sum_{\beta=1}^{r_2} \left(\sum_{i=1}^n b_{i\beta} \zeta_i \right)^2. \quad (4.1.9)$$

Распределение случайных величин ζ_i , $i = 1, \dots, n$, определено в (2.3.1).

Вместо статистики (2.2.19), которая была использована в (4.1.8), можно употребить любую другую квадратичную знаковую статистику, например (2.2.22). В таком случае шаг 3 заменяется на

Шаг 3'. Отвергнуть (4.1.1) на приближенном уровне значимости ε , если

$$S^T(\widehat{X})B(B^T B)^{-1}B^T S(\widehat{X}) > \tilde{q}_{1-\varepsilon}, \quad (4.1.10)$$

где $\tilde{q}_{1-\varepsilon}$ есть $(1 - \varepsilon)$ -квантиль случайной величины

$$\zeta^T B (B^T B)^{-1} B^T \zeta. \quad (4.1.11)$$

С геометрической точки зрения статистика в (4.1.10) есть

$$|\text{proj}_{\mathbb{L}_2} S(\widehat{X})|^2. \quad (4.1.12)$$

Эта форма представления статистики (4.1.10) позволяет увидеть сходство знакового правила (4.1.5), (4.1.6) и (4.1.10) с традиционными правилами проверки линейных гипотез в гауссовских линейных моделях, которые основываются на F -отношениях. Статистика (4.1.12) совпадает с точностью до множителя с числителем соответствующей F -статистики (при замене видимых остатков их знаками). Мы уже отмечали это сходство в § 2.3.

Использование (4.1.10) вместо (4.1.8) может упростить процедуру проверки, ибо (4.1.12) можно вычислить, не вводя явно подпространство \mathbb{L}_2 и его базис (т.е. не вводя матрицу B), так как

$$|\text{proj}_{\mathbb{L}_2} S(\widehat{X})|^2 = |\text{proj}_{\mathbb{L}} S(\widehat{X})|^2 - |\text{proj}_{\mathbb{L}_1} S(\widehat{X})|^2. \quad (4.1.13)$$

Подпространства \mathbb{L} и \mathbb{L}_1 , участвующие в формулировке линейной гипотезы и линейной модели, обычно бывают заданы явно, и проекции на них легко вычисляются.

К сожалению, описанные правила проверки линейных гипотез не являются свободными от распределения (непараметрическими): точный уровень значимости может отличаться от номинального значения ε . Однако при естественных свойствах матриц A и B и распределении ошибок истинный уровень значимости (остающийся нам неизвестным) при $n \rightarrow \infty$ стремится к ε (как мы покажем ниже). Впервые для ранговых критериев такой метод действий был предложен Дж.Ходжесом и Э.Леманом (см. [78]).

§ 4.2. Асимптотические свойства знаковых критериев для линейных гипотез

Изучим асимптотические свойства правил из § 4.1 при альтернативах, близких к гипотезе (4.1.4). *Близкими к гипотезе (4.1.4)* мы будем называть такие альтернативы, когда при $n \rightarrow \infty$ вектор τ в (4.1.3) принадлежит области

$$\{\tau: |\tau| < Cn^{-1/2+\gamma}\}, \quad (4.2.1)$$

где $C > 0$ — произвольная постоянная, γ фиксировано, $0 < \gamma < 1/4$.

Относительно матриц A и B примем условия, аналогичные условиям 2.1.7, 2.1.8.

У с л о в и е 4.2.1.

$$\max_{1 \leq i \leq n} \max_{1 \leq \alpha \leq r_1} |a_{i\alpha}| < K, \quad \max_{1 \leq i \leq n} \max_{1 \leq \beta \leq r_2} |b_{i\beta}| < K.$$

$= 1\}$. Умножим вектор $\sum_{i=1}^n \mathbf{a}_i \operatorname{sign}(x_i - \mathbf{a}_i \sigma)$ на \mathbf{e} . К полученному скалярному произведению применим неравенство Коши:

$$\frac{1}{n} \left(\sum_{i=1}^n (\mathbf{a}_i \mathbf{e}) \operatorname{sign}(\xi_i + b_i \tau - (\mathbf{a}_i \mathbf{e}) |\sigma|) \right)^2 \leq \frac{1}{n} \left(\sum_{i=1}^n \mathbf{a}_i \operatorname{sign}(x_i - \mathbf{a}_i \sigma) \right)^2.$$

Положим $d_i = \mathbf{a}_i \mathbf{e}$ и рассмотрим

$$Z_n := -\frac{1}{\sqrt{n}} \sum_{i=1}^n d_i \operatorname{sign}(\xi_i + b_i \tau - d_i |\sigma|). \quad (4.2.4)$$

Так как относительно $|\sigma|$ эта функция не убывает, своего наименьшего значения в области $\{\sigma: |\sigma| \geq n^{-1/2+\gamma}\}$ при фиксированном \mathbf{e} , $|\mathbf{e}| = 1$, она достигает при $|\sigma| = n^{-1/2+\gamma}$. Мы покажем, что при этом $Z_n \xrightarrow{P} \infty$. Так как дисперсия Z_n ограничена, для этого достаточно доказать, что $\mathbf{E}Z_n \rightarrow \infty$ для близких альтернатив. Имеем для $|\sigma| = n^{-1/2+\gamma}$

$$\begin{aligned} \mathbf{E}Z_n &= -\frac{1}{\sqrt{n}} \sum_{i=1}^n d_i (1 - 2F(-b_i \tau + d_i n^{-1/2+\gamma})) = \\ &= 2 \frac{1}{\sqrt{n}} \sum_{i=1}^n d_i f(\kappa_i) (b_i \tau - d_i n^{-1/2+\gamma}), \end{aligned}$$

где κ_i — некоторая точка между 0 и $b_i \tau - d_i n^{-1/2+\gamma}$. Пусть $\omega_n := \max(|\kappa_1|, \dots, |\kappa_n|)$, причем $\omega_n \rightarrow 0$. Заметив, что $\sum_{i=1}^n d_i b_i \tau = \mathbf{e}^T \left(\sum_{i=1}^n \mathbf{a}_i^T b_i \right) \tau = 0$ в силу (4.1.2), находим

$$\mathbf{E}Z_n = \frac{2}{\sqrt{n}} f(0) n^{-1/2+\gamma} \sum_{i=1}^n d_i + 2 \sum_{i=1}^n (f(\kappa_i) - f(0)) d_i (b_i \tau - d_i n^{-1/2+\gamma}). \quad (4.2.5)$$

Первое слагаемое в (4.2.5) эквивалентно

$$2 f(0) \mathbf{e}^T \left(\frac{1}{n} \mathbf{A}^T \mathbf{A} \right) \mathbf{e} n^\gamma \rightarrow \infty.$$

Второе слагаемое бесконечно мало по сравнению с первым, так как по абсолютному значению не превосходит (с учетом условия 2.1.6)

$$\frac{2}{\sqrt{n}} n \omega_n \left(\sum_{i=1}^n |d_i| + \sum_{i=1}^n |d_i|^2 \right) n^{-1/2+\gamma} \leq \text{const} \cdot \omega_n n^\gamma.$$

Таким образом, для каждого $e \in S^{r-1}$ мы получили, что $Z_n \xrightarrow{P} \infty$. Рассуждая, как при доказательстве теоремы 3.2.2, нетрудно показать, что $\sup_{e \in S^{r-1}} Z_n \xrightarrow{P} \infty$.

Обратимся к области $\{\sigma: |\sigma| < n^{-1/2+\gamma}\}$. Пусть $s = \sqrt{n}\sigma$, $|s| < n^\gamma$. Рассмотрим в этой области экстремальную задачу (4.1.5), которая теперь имеет вид

$$\frac{1}{\sqrt{n}} \left| A^T S \left(X - \frac{1}{\sqrt{n}} A s \right) \right| \Rightarrow \min_{\{s \in \mathbb{R}^{r_1}, |s| < n^\gamma\}}. \quad (4.2.6)$$

Обозначим решение (4.2.6) через \hat{s}_n . В силу (4.2.2) решение задачи (4.2.6) для близких альтернатив эквивалентно решению задачи

$$\left| \frac{1}{\sqrt{n}} A^T S(\xi) - 2f(0) \left(\frac{1}{n} A^T A \right) s \right| \Rightarrow \min_{\{s \in \mathbb{R}^{r_1}, |s| < n^\gamma\}}. \quad (4.2.7)$$

Если отвлечься от условия $|s| < n^\gamma$, то решение (4.2.7) есть

$$s_n^* = \frac{1}{2f(0)} \left(\frac{1}{n} A^T A \right)^{-1} \left(\frac{1}{\sqrt{n}} A^T S(\xi) \right). \quad (4.2.8)$$

Асимптотически (4.2.8) подчиняется нормальному распределению с параметрами 0 и $(2f(0))^{-2} \Sigma_1^{-1}$. Поэтому для достаточно больших n этот случайный вектор со сколь угодно близкой к 1 вероятностью удовлетворяет условию $|s_n^*| < n^\gamma$. Следовательно, случайный вектор s_n^* эквивалентен \hat{s}_n — решению задачи (4.2.6). При этом значение целевой функции из (4.2.6) в точке минимума ограничено по вероятности. Мы получили, что \hat{s}_n действительно асимптотически эквивалентно точке глобального минимума (4.1.5) (в том смысле, что разность между ними по вероятности стремится к нулю), когда τ удовлетворяет (4.2.1). Теорема доказана. \square

Теорема 4.2.1 позволяет указать асимптотическую мощность при близких альтернативах для различных знаковых критериев. Близкие альтернативы (4.2.1) при этом будет удобнее представлять в виде

$$\{t: |t| < Cn^\gamma, t \in \mathbb{R}^{r_2}\} \quad (4.2.9)$$

и $\tau = t/\sqrt{n}$.

Теорема 4.2.2. Для модели (4.1.3) в условиях теоремы 4.2.1 при близких альтернативах (4.2.9) нормированная статистика из критерия (4.1.8) сходится по вероятности:

$$\frac{1}{n} \left| B^T S(\widehat{X}) \right|^2 \xrightarrow{P} |Y|^2, \quad Y \sim N(2f(0) \Sigma_2 t, \Sigma_2); \quad (4.2.10)$$

статистика из критерия (4.1.10) сходится по вероятности:

$$S^T(\widehat{X})B(B^T B)^{-1}B^T S(\widehat{X}) \xrightarrow{P} \chi^2(\Delta, r), \quad (4.2.11)$$

где параметр нецентральности

$$\Delta = 2f(0)\sqrt{t^T \Sigma_2 t}. \quad (4.2.12)$$

С л е д с т в и е 4.2.1. Реальные уровни правил (4.1.8), (4.1.10) стремятся к номинальным при $n \rightarrow \infty$.

Д о к а з а т е л ь с т в о. Для доказательства достаточно заметить, что в перечисленных условиях теорема 3.2.4 (аналогично (4.2.2)) дает

$$\frac{1}{\sqrt{n}} B^T S(\widehat{X}) = \frac{1}{\sqrt{n}} B^T S(\xi) + 2f(0) \left(\frac{1}{n} B^T B \right) t + \omega_n,$$

где $\sup_{|t| < Cn^\gamma} |\omega_n| \xrightarrow{P} 0$, и

$$\frac{1}{\sqrt{n}} B^T S(\xi) \xrightarrow{d} N(0, \Sigma). \quad \square$$

Теорема 4.2.2 позволяет получить полезные следствия и для простых гипотез. В модели (2.1.5) при проверке знаковыми критериями (2.2.18) и (2.2.21) простой гипотезы $H_0: \theta = \theta^0$ против близких альтернатив $H: \theta = \theta^0 + \frac{1}{\sqrt{n}} t$, $|t| < \text{const} \cdot n^\gamma$, $0 < \gamma < 1/4$, мы получаем аналогичные результаты, заменяя в (4.2.10)–(4.2.12) B и Σ_2 на C и Σ соответственно.

§ 4.3. Примеры

П р и м е р 1. (проверка гипотезы об однородности двух выборок, которые могут отличаться сдвигом). Пусть x_1, \dots, x_m и y_1, \dots, y_n — две независимые выборки. Представим их элементы в виде

$$\begin{aligned} x_i &= a + \varepsilon_i, & i &= 1, \dots, m, \\ y_j &= b + \xi_j, & j &= 1, \dots, n, \end{aligned} \quad (4.3.1)$$

где a и b суть медианы распределений, а независимые случайные величины $\varepsilon_1, \dots, \varepsilon_m, \xi_1, \dots, \xi_n$ обладают свойством

$$P\{\varepsilon_i > 0\} = P\{\varepsilon_i < 0\} = P\{\xi_j > 0\} = P\{\xi_j < 0\} = 0.5.$$

Обсуждаемая гипотеза однородности сводится к

$$H: a = b. \quad (4.3.2)$$

Модель (4.3.1) можно представить в виде линейной модели (2.1.5), если положить

$$X = (x_1, \dots, x_m, y_1, \dots, y_n)^T, \quad L = \mathbb{L}\{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}, \quad (4.3.3)$$

где $(\mathbf{1}, \mathbf{0})^T = (\underbrace{1, \dots, 1}_m, \underbrace{0, \dots, 0}_n)$, $(\mathbf{0}, \mathbf{1})^T = (\underbrace{0, \dots, 0}_m, \underbrace{1, \dots, 1}_n)$.

Участвующее в формулировке линейной гипотезы (4.1.1) линейное подпространство \mathbb{L}_1 порождается вектором $(\mathbf{1}, \mathbf{1})^T$. Легко убедиться, что ортогональное дополнение \mathbb{L}_1 до \mathbb{L} , т.е. подпространство \mathbb{L}_2 , порождается, например, вектором $m^{-1}(\mathbf{1}, \mathbf{0}) - n^{-1}(\mathbf{0}, \mathbf{1})$. В форме (4.1.3) обсуждаемая задача (4.3.1) имеет вид

$$\begin{aligned} x_i &= \sigma + \frac{\tau}{m} + \varepsilon_i, & i &= 1, \dots, m, \\ y_j &= \sigma - \frac{\tau}{n} + \xi_j, & j &= 1, \dots, n, \end{aligned} \quad (4.3.4)$$

а гипотеза (4.3.2) в форме (4.1.4) теперь есть

$$H: \tau = 0. \quad (4.3.5)$$

Проведем проверку гипотезы (4.3.5) по схеме из § 4.1.

Шаг 1. Считая $\tau = 0$, оцениваем σ . Знаковая оценка σ есть

$$\hat{\sigma} = \text{med}(x_1, \dots, x_m, y_1, \dots, y_n). \quad (4.3.6)$$

Шаг 2. От исходных наблюдений (4.3.1) переходим к

$$\begin{aligned} \hat{x}_i &= x_i - \hat{\sigma}, & i &= 1, \dots, m, \\ \hat{y}_j &= y_j - \hat{\sigma}, & j &= 1, \dots, n. \end{aligned} \quad (4.3.7)$$

Шаг 3. К выровненным наблюдениям (4.3.7) как к независимым переменным применим знаковый критерий для проверки (4.3.5) в схеме

$$\begin{aligned} \hat{x}_i &= \frac{\tau}{m} + \varepsilon_i, & i &= 1, \dots, m, \\ \hat{y}_j &= -\frac{\tau}{n} + \xi_j, & j &= 1, \dots, n. \end{aligned}$$

Статистика такого знакового критерия есть

$$\frac{mn}{m+n} \left(\frac{1}{m} \sum_{i=1}^m \text{sign } \hat{x}_i - \frac{1}{n} \sum_{j=1}^n \text{sign } \hat{y}_j \right)^2. \quad (4.3.8)$$

Окончательно правило проверки (4.3.2) в модели (4.3.1) выглядит так: отвергнуть гипотезу об однородности двух выборок (против двусторонних альтернатив сдвига) на приближенном уровне значимости 2α , если

$$\sqrt{\frac{mn}{m+n}} \left| \frac{1}{m} \sum_{i=1}^m \text{sign } \hat{x}_i - \frac{1}{n} \sum_{j=1}^n \text{sign } \hat{y}_j \right| > z_{\alpha}, \quad (4.3.9)$$

где z_{α} — верхняя 100α -процентная точка распределения случайной величины

$$\sqrt{\frac{mn}{m+n}} \left(\frac{1}{m} \sum_{i=1}^m \zeta_i - \frac{1}{n} \sum_{j=1}^n \zeta'_j \right). \quad (4.3.10)$$

Здесь случайные величины $\zeta_1, \dots, \zeta_m, \zeta'_1, \dots, \zeta'_n$ независимы и одинаково распределены, причем

$$P\{\zeta_i = 1\} = P\{\zeta_i = -1\} = 1/2.$$

При больших m и n для (4.3.10), естественно, действует нормальная аппроксимация $N(0, 1)$. Против односторонних альтернатив критерий (4.3.9) необходимо соответствующим образом изменить.

Статистику (4.3.8) можно получить и по методу (4.1.13). Пусть

$$S(\widehat{X}) = (\text{sign } \hat{x}_1, \dots, \text{sign } \hat{x}_m, \text{sign } \hat{y}_1, \dots, \text{sign } \hat{y}_n)^T.$$

Легко видеть, что

$$|\text{proj}_{\mathbb{E}} S(\widehat{X})|^2 = \frac{1}{m} \left(\sum_{i=1}^m \text{sign } \hat{x}_i \right)^2 + \frac{1}{n} \left(\sum_{j=1}^n \text{sign } \hat{y}_j \right)^2,$$

$$|\text{proj}_{\mathbb{L}_1} S(\widehat{X})|^2 = \frac{1}{m+n} \left(\sum_{i=1}^m \text{sign } \hat{x}_i + \sum_{j=1}^n \text{sign } \hat{y}_j \right)^2.$$

Теперь нетрудно проверить, что (4.1.13) превращается в (4.3.8).

Статистика (4.3.8) по структуре напоминает другие известные статистики, применяемые для проверки (4.3.2). Например, статистику Стьюдента (стьюдентово отношение) и статистику ранговых сумм Уилкоксона.

Критерий Стьюдента можно применять для проверки (4.3.2), если случайные ошибки в (4.3.1) распределены одинаково и нормально. В числителе стьюдентова отношения стоит

$$|\bar{x} - \bar{y}| = \left| \frac{1}{m} \sum_{i=1}^m \hat{x}_i - \frac{1}{n} \sum_{j=1}^n \hat{y}_j \right|.$$

От этого выражения (4.3.8) отличается тем, что величины \hat{x}_i, \hat{y}_j заменены в нем функциями sign , т.е. знаками.

Критерий ранговых сумм Уилкоксона можно применять для проверки (4.3.2), когда ошибки в (4.3.1) распределены одинаково и непрерывно. Традиционная форма статистики Уилкоксона — это

$$\left| \sum_{j=1}^n \text{rank } y_j - \frac{n(m+n+1)}{2} \right|,$$

где $\text{rank } y_j$ обозначает ранг y_j среди всех $m+n$ наблюдений из (4.3.1). Легко видеть, что эта статистика лишь постоянным множителем отличается от статистики

$$\left| \frac{1}{n} \sum_{j=1}^n \text{rank } y_j - \frac{1}{m} \sum_{i=1}^m \text{rank } x_i \right|.$$

Эта последняя имеет ту же структуру, что и (4.3.8), с той разницей, что случайные величины (4.3.7) в ней заменены их рангами, а не знаками. (Заметим, что ранги \hat{x}_i, \hat{y}_j в совокупности (4.3.7) совпадают с рангами x_i, y_j в совокупности (4.3.1).)

Пример 2 (параллельность двух линий регрессии). Предположим, что имеются две группы наблюдений, для каждой из которых справедлива модель простой линейной регрессии. Наша цель — проверить гипотезу о равенстве коэффициентов наклона. Рассмотрим упрощенную постановку задачи, когда обе регрессионные прямые проходят через начало координат. В этом случае все необходимые статистические оценки можно указать явно.

Пусть наблюдаются случайные величины x_i, y_j , для которых действуют соотношения

$$\begin{aligned} x_i &= as_i + \varepsilon_i, & i &= 1, \dots, m, \\ y_j &= bt_j + \xi_j, & j &= 1, \dots, n. \end{aligned} \quad (4.3.11)$$

Здесь a и b — неизвестные параметры, значения $s_1, \dots, s_m, t_1, \dots, t_n$ заданы, $\varepsilon_1, \dots, \varepsilon_m, \xi_1, \dots, \xi_n$ — независимые случайные величины, медианы которых равны нулю. Подлежащая проверке гипотеза может быть выражена так:

$$H: a = b. \quad (4.3.12)$$

Модель (4.3.11) можно представить как линейную модель (2.1.5), если положить

$$X = (x_1, \dots, x_m, y_1, \dots, y_n)^T, \quad L = L\{(s, 0)^T, (0, t)^T\},$$

где

$$(s, 0)^T = (s_1, \dots, s_m, \underbrace{0, \dots, 0}_n), \quad (0, t)^T = (\underbrace{0, \dots, 0}_m, t_1, \dots, t_n).$$

Линейное подпространство \mathbb{L}_1 в представлении гипотезы (4.1.1) порождается вектором $(s, t)^T = (s_1, \dots, s_m, t_1, \dots, t_n)^T$. Ортогональное дополнение \mathbb{L}_1 до всего \mathbb{L} , т.е. подпространство \mathbb{L}_2 , порождается вектором

$$|t|^2(s, 0)^T - |s|^2(0, t)^T.$$

В новой параметризации, создаваемой этими подпространствами, модель (4.3.11) приобретает вид

$$\begin{aligned} x_i &= \sigma s_i + \tau |t|^2 s_i + \varepsilon_i, & i = 1, \dots, m, \\ y_j &= \sigma t_j - \tau |s|^2 t_j + \xi_j, & j = 1, \dots, n. \end{aligned} \quad (4.3.13)$$

Соответственно, гипотеза (4.3.12) превращается в

$$H: \tau = 0. \quad (4.3.14)$$

Применим для ее проверки схему из § 4.1.

Шаг 1. Считая, что $\tau = 0$, оцениваем σ . Согласно (1.2.12) знаковая оценка σ есть медиана дискретного распределения вероятностей, когда вероятности сосредоточены в точках вида x_i/s_i , $i = 1, \dots, m$, и y_j/t_j , $j = 1, \dots, n$, причем вероятность точки x_i/s_i пропорциональна $|s_i|$, а вероятность точки y_j/t_j пропорциональна $|t_j|$.

Шаг 2. Производим выравнивание наблюдений:

$$\begin{aligned} \hat{x}_i &= x_i - \hat{\sigma} s_i, & i = 1, \dots, m, \\ \hat{y}_j &= y_j - \hat{\sigma} t_j, & j = 1, \dots, n. \end{aligned} \quad (4.3.15)$$

Шаг 3. Предполагая, что наблюдения (4.3.15) независимы и удовлетворяют модели

$$\begin{aligned} \hat{x}_i &= \tau |t|^2 s_i + \varepsilon_i, & i = 1, \dots, m, \\ \hat{y}_j &= -\tau |s|^2 t_j + \xi_j, & j = 1, \dots, n, \end{aligned}$$

проверяем по этим данным гипотезу (4.3.14). Статистика знакового критерия согласно (2.1.19) или (2.1.21) имеет вид

$$\frac{1}{|s|^2} \sum_{i=1}^m s_i \operatorname{sign} \hat{x}_i - \frac{1}{|t|^2} \sum_{j=1}^n t_j \operatorname{sign} \hat{y}_j.$$

Окончательно правило для проверки гипотезы (4.3.12) можно сформулировать так: в схеме (4.3.11) отвергнуть гипотезу (4.3.12) на приближенном уровне значимости 2α , если

$$\left| \frac{1}{|s|^2} \sum_{i=1}^m s_i \operatorname{sign} \hat{x}_i - \frac{1}{|t|^2} \sum_{j=1}^n t_j \operatorname{sign} \hat{y}_j \right| > z_{\alpha},$$

Упомянутые в § 4.1 подпространства L_1 и L_2 здесь таковы: L_1 — одномерное подпространство, порожденное вектором e_0 ;

$$L_2 = \left\{ l: l = \sum_{j=1}^t \frac{\alpha_j}{n_j} e_j, \sum_{j=1}^t \alpha_j = 0 \right\}.$$

Легко проверить, что при условии (4.4.2) соблюдаются условия ортогональности (4.1.2). Применим для проверки линейной гипотезы (4.4.3) процедуру из § 4.2.

Шаг 1. Считаем, что гипотеза H верна. В таком случае в задаче остается единственный неизвестный параметр μ , знаковой оценкой $\hat{\mu}$ которого служит медиана всей выборочной совокупности, т.е.

$$\hat{\mu} = \text{med}(x_{11}, x_{21}, \dots, x_{n_1 1}, x_{12}, \dots, x_{n_2 2}, \dots, x_{1t}, \dots, x_{n_t t}).$$

Шаг 2. Переходим к выравненным наблюдениям

$$\hat{x}_{ij} = x_{ij} - \hat{\mu}.$$

Шаг 3. К величинам \hat{x}_{ij} как к независимым применяем знаковый критерий в схеме (4.4.1), (4.4.2) для проверки гипотезы (4.4.3), считая $\mu = 0$.

Чтобы конкретизировать этот шаг, надо выбрать определенный знаковый критерий. Здесь у нас есть по крайней мере две возможности (о них уже упоминалось в § 4.1) — критерии (2.2.18) и (2.2.21). Начнем с критерия типа (2.2.18). Забудем на время о связи (4.4.2), наложенной на параметры, и рассмотрим t выборок

$$y_{ij} = \frac{\theta_j}{n_j} + \varepsilon_{ij}, \quad i = 1, \dots, n_j,$$

где $\theta_1, \dots, \theta_t$ — независимые переменные. Действуя так же, как в § 2.2, можно (при условиях 2.1.1–2.1.3) получить асимптотическое представление для правдоподобия статистики знаков

$$S(Y) = \|\text{sign } y_{ij}\|, \quad i = 1, \dots, n_j, \quad j = 1, \dots, t.$$

Оно равно

$$\begin{aligned} & \left(\frac{1}{2}\right)^N \left\{ 1 + (2f(0)) \sum_{j=1}^t \frac{\theta_j}{n_j} \sum_{i=1}^{n_j} \text{sign } y_{ij} + \right. \\ & \left. + \frac{1}{2} (2f(0))^2 \left[\left(\sum_{j=1}^t \frac{\theta_j}{n_j} \sum_{i=1}^{n_j} \text{sign } y_{ij} \right)^2 - \sum_{j=1}^t \frac{\theta_j^2}{n_j} \right] + o(|\theta|^2) \right\}, \quad (4.4.4) \end{aligned}$$

где $N = n_1 + \dots + n_t$. Статистика (2.2.18) знакового критерия максимальной средней кривизны при этом равна

$$\sum_{j=1}^t \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \text{sign } y_{ij} \right)^2. \quad (4.4.5)$$

Нас, однако, интересуют свойства функции мощности (а именно ее средняя кривизна в точке $\theta = 0$) при наличии связи (4.4.2). Чтобы найти ее, заметим, что среднюю кривизну можно представить как сумму средних кривизны на подпространстве, заданном (4.4.2), и ортогональном дополнении к нему. Последнее подпространство одномерно. Его направляющий орт есть $(t^{-1/2}, \dots, t^{-1/2})^T$. Рассмотрим правдоподобие (4.4.5) вдоль этого направления, положив $\theta_j = u/\sqrt{t}$, $j = 1, \dots, m$, $u \in \mathbb{R}^1$. Статистика знакового критерия максимальной кривизны (определяемая коэффициентом при u^2) при этом оказывается равной

$$\frac{1}{t} \left(\sum_{j=1}^t \frac{1}{n_j} \sum_{i=1}^{n_j} \text{sign } y_{ij} \right)^2. \quad (4.4.6)$$

Теперь статистикой знакового критерия (2.2.18) в схеме (4.4.1), (4.4.2) для проверки гипотезы (4.4.3) (при $\mu = 0$, как того требуют условия шага 3) служит разность выражений (4.4.5) и (4.4.6). Окончательно правило проверки гипотезы (4.4.3) выглядит так:

отвергнуть гипотезу H (4.4.3) на приближенном уровне значимости α , $0 < \alpha < 1$, если

$$\sum_{j=1}^t \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \text{sign } \hat{x}_{ij} \right)^2 - \frac{1}{t} \left(\sum_{j=1}^t \frac{1}{n_j} \sum_{i=1}^{n_j} \text{sign } \hat{x}_{ij} \right)^2 > q_{1-\alpha},$$

где $q_{1-\alpha}$ есть верхнее α -значение случайной величины

$$\sum_{j=1}^t \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \zeta_{ij} \right)^2 - \frac{1}{t} \left(\sum_{j=1}^t \frac{1}{n_j} \sum_{i=1}^{n_j} \zeta_{ij} \right)^2. \quad (4.4.7)$$

Здесь ζ_{ij} ($i = 1, \dots, n_j$; $j = 1, \dots, t$) — независимые случайные величины, принимающие значения 1 и -1 с вероятностями 0.5.

Обратимся теперь ко второй из отмеченных выше возможностей и выберем для шага 3 знаковый критерий (2.2.21). Как было указано в (4.1.12), статистикой этого критерия служит квадрат длины проекции вектора

$$\|\text{sign } \hat{x}_{ij}\|^T, \quad i = 1, \dots, n_j, \quad j = 1, \dots, t,$$

на подпространство \mathbb{L}_2 . Эта величина легко вычисляется. Она равна

$$\sum_{j=1}^t n_j \left(n_j^{-1} \sum_{i=1}^{n_j} \text{sign } \hat{x}_{ij} - \left(\sum_{j=1}^t n_j \right)^{-1} \sum_{j=1}^t \sum_{i=1}^{n_j} \text{sign } \hat{x}_{ij} \right)^2.$$

Заметим, что в силу определения $\hat{\mu}$ как медианы всей совокупности наблюдений x_{ij}

$$\sum_{j=1}^t \sum_{i=1}^{n_j} \text{sign } \hat{x}_{ij} = 0.$$

Поэтому правило для проверки гипотезы (4.4.3) в модели (4.4.1), (4.4.2) с применением критерия (2.2.21) формулируется так:

отвергнуть гипотезу H (4.4.3) на приближенном уровне значимости α , если

$$\sum_{j=1}^t n_j^{-1} \left(\sum_{i=1}^{n_j} \text{sign } \hat{x}_{ij} \right)^2 > q_{1-\alpha}, \quad (4.4.8)$$

где $q_{1-\alpha}$ — верхнее α -значение случайной величины

$$\sum_{j=1}^t n_j \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \zeta_{ij} - \sum_{j=1}^t \sum_{i=1}^{n_j} \zeta_{ij} / \sum_{j=1}^t n_j \right)^2. \quad (4.4.9)$$

Случайные величины ζ_{ij} были определены выше. При больших n_1, \dots, n_t в качестве $q_{1-\alpha}$ можно использовать квантили распределения χ^2 с $t - 1$ степенями свободы, так как статистика (4.4.8) при гипотезе асимптотически распределена как хи-квадрат с числом степеней свободы, равным размерности подпространства \mathbb{L}_2 , когда $n_1, \dots, n_t \rightarrow \infty$.

4.4.2. Двухфакторный анализ. Рассмотрим аддитивную схему действия факторов и план эксперимента с одинаковым числом наблюдений при каждой комбинации уровней факторов (m наблюдений в клетке). Мы уже рассматривали эту схему в § 2.4. Воспользуемся введенными там обозначениями. Анализ двухфакторной таблицы обыкновенно начинают с проверки гипотезы об отсутствии действия одного из факторов (отсутствие эффекта обработки), т.е. гипотезы

$$H: \beta_1 = \dots = \beta_t = 0. \quad (4.4.10)$$

По схеме § 4.1 проверка гипотезы (4.4.10) начинается со знакового оценивания остальных параметров и выравнивания наблюдений. Очевидно, что знаковая оценка $\mu + \alpha_i$, $i = 1, \dots, r$, является медианой всех наблюдений из строки i , т.е.

$$(\mu + \alpha_i)^* = \text{med}(x_{i11}, \dots, x_{i1m}, x_{i21}, \dots, x_{itm}).$$

Поэтому

$$\hat{x}_{ijk} = x_{ijk} - (\mu + \alpha_i)^*$$

для всех i, j, k .

На шаге 3 проверки гипотезы используем критерий (2.2.21), которым и ограничимся на этот раз.

В § 4.1 мы отметили, что эта знаковая статистика совпадает (с точностью до постоянного множителя) с числителем F -статистики, которая в гауссовской постановке служит для проверки интересующей нас гипотезы. Применительно к наблюдениям (2.4.1) F -отношение для проверки (4.4.10) равно

$$\frac{trm - r - t - 1}{t - 1} \frac{mr \sum_{j=1}^t (x_{.j} - x_{...})^2}{\sum_{i=1}^r \sum_{j=1}^t \sum_{k=1}^m (x_{ijk} - x_{i..} - x_{.j.} + x_{...})^2}, \quad (4.4.11)$$

где замена индекса точкой означает результат усреднения по этому индексу (см., например, [3, 48, § 7.3, В] или [17, § 5.5.3]). Остается лишь заменить наблюдения x_{ijk} на $\text{sign } x_{ijk}$. Для удобства положим

$$z_{ij} = \frac{1}{\sqrt{m}} \sum_{k=1}^m \text{sign } \hat{x}_{ijk}.$$

Теперь правило проверки гипотезы H (4.4.10) выглядит так:

отвергнуть гипотезу H (4.4.10) на приближенном уровне значимости α , если

$$r \sum_{j=1}^t (z_{.j} - z_{...})^2 > q_{1-\alpha}, \quad (4.4.12)$$

где $q_{1-\alpha}$ — верхнее α -значение случайной величины

$$r m \sum_{j=1}^t (\zeta_{.j} - \zeta_{...})^2. \quad (4.4.13)$$

Относительно независимых одинаково распределенных случайных величин ζ_{ijk} , как всегда, действует предположение

$$P\{\zeta_{ijk} = 1\} = P\{\zeta_{ijk} = -1\} = 1/2.$$

Асимптотически, при $m \rightarrow \infty$, случайная величина (4.4.13) распределена как $\chi^2(t-1)$. Поэтому при достаточно больших m в качестве $q_{1-\alpha}$ можно взять соответствующую квантиль распределения хи-квадрат с $t-1$ степенями свободы.

§ 4.5. Вычисление критических значений
в задачах проверки линейных гипотез

В этом параграфе мы обсудим методы вычисления критических значений в задачах проверки линейных гипотез и приведем таблицы таких критических значений при некоторых значениях параметров.

4.5.1. Однофакторный анализ. Мы остановимся на вычислении критических значений случайной величины (4.4.9).

Пусть $B_j, j = 1, \dots, t$, суть независимые случайные величины, распределенные по биномиальным законам $Bi(n_j, 1/2)$. Заметим, что $(\zeta_{ij} + 1)/2$ принимает значения 1 и 0 с вероятностями $1/2$. Поэтому

$$\sum_{i=1}^{n_j} \zeta_{ij} = 2B_j + n_j.$$

Следовательно, (4.4.9) равно

$$\zeta_t = 4 \sum_{j=1}^t n_j \left(\frac{B_j}{n_j} - \sum_{j=1}^t B_j / \sum_{j=1}^t n_j \right)^2. \quad (4.5.1)$$

Квантильные значения статистики (4.5.1) при малых и умеренных $n := \prod_{j=1}^m (n_j + 1)$ можно вычислить точно путем перебора. При других n приближенные квантильные значения нетрудно вычислить моделированием распределения случайной величины (4.5.1) методом Монте-Карло. В табл. 4.5.1 приведены некоторые «хвостовые» верхние значения q и соответствующие им точные вероятности $P\{\zeta_t \leq q\}$ для случайной величины (4.5.1) при $t = 2, n_1 = 5, \dots, 15, n_2 = 5, \dots, 15$.

Таблица 4.5.1

Квантили случайной величины ζ_2 (4.5.1)

	q	$P\{\zeta_2 \leq q\}$		q	$P\{\zeta_2 \leq q\}$
$n_1=5, n_2=5$	0.400	0.656	$n_1=5, n_2=6$	2.048	0.842
	1.600	0.891		2.376	0.915
	3.600	0.979		4.376	0.974
$n_2=7$	2.143	0.859	$n_2=8$	2.223	0.893
	3.086	0.944		2.777	0.913
	3.810	0.961		3.723	0.964
$n_2=9$	1.834	0.861	$n_2=10$	1.200	0.796
	2.800	0.928		2.133	0.901
	3.968	0.955		3.333	0.959
$n_2=11$	2.405	0.865	$n_2=12$	2.075	0.877
	2.618	0.916		3.075	0.937
	3.823	0.958		3.529	0.951

Таблица 4.5.1 (продолжение)

Квантили случайной величины ζ_2 (4.5.1)

	q	$P\{\zeta_2 \leq q\}$		q	$P\{\zeta_2 \leq q\}$
$n_2=13$	2.137	0.858	$n_2=14$	2.033	0.867
	2.492	0.907		2.890	0.930
	3.501	0.953		3.898	0.966
$n_2=15$	1.667	0.841	$n_2=16$	2.002	0.866
	2.400	0.907		2.752	0.918
	4.267	0.976		3.621	0.950
$n_1=6, n_2=6$	0.333	0.612	$n_1=6, n_2=7$	2.117	0.899
	1.333	0.854		2.374	0.908
	3.000	0.961		3.875	0.969
$n_2=8$	2.381	0.893	$n_2=9$	2.178	0.886
	2.881	0.934		2.844	0.926
	4.024	0.960		3.600	0.958
$n_2=10$	2.017	0.871	$n_2=11$	2.228	0.869
	2.817	0.922		2.410	0.907
	3.750	0.951		3.651	0.960
$n_2=12$	1.778	0.866	$n_2=13$	2.116	0.856
	2.778	0.935		2.429	0.901
	4.000	0.973		3.498	0.954
$n_2=14$	1.867	0.851	$n_2=15$	2.305	0.888
	2.752	0.922		2.743	0.913
	3.810	0.961		3.733	0.954
$n_1=7, n_2=7$	1.143	0.820	$n_1=7, n_2=8$	1.905	0.878
	2.571	0.943		2.976	0.901
	4.571	0.987		3.471	0.961
$n_2=9$	2.286	0.892	$n_2=10$	2.101	0.859
	2.683	0.919		2.827	0.925
	3.813	0.957		3.661	0.953
$n_2=11$	2.104	0.876	$n_2=12$	2.256	0.871
	2.773	0.918		2.566	0.910
	3.951	0.951		3.812	0.961
$n_2=13$	1.978	0.862	$n_2=14$	1.524	0.834
	2.848	0.907		2.381	0.911
	4.064	0.962		3.429	0.957
$n_2=15$	2.002	0.871	$n_2=16$	2.242	0.880
	2.771	0.911		2.484	0.902
	3.990	0.967		3.728	0.952
$n_1=8, n_2=8$	1.000	0.790	$n_1=8, n_2=9$	1.729	0.855
	2.250	0.923		2.941	0.934
	4.000	0.979		3.346	0.951

Т а б л и ц а 4.5.1 (продолжение)

Квантили случайной величины ζ_2 (4.5.1)

	q	$P\{\zeta_2 \leq q\}$		q	$P\{\zeta_2 \leq q\}$	
$n_2=10$	2.178	0.887	$n_2=11$	2.299	0.884	
	2.500	0.903		2.766	0.921	
	3.600	0.954		3.828	0.951	
$n_2=12$	2.133	0.879	$n_2=13$	1.995	0.859	
	2.700	0.917		2.645	0.912	
	4.033	0.966		3.875	0.961	
$n_2=14$	2.104	0.866	$n_2=15$	1.878	0.855	
	2.864	0.910		2.557	0.901	
	4.058	0.960		3.623	0.955	
$n_1=9, n_2=9$	0.889	0.762	$n_1=9, n_2=10.$	2.395	0.860	
	2.000	0.904		2.704	0.927	
	3.556	0.969		4.126	0.956	
	$n_2=11$	2.069	0.876	$n_2=12$	2.286	0.886
		2.766	0.901		2.683	0.914
		3.735	0.958		4.063	0.966
$n_2=13$	2.127	0.863	$n_2=14$	2.099	0.869	
	2.741	0.917		2.672	0.912	
	3.885	0.952		3.877	0.961	
$n_2=15$	2.178	0.871	$n_2=16$	1.960	0.858	
	2.500	0.901		2.668	0.904	
	3.600	0.951		3.738	0.954	
$n_1=10, n_2=10$	0.800	0.737	$n_1=10, n_2=11$	2.244	0.859	
	1.800	0.885		2.500	0.917	
	3.200	0.959		3.825	0.954	
	$n_2=12$	1.964	0.862	$n_2=13$	2.259	0.884
		2.933	0.926		2.590	0.903
		4.097	0.956		3.901	0.961
$n_2=14$	2.100	0.858	$n_2=15$	2.160	0.879	
	2.743	0.916		2.667	0.914	
	4.005	0.962		3.840	0.960	
$n_1=11, n_2=11$	1.636	0.866	$n_1=11, n_2=12$	2.108	0.861	
	2.909	0.948		2.324	0.905	
	4.545	0.983		3.563	0.952	
	$n_2=13$	2.155	0.853	$n_2=14$	2.198	0.879
		2.798	0.922		2.702	0.903
		3.921	0.952		3.740	0.955
$n_2=15$	2.060	0.856	$n_2=16$	2.189	0.870	
	2.719	0.913		2.735	0.915	
	3.939	0.961		3.779	0.950	

Таблица 4.5.1 (продолжение)

Квантили случайной величины ζ_2 (4.5.1)

	q	$P\{\zeta_2 \leq q\}$		q	$P\{\zeta_2 \leq q\}$		
$n_1=12, n_2=12$	1.500	0.848	$n_1=12, n_2=13$	1.986	0.862		
	2.667	0.936		3.103	0.911		
	4.167	0.977		3.450	0.955		
	$n_2=14$	2.289		0.867	$n_2=15$	2.141	0.872
		2.670		0.917		2.674	0.902
	3.751	0.951		3.919	0.955		
$n_1=13, n_2=13$	1.385	0.831	$n_1=13, n_2=14$	1.875	0.859		
	2.462	0.924		2.930	0.909		
	3.846	0.971		4.220	0.950		
	$n_2=15$	2.216		0.868	$n_2=16$	2.080	0.861
		2.550		0.909		2.632	0.901
	3.798	0.960		3.830	0.952		
$n_1=14, n_2=14$	1.286	0.815	$n_1=14, n_2=15$	1.776	0.853		
	2.286	0.913		2.775	0.909		
	3.571	0.964		4.099	0.955		
$n_1=15, n_2=15$	1.200	0.800	$n_1=15, n_2=16$	1.809	0.850		
	2.133	0.901		2.634	0.908		
	3.333	0.957		3.884	0.954		

В табл. 4.5.2 мы приводим «хвостовые» верхние значения случайной величины (4.5.1) для $t > 2$ при одинаковых объемах выборок: $n = n_1 = \dots = n_t$.

Таблица 4.5.2

Квантили ζ_t (4.5.1) для $n = n_1 = \dots = n_t$

	q	$P\{\zeta_2 \leq q\}$		q	$P\{\zeta_2 \leq q\}$
$t=3, n=2$	1.333	0.719	$t=3, n=3$	2.667	0.777
	4.000	0.906		3.556	0.918
	5.333	1.000		6.222	0.988
$n=4$	2.667	0.832	$n=5$	3.733	0.900
	4.667	0.949		4.800	0.943
	6.000	0.978		6.400	0.961
$n=6$	3.111	0.850	$n=7$	3.429	0.860
	4.000	0.902		4.952	0.950
	5.778	0.972		6.095	0.965
$n=8$	4.000	0.858	$n=9$	3.852	0.897
	4.333	0.924		4.741	0.921
	6.333	0.968		5.630	0.951

Таблица 4.5.2 (продолжение)

Квантили ζ_t (4.5.1) для $n = n_1 = \dots = n_t$

	q	$P\{\zeta_2 \leq q\}$		q	$P\{\zeta_2 \leq q\}$
$n=10$	3.467	0.869	$n=11$	3.879	0.872
	5.067	0.933		4.606	0.913
	5.600	0.960		6.061	0.955
$n=12$	4.222	0.893	$n=13$	3.897	0.872
	4.667	0.929		4.308	0.912
	6.000	0.950		5.744	0.956
$n=14$	3.619	0.851	$n=15$	3.733	0.877
	4.762	0.910		4.800	0.907
	5.905	0.961		5.511	0.951
$t=4, n=2$	4.000	0.758	$t=4, n=3$	5.333	0.856
	5.500	0.945		6.333	0.909
	6.000	0.977		6.667	0.961
$n=4$	5.000	0.874	$n=5$	5.400	0.894
	6.000	0.915		6.400	0.906
	6.750	0.956		7.200	0.959
$n=6$	5.833	0.891	$n=7$	5.143	0.871
	6.000	0.918		6.143	0.907
	7.333	0.957		7.429	0.952
$n=8$	5.375	0.864	$n=9$	5.667	0.875
	6.375	0.911		6.222	0.909
	7.375	0.959		7.556	0.953
$n=10$	5.200	0.853			
	5.900	0.909			
	7.600	0.954			
$t=5, n=2$	5.600	0.834	$t=5, n=3$	6.400	0.821
	6.400	0.922		6.933	0.920
	8.000	0.980		9.067	0.973
$n=4$	6.800	0.889	$n=5$	7.040	0.867
	7.200	0.914		7.360	0.910
	9.200	0.968		8.960	0.963

Табл. 4.5.1, 4.5.2 можно использовать как в задачах однофакторного анализа, так и в задачах, сводящихся к ним. Например, в рассмотренной выше задаче проверки однородности двух выборок (см. § 4.3, пример 1) нетрудно показать, что статистика знакового критерия (4.3.8) совпадает со статистикой (4.4.8) при $t = 2$ и $n_1 = m, n_2 = n$. Значит, критические значения можно выбирать из табл. 4.5.1.

4.5.2. Двухфакторный анализ. Случайную величину из (4.4.13) в задаче проверки гипотезы об отсутствии эффекта обработки можно записать в виде

$$\frac{4}{mr} \sum_{j=1}^t \left(B_j - \frac{1}{t} \sum_{j=1}^t B_j \right)^2,$$

где B_j независимы и распределены по биномиальным законам с параметрами mr и $p = 0.5$; т.е. $B_j \sim \text{Bi}(mr, 1/2)$. Это выражение совпадает со случайной величиной (4.5.1) при $n_1 = \dots = n_t = mr$. Поэтому при проверке такой гипотезы из двухфакторного анализа для вычисления критических значений можно использовать табл. 4.5.2.

В упоминавшийся пакет Sign включены программы для вычисления распределений случайных величин (4.4.7) и (4.4.9) при различных значениях параметров m, n_1, \dots, n_m .

**АВТОРЕГРЕССИОННЫЕ МОДЕЛИ
ВРЕМЕННЫХ РЯДОВ**

Глава 5

**ПРОЦЕДУРЫ НАИМЕНЬШИХ КВАДРАТОВ
И НАИМЕНЬШИХ МОДУЛЕЙ
В ПРОСТЕЙШЕЙ АВТОРЕГРЕССИОННОЙ МОДЕЛИ**

§ 5.1. Введение

В первой части главы рассмотрена наиболее распространенная модель стационарного временного ряда — стационарная модель авторегрессии, которая в простейшем однопараметрическом случае определяется стохастическим разностным уравнением

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i \in \mathbb{Z} = \{0, \pm 1, \dots\}, \quad |\beta| < 1.$$

Случайная последовательность $\{u_i\}$, удовлетворяющая этому уравнению, называется *процессом авторегрессии*. Случайную последовательность $\{\varepsilon_i\}$ будем предполагать состоящей из независимых одинаково распределенных случайных величин (н.о.р.сл.в.) с неизвестной функцией распределения. В такой непараметрической ситуации мы будем решать задачи оценивания неизвестного коэффициента β и проверки гипотез относительно β . Наши выводы будут основываться на выборке u_0, \dots, u_n из строго стационарного решения уравнения авторегрессии. Существование и единственность такого решения гарантируются предположением $|\beta| < 1$ и моментными ограничениями на величины ε_i . Эти вопросы рассмотрены в § 5.2. В § 5.3 мы изучаем простейшую и наиболее распространенную оценку β в стационарной авторегрессии — оценку наименьших квадратов, а также свойства основанных на этой оценке тестов. Изучаются фиксированные и близкие альтернативы.

Многие из асимптотических результатов, справедливых для стационарной авторегрессии, допускают распространение и на нестационарный случай. Существует много способов задать нестационарный процесс авторегрессии, и мы не будем рассматривать в этой книге нестационарный случай систематически. Мы ограничиваемся несколькими избранными результатами. В частности, в § 5.4 мы рассматриваем простейшую нестационарную авторегрессию

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i = 1, 2, \dots, \quad \beta \in \mathbb{R}^1, \quad u_0 = 0$$

(которая отличается от стационарного случая фиксированным начальным значением u_0). В предположении, что $\{\varepsilon_i\}$ — гауссовская

последовательность н.о.р.сл.в., мы приводим результаты об асимптотическом поведении оценки наименьших квадратов (которая, кстати, совпадает в рассматриваемой ситуации с оценкой максимального правдоподобия), построенной по наблюдениям u_1, \dots, u_n . Эти результаты совершенно различны в случаях $|\beta| < 1$ (этот случай асимптотически эквивалентен стационарному), $|\beta| = 1$ (критический случай), $|\beta| > 1$ (взрывная авторегрессия).

В § 5.5 мы возвращаемся к рассмотрению стационарной авторегрессии и изучаем оценку наименьших модулей и связанные с ней тесты. Асимптотическое исследование мы проводим с помощью равномерных стохастических разложений, но не самой выпуклой целевой функции, определяющей оценку наименьших модулей (как это обычно делается), а ее «производной». Аналогичный подход систематически используется далее в гл. 6 и 7 при изучении знаковых процедур, с которыми не удается связать никакой выпуклой задачи оптимизации.

В § 5.6 мы изучаем устойчивость оценок наименьших квадратов и наименьших модулей в стационарной схеме авторегрессии к искажению данных. Мы рассматриваем схему засорения данных независимыми выбросами. В этом случае модель имеет вид

$$y_i = u_i + z_i^{\gamma} \xi_i, \quad i \in \mathbb{Z},$$

где $\{z_i^{\gamma}\}$ — последовательность независимых биномиальных величин $\text{Bi}(1, \gamma)$, γ — уровень засорения; $\{\xi_i\}$ — последовательность н.о.р.сл.в., представляющих собой выбросы; последовательности $\{u_i\}$, $\{z_i^{\gamma}\}$, $\{\xi_i\}$ предполагаются независимыми между собой. Устойчивость оценок, строящихся по y_1, \dots, y_n , мы характеризуем с помощью функционалов влияния и показываем, что оценки наименьших квадратов и наименьших модулей имеют неограниченные функционалы влияния (а значит, и чувствительности к большому ошибкам) на естественных классах распределений величин ξ_i . Качественно последнее означает, что эти оценки неустойчивы даже к малой доле больших выбросов в данных.

Естественная альтернатива простейшей стационарной схеме авторегрессии (наряду с рассмотренной в § 5.4) — нестационарная схема вида

$$u_i = \beta_i u_{i-1} + \varepsilon_i, \quad i \in \mathbb{Z}, \quad |\beta_i| < 1,$$

с коэффициентом β_i , зависящим от (временного) параметра i . В § 5.7 мы рассматриваем такую схему и строим тесты для проверки гипотезы о постоянстве последовательности $\{\beta_i\}$, т.е. гипотезы стационарности. Мы изучаем мощность тестов при близких альтернативах. Статистики тестов являются функционалами типа Колмогорова и омега-квадрат от случайных процессов, похожих на рассматриваемые в известном принципе инвариантности Донскера—Прохорова.

В этой главе мы рассматриваем главным образом те результаты, которые понадобятся нам далее в гл. 6 и 7 при изучении знаковых процедур. Мы сознательно оставляем в стороне многие аспекты, в частности, случай неизвестного i , вообще говоря, переменного среднего y

величин $\{\varepsilon_i\}$, случай, когда последовательность $\{\varepsilon_i\}$ есть последовательность скользящего среднего, и т.д. Далее по ходу изложения мы дадим ссылки, которые помогут читателю найти подобные результаты в литературе.

Напомним, что все предельные переходы совершаются при $n \rightarrow \infty$, если не оговорено противное.

§ 5.2. Простейшее стационарное уравнение авторегрессии и его решения

В следующих двух параграфах этой главы мы будем рассматривать простейшую стационарную однопараметрическую схему авторегрессии

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i \in \mathbb{Z}. \quad (5.2.1)$$

В (5.2.1) $\{\varepsilon_i\}$ — невырожденные н.о.р.сл.в. с функцией распределения $G(x) = \mathbf{P}\{\varepsilon_1 < x\}$; β — неизвестный неслучайный параметр, $|\beta| < 1$.

Нашей целью будет построение непараметрических (т.е. действенных при неизвестной $G(x)$) процедур проверки гипотез относительно параметра β и методов его оценивания. Эти процедуры будут основаны на выборке u_0, u_1, \dots, u_n из строго стационарного решения уравнения (5.2.1). Существование и единственность такого решения гарантируется моментными ограничениями на $G(x)$. Обычно считают (см., например, [2, 45, § 5.2]), что выполнено

Условие 5.2.1. $\mathbf{E}\varepsilon_1 = 0, \quad \mathbf{E}\varepsilon_1^2 < \infty.$

Условие 5.2.1 обеспечивает существование строго стационарного решения уравнения (5.2.1), которое представимо в виде

$$u_i = \sum_{j=0}^{\infty} \beta^j \varepsilon_{i-j}. \quad (5.2.2)$$

Ряд в (5.2.2) сходится в \mathbb{L}^2 , и это п.н. единственное строго стационарное решение. Мы докажем похожее утверждение при более слабом ограничении.

Условие 5.2.2. $\mathbf{E}|\varepsilon_1|^{1+\Delta} < \infty$ при некотором $\Delta > 0$.

Лемма 5.2.1. При условии 5.2.2 существует п.н. единственное строго стационарное решение уравнения (5.2.1), представимое в виде (5.2.2), причем ряд сходится в $\mathbb{L}^{1+\Delta}$.

Доказательство. Величины $\{u_i\}$ из (5.2.2) формально удовлетворяют (5.2.1). Для сходимости в $\mathbb{L}^{1+\Delta}$ ряда (5.2.2) достаточно (в силу полноты $\mathbb{L}^{1+\Delta}$) фундаментальности последовательности частных сумм

$$S_n := \sum_{j=0}^n \beta^j \varepsilon_{i-j}.$$

В силу неравенства Минковского [44, гл. 2, § 6] имеем при $m > n$

$$\begin{aligned} \left\{ \mathbf{E}_\beta |S_m - S_n|^{1+\Delta} \right\}^{1/(1+\Delta)} &= \left\{ \mathbf{E}_\beta \left| \sum_{j=n+1}^m \beta^j \varepsilon_{i-j} \right|^{1+\Delta} \right\}^{1/(1+\Delta)} \leq \\ &\leq \sum_{j=n+1}^m \left\{ \mathbf{E}_\beta |\beta^j \varepsilon_{i-j}|^{1+\Delta} \right\}^{1/(1+\Delta)} = \\ &= \left\{ \mathbf{E} |\varepsilon_1|^{1+\Delta} \right\}^{1/(1+\Delta)} \sum_{j=n+1}^m |\beta|^j = O(|\beta|^n) = o(1). \end{aligned}$$

Итак, соотношение (5.2.2) действительно определяет решение, принадлежащее $\mathbb{L}^{1+\Delta}$. Его строгая стационарность очевидна. Осталось проверить единственность. В силу (5.2.1) для любого строго стационарного решения $\{\tilde{u}_i\}$ имеем

$$\tilde{u}_i = \beta(\beta \tilde{u}_{i-2} + \varepsilon_{i-1}) + \varepsilon_i = \dots = \beta^k \tilde{u}_{i-k} + \sum_{j=0}^{k-1} \beta^j \varepsilon_{i-j}. \quad (5.2.3)$$

В силу строгой стационарности последовательности $\{\tilde{u}_i\}$ и условия $|\beta| < 1$ имеет место сходимость

$$\beta^k \tilde{u}_{i-k} \xrightarrow{\mathbf{P}_\beta} 0, \quad k \rightarrow \infty.$$

Следовательно, в силу (5.2.3) $\{\tilde{u}_i\}$ представляется правой частью (5.2.2), сходящейся по вероятности. Поскольку предел последовательности, сходящейся по вероятности, единственен п.н., то лемма доказана. \square

Отметим, что если $\mathbf{E}\varepsilon_1 = 0$, то лемма 5.2.1 влечет, что $\mathbf{E}_\beta u_1 = 0$, а при $\Delta = 1$ легко находим и дисперсию

$$\mathbf{D}_\beta u_1 = \mathbf{E}_\beta \left(\sum_{j=0}^{\infty} \beta^j \varepsilon_{i-j} \right)^2 = \sum_{j=0}^{\infty} \mathbf{E}_\beta (\beta^j \varepsilon_{i-j})^2 = \mathbf{E}\varepsilon_1^2 (1 - \beta^2)^{-1}. \quad (5.2.4)$$

Уместно отметить, что случаи $\Delta \geq 1$ и $\Delta < 1$ окажутся существенно различными для задач статистического анализа, которыми мы будем заниматься. Эти вопросы будут обсуждаться далее систематически, начиная со следующего § 5.3.

Здесь же естественно задаться вопросом о минимальных моментных ограничениях на $G(x)$, обеспечивающих существование и п.н. единственность строго стационарного решения (5.2.1). Близким к таковым является

У с л о в и е 5.2.3. $\mathbf{E} \log^+ |\varepsilon_1| < \infty$, где $\log^+ x = \max(0, \log x)$.

Следующий результат заимствован из [111].

Лемма 5.2.2. *При условии 5.2.3 существует п.н. единственное строго стационарное решение (5.2.1), представимое скользящим средним (5.2.2), и ряд в (5.2.2) сходится п.н. абсолютно.*

Доказательство. Для абсолютной сходимости п.н. ряда (5.2.2) достаточно, чтобы

$$\overline{\lim}_{j \rightarrow \infty} |\varepsilon_{i-j}|^{1/j} \leq 1 \text{ п.н.}$$

Справедливость последнего соотношения в силу леммы Бореля—Кантелли будет следовать из сходимости ряда

$$\sum_{j=0}^{\infty} P\{|\varepsilon_{i-j}|^{1/j} \geq 1 + \delta\}$$

при любом $\delta > 0$. Последний ряд равен ряду

$$\sum_{j=0}^{\infty} P\{\log^+ |\varepsilon_1| \geq j \log(1 + \delta)\},$$

который сходится в силу условия 5.2.3.

Единственность строго стационарного решения доказывается как в лемме 5.2.1. Лемма 5.2.2 доказана. \square

Существует огромное число содержательных задач из разных предметных областей (в частности, экономики, управления, гидрологии, геофизики, медицины, обработки звуковых сигналов и изображений и т.д.), приводящих к схемам авторегрессии и более общим схемам авторегрессии—скользящего среднего. Некоторые ссылки можно найти в обзорной статье [1]; выделим также работы [60, 96, 97, 110], где рассматриваются прикладные задачи, приводящие к авторегрессионным схемам с бесконечной дисперсией. Систематическому изложению собственно теории статистического анализа процессов авторегрессии—скользящего среднего также посвящено большое число монографий и статей. Здесь выделим лишь общеизвестные монографии [2, 45; 5, 57; 37, 74; 13, 83; 84]. Дальнейшие ссылки появятся по ходу изложения.

§ 5.3. Процедуры наименьших квадратов

Рассмотрим схему авторегрессии (5.2.1) при условии 5.2.1. Пусть u_0, u_1, \dots, u_n — выборка из строго стационарного решения (5.2.2). Такое решение существует в силу леммы 5.2.1.

В этом параграфе мы изучим оценку наименьших квадратов неизвестного параметра β , построенную по u_0, u_1, \dots, u_n , а также соответствующие тесты для проверки гипотез относительно β .

5.3.1. Оценка наименьших квадратов. *Оценкой наименьших квадратов* параметра β (будем обозначать ее $\hat{\beta}_{n,LS}$) называется решение задачи

$$\sum_{i=1}^n (u_i - \theta u_{i-1})^2 \Rightarrow \inf_{\theta \in \mathbb{R}^1}$$

или эквивалентного ей уравнения

$$\sum_{i=1}^n (u_i - \theta u_{i-1}) u_{i-1} = 0. \quad (5.3.1)$$

Поскольку

$$n^{-1} \sum_{i=1}^n u_{i-1}^2 \xrightarrow{P_\beta} E_\beta u_1^2 = E \varepsilon_1^2 (1 - \beta^2)^{-1} \neq 0$$

(см. далее соотношение (5.3.6)), то с вероятностью, стремящейся к единице, решение $\hat{\beta}_{n,LS}$ уравнения (5.3.1) существует, единственно и задается соотношением

$$\hat{\beta}_{n,LS} = \sum_{i=1}^n u_i u_{i-1} / \sum_{i=1}^n u_{i-1}^2. \quad (5.3.2)$$

Следующая теорема устанавливает асимптотическую нормальность $\hat{\beta}_{n,LS}$. Ее более общий вариант для векторной авторегрессии можно найти в [2, 45, п. 5.5.4].

Теорема 5.3.1. Пусть выполнено условие 5.2.1.

Тогда

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta) \xrightarrow{d_\beta} N(0, \sigma_{LS}^2(\beta)), \quad (5.3.3)$$

где $\sigma_{LS}^2(\beta) = 1 - \beta^2$.

Доказательство. Заменяя в (5.3.2) u_i на $\beta u_{i-1} + \varepsilon_i$ получим

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta) = n^{-1/2} \sum_{i=1}^n \varepsilon_i u_{i-1} / \left(n^{-1} \sum_{i=1}^n u_{i-1}^2 \right). \quad (5.3.4)$$

Из (5.2.1) следует

$$n^{-1} \sum_{i=1}^n u_i^2 = \beta^2 n^{-1} \sum_{i=1}^n u_{i-1}^2 + 2\beta n^{-1} \sum_{i=1}^n \varepsilon_i u_{i-1} + n^{-1} \sum_{i=1}^n \varepsilon_i^2. \quad (5.3.5)$$

Соотношение (5.2.2) влечет независимость величин u_i и $\{\varepsilon_j, j > i\}$ (этот факт систематически используется в дальнейшем). Поэтому

$$\mathbf{E}_\beta n^{-1} \sum_{i=1}^n \varepsilon_i u_{i-1} = 0, \quad \mathbf{E}_\beta (n^{-1} \sum_{i=1}^n \varepsilon_i u_{i-1})^2 = n^{-1} \mathbf{E} \varepsilon_1^2 \mathbf{E}_\beta u_0^2$$

и, значит,

$$n^{-1} \sum_{i=1}^n \varepsilon_i u_{i-1} \xrightarrow{P_\beta} 0.$$

Поскольку в силу закона больших чисел

$$n^{-1} \sum_{i=1}^n \varepsilon_i^2 \xrightarrow{P} \mathbf{E} \varepsilon_1^2,$$

то из (5.3.5) следует, что

$$n^{-1} \sum_{i=1}^n u_i^2 \xrightarrow{P_\beta} \mathbf{E} \varepsilon_1^2 (1 - \beta^2)^{-1} = \mathbf{E}_\beta u_1^2 \neq 0. \quad (5.3.6)$$

В силу (5.3.4) и (5.3.6) для доказательства теоремы достаточно показать, что

$$S_n := n^{-1/2} \sum_{i=1}^n \varepsilon_i u_{i-1} \xrightarrow{d_\beta} N(0, \mathbf{E} \varepsilon_1^2 \mathbf{E}_\beta u_0^2). \quad (5.3.7)$$

Введем величины

$$u_{ik} = \sum_{j=0}^k \beta^j \varepsilon_{i-j}, \quad i \in \mathbb{Z}.$$

Очевидно, при любом $k \geq 0$ последовательность $\{\varepsilon_i u_{i-1, k}, i \in \mathbb{Z}\}$ является строго стационарной $(k+1)$ -зависимой последовательностью, $\mathbf{E}_\beta \varepsilon_1 u_{0k} = 0$, $\mathbf{E}_\beta (\varepsilon_1 u_{0k})^2 = (\mathbf{E} \varepsilon_1^2)^2 \frac{1 - \beta^{2(k+1)}}{1 - \beta^2}$. В силу центральной предельной теоремы для k -зависимых последовательностей [2, 45, теорема 7.7.5]

$$Z_{kn} := n^{-1/2} \sum_{i=1}^n \varepsilon_i u_{i-1, k} \xrightarrow{d_\beta} Z_k, \quad (5.3.8)$$

где

$$Z_k \sim N\left(0, (\mathbf{E} \varepsilon_1^2)^2 \frac{1 - \beta^{2(k+1)}}{1 - \beta^2}\right).$$

Очевидно,

$$Z_k \xrightarrow{d_\beta} Z \sim N(0, (\mathbf{E} \varepsilon_1^2)^2 (1 - \beta^2)^{-1}), \quad k \rightarrow \infty.$$

Далее, если $X_{kn} := S_n - Z_{kn}$, то

$$\begin{aligned} \sup_n \mathbf{E}_\beta X_{kn}^2 &= \mathbf{E} \varepsilon_1^2 \mathbf{E}_\beta (u_0 - u_{0k})^2 = \\ &= \mathbf{E} \varepsilon_1^2 \mathbf{E}_\beta \left(\sum_{j=k+1}^{\infty} \beta^j \varepsilon_{-j} \right)^2 = (\mathbf{E} \varepsilon_1^2)^2 \frac{\beta^{2(k+1)}}{1 - \beta^2} \rightarrow 0, \quad k \rightarrow \infty. \end{aligned}$$

Далее нам потребуется следующее утверждение, являющееся перефразировкой теоремы 7.7.1 из [2, 45].

Л е м м а 5.3.1. Пусть $S_n = Z_{kn} + X_{kn}$, $n, k = 1, 2, \dots$. Предположим:

- 1) $\sup_n \mathbf{E} X_{kn}^2 \rightarrow 0, \quad k \rightarrow \infty;$
- 2) $Z_{kn} \xrightarrow{d} Z_k, \quad n \rightarrow \infty;$
- 3) $Z_k \xrightarrow{d} Z, \quad k \rightarrow \infty.$

Тогда $S_n \xrightarrow{d} Z, \quad n \rightarrow \infty.$

Возвращаясь к доказательству теоремы 5.3.1, заметим, что условия леммы 5.3.1 в нашем случае выполнены. В силу этой леммы (5.3.7) выполнено. Теорема 5.3.1 полностью доказана. \square

Замечательной особенностью оценки $\hat{\beta}_{n,LS}$ является то, что предельная дисперсия $\sigma_{LS}^2(\beta) = 1 - \beta^2$ свободна от $G(x)$.

Из (5.3.3) следует, что интервал

$$\left\{ \theta: \hat{\beta}_{n,LS} - n^{-1/2} (1 - \hat{\beta}_{n,LS}^2)^{1/2} \xi_{1-\alpha/2} < \theta < \hat{\beta}_{n,LS} + n^{-1/2} (1 - \hat{\beta}_{n,LS}^2)^{1/2} \xi_{1-\alpha/2} \right\} \quad (5.3.9)$$

является асимптотическим доверительным интервалом для β надежности $1 - \alpha$, если $\xi_{1-\alpha/2}$ — квантиль стандартного нормального закона уровня $1 - \alpha/2$.

Оценка $\hat{\beta}_{n,LS}$ является непараметрической. Интересно сравнить ее с параметрической асимптотически эффективной по Фишеру оценкой максимального правдоподобия $\hat{\beta}_{n,ML}$. Для построения последней предположим, что $G(x)$ имеет регулярную плотность вероятности $g(x)$ (см. [18, 59, § 9.1]). Обозначим $I(g)$ информацию Фишера $g(x)$ относительно сдвига. Пользуясь марковским свойством последовательности $\{u_i\}$, легко получить, что совместная плотность величин $u_0, u_1, \dots, u_n, \quad n \geq 1$, равна

$$g^n(x_0, x_1, \dots, x_n, \beta) = g_{u_0}(x_0, \beta) \prod_{i=1}^n g(x_i - \beta x_{i-1}),$$

где $g_u(x_0, \beta)$ — плотность u_0 . Для регулярных $g(x)$ состоятельное решение $\hat{\beta}_{n,ML}$ уравнения правдоподобия

$$\frac{\partial}{\partial \theta} \log g^n(u_0, u_1, \dots, u_n, \theta) = 0$$

представимо в виде

$$\sqrt{n}(\hat{\beta}_{n,ML} - \beta) = (I(g)E_{\beta}u_1^2)^{-1} n^{-1/2} \sum_{i=1}^n u_{i-1} \frac{g'(\varepsilon_i)}{g(\varepsilon_i)} + o_p(1),$$

откуда

$$\sqrt{n}(\hat{\beta}_{n,ML} - \beta) \xrightarrow{d} N(0, \sigma_{ML}^2(\beta)), \quad (5.3.10)$$

где $\sigma_{ML}^2(\beta) = (1 - \beta^2)(I(g)E_{\beta}u_1^2)^{-1}$ (см., например, [18, 59, п. 9.2.3, пример 9.11]). Из (5.3.10) и (5.3.3) следует, что асимптотическая относительная эффективность (АОЭ) оценки $\hat{\beta}_{n,LS}$ относительно $\hat{\beta}_{n,ML}$ есть

$$e_{LS,ML} := \frac{\sigma_{ML}^2(\beta)}{\sigma_{LS}^2(\beta)} = (I(g)E_{\beta}u_1^2)^{-1}.$$

Эта эффективность инвариантна относительно β и параметра масштаба $G(x)$. При $G(x) \sim N(0, 1)$ $e_{LS,ML}$ равна единице и $\hat{\beta}_{n,LS}$ асимптотически эффективна по Фишеру.

5.3.2. Тесты наименьших квадратов. Соотношение (5.3.3) можно использовать для проверки гипотезы $H_0: \beta = \beta_0$, беря в качестве тестовой статистики

$$T_{n,LS}(\beta_0) = \sqrt{n} \sigma_{LS}^{-1}(\beta_0) (\hat{\beta}_{n,LS} - \beta_0).$$

При H_0 и условии 5.2.1 эта статистика асимптотически является стандартной гауссовской. При фиксированной альтернативе $H_1: \beta \neq \beta_0$ в силу теоремы 5.3.1 $T_{n,LS}(\beta_0)$ расходитя к бесконечности и соответствующий тест состоятелен против H_1 .

Интересен вопрос об асимптотической мощности теста со статистикой $T_{n,LS}(\beta_0)$ при альтернативах

$$H_{1n}: \beta = \beta_n, \quad \beta_n \rightarrow \beta_0,$$

сближающихся с гипотезой H_0 .

При изучении поведения тестовой статистики в случае меняющегося вместе с n параметра авторегрессионной схемы полезна следующая

Теорема 5.3.2. Пусть выполнено условие 5.2.1.
Тогда при альтернативе H_{1n}

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta_n) \xrightarrow{d_{\beta_n}} N(0, \sigma_{LS}^2(\beta_0)), \quad (5.3.11)$$

где $\sigma_{LS}^2(\beta) = 1 - \beta^2$.

Прежде чем доказывать теорему 5.3.2, обсудим несколько ее следствий.

Следствие 5.3.1. Если последовательность $\sqrt{n}(\beta_n - \beta_0)$ расходится к бесконечности, то это же верно для $T_{n,LS}(\beta_0)$ и, следовательно, соответствующий тест состоятелен против H_{1n} .

Следствие 5.3.2. Если $\sqrt{n}(\beta_n - \beta_0) = a + o(1)$ для некоторой постоянной a , то

$$T_{n,LS}(\beta_0) \xrightarrow{d_{\beta_n}} N(a(1 - \beta_0^2)^{-1/2}, 1). \quad (5.3.12)$$

Будем в этом случае обозначать H_{1n} через $H_{1n}(a)$. Итак, при $H_{1n}(a)$ в случае $a \neq 0$ асимптотическая мощность нетривиальна (т.е. больше уровня значимости и меньше единицы). В частности, если H_0 проверяется на уровне значимости α против альтернативы

$$H_{1n}^+(a): \beta = \beta_n := \beta_0 + a n^{-1/2} + o(n^{-1/2}), \quad a > 0,$$

то для мощности теста справедливо

$$P_{\beta_n}\{T_{n,LS}(\beta_0) > \xi_{1-\alpha}\} \rightarrow \Phi(a(1 - \beta_0^2)^{-1/2} - \xi_{1-\alpha}),$$

где $\Phi(x)$ — стандартная гауссовская функция распределения.

Если $a = 0$, то асимптотические распределения тестовой статистики при H_0 и $H_{1n}(0)$ совпадают, и гипотезы H_0 и $H_{1n}(0)$ асимптотически неразличимы.

Доказательство теоремы 5.3.2. При $\beta = \beta_n$ аналогично (5.3.4) получаем

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta_n) = n^{-1/2} \sum_{i=1}^n \varepsilon_i u_{i-1} / \left(n^{-1} \sum_{i=1}^n u_{i-1}^2 \right). \quad (5.3.13)$$

Из (5.2.2) следует, что

$$E_{\beta_n} u_1^2 = E \varepsilon_1^2 (1 - \beta_n^2)^{-1} \rightarrow E \varepsilon_1^2 (1 - \beta_0^2)^{-1},$$

так что $n^{-1} E_{\beta_n} u_1^2 = o(1)$. Это замечание совместно с соотношением (5.3.5), в котором следует заменить β на β_n , влечет, как при доказа-

тельстве теоремы 5.3.1

$$n^{-1} \sum_{i=1}^n u_i^2 \xrightarrow{P_{\beta_n}} \mathbf{E} \varepsilon_1^2 (1 - \beta_0^2)^{-1}. \quad (5.3.14)$$

В силу (5.3.13) и (5.3.14) для доказательства (5.3.11) достаточно показать

$$S_n := n^{-1/2} \sum_{i=1}^n \varepsilon_i u_{i-1} \xrightarrow{d_{P_n}} N(0, (\mathbf{E} \varepsilon_1^2)^2 (1 - \beta_0^2)^{-1}). \quad (5.3.15)$$

Положим $u_i^0 = \sum_{j=0}^{\infty} \beta_0^j \varepsilon_{i-j}$. Тогда в силу (5.3.7)

$$S_{n0} := n^{-1/2} \sum_{i=1}^n \varepsilon_i u_{i-1}^0 \xrightarrow{d_{P_n}} N(0, (\mathbf{E} \varepsilon_1^2)^2 (1 - \beta_0^2)^{-1}). \quad (5.3.16)$$

Покажем теперь, что

$$S_n - S_{n0} = o_p(1). \quad (5.3.17)$$

Для этого достаточно показать, что

$$\mathbf{E}_{\beta_n} (S_n - S_{n0})^2 = \mathbf{E} \varepsilon_1^2 \mathbf{E}_{\beta_n} (u_0 - u_0^0)^2 \rightarrow 0.$$

Действительно, при $\max(|\beta_0|, |\beta_n|) \leq b < 1$

$$\begin{aligned} \mathbf{E}_{\beta_n} (u_0 - u_0^0)^2 &= \mathbf{E} \varepsilon_1^2 \sum_{j=1}^{\infty} (\beta_n^j - \beta_0^j)^2 = \\ &= \mathbf{E} \varepsilon_1^2 (\beta_n - \beta_0)^2 \sum_{j=1}^{\infty} \left(\sum_{s=0}^{j-1} \beta_n^{j-s-1} \beta_0^s \right)^2 \leq \\ &\leq \mathbf{E} \varepsilon_1^2 (\beta_n - \beta_0)^2 \sum_{j=1}^{\infty} j^2 b^{2(j-1)} = o(1). \end{aligned}$$

Теперь из (5.3.16)–(5.3.17) следует (5.3.15) и утверждение теоремы. \square

Уместно отметить, что справедливость (5.3.12) при $H_{1n}(a)$ можно установить также с помощью общих результатов для контигуальных гипотез, но при несколько более жестких, чем $\mathbf{E} \varepsilon_1 = 0$, $0 < \mathbf{E} \varepsilon_1^2 < \infty$, условиях. А именно, пусть дополнительно $G(x)$ имеет абсолютно непрерывную плотность $g(x) > 0$ для всех x и конечную ненулевую информацию Фишера $I(g)$. Тогда (см. [91]) при $\beta_n =$

$$= \beta_0 + an^{-1/2} + o(n^{-1/2})$$

$$\log \frac{d\mathbf{P}_{n,\beta_n}}{d\mathbf{P}_{n,\beta_0}} = a \Delta_n(\beta_0) - \frac{1}{2} a^2 I(g) \mathbf{E}_{\beta_0} u_1^2 + \delta_n$$

и

$$\Delta_n(\beta_0) \xrightarrow{d_{\beta_0}} N(0, I(g) \mathbf{E}_{\beta_0} u_1^2), \quad \delta_n \xrightarrow{P_{\beta_0}} 0.$$

Здесь $\mathbf{P}_{n,\beta}$ — распределение вектора наблюдений (u_0, u_1, \dots, u_n) ,

$$\Delta_n(\theta) = n^{-1/2} \sum_{i=1}^n \frac{\partial}{\partial \theta} \log g(u_i - \theta u_{i-1}). \quad (5.3.18)$$

Таким образом, наблюдения u_0, u_1, \dots, u_n обладают свойством локальной асимптотической нормальности (ЛАН) (определение см., например, в [15, гл. 2, § 2]). Это свойство влечет контигуальность H_0 и $H_{1n}(a)$ [11, 67, гл. 6, § 1]. С помощью третьей леммы Ле Кама [11, 67, гл. 6, § 1] и разложения

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0) = \frac{1 - \beta_0}{\mathbf{E}\varepsilon_1^2} n^{-1/2} \sum_{i=1}^n u_{i-1}\varepsilon_i + o_p(1),$$

справедливого при гипотезе H_0 в силу доказательства теоремы 5.3.1, получаем: при альтернативе $H_{1n}(a)$ (и перечисленных условиях регулярности) $\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0)$ — асимптотически гауссовская величина с предельной дисперсией $\sigma_{LS}^2(\beta_0)$ и асимптотическим сдвигом, равным ковариации предельного гауссовского распределения вектора $(\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0), \log \frac{d\mathbf{P}_{n,\beta_n}}{d\mathbf{P}_{n,\beta_0}})$ при H_0 , которая и равна a , т.е.

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0) \xrightarrow{d_{\beta_0}} N(a, \sigma_{LS}^2(\beta_0)).$$

Последнее соотношение эквивалентно (5.3.12).

Мы используем (5.3.12) далее для вычисления АОЭ теста, основанного на $\hat{\beta}_{n,LS}$, относительно других тестов.

В качестве примера рассмотрим параметрический асимптотически оптимальный тест со статистикой

$$T_{n,ML}(\beta_0) := (I(g) \mathbf{E}_{\beta_0} u_1^2)^{-1/2} \Delta_n(\beta_0),$$

в которой $\Delta_n(\theta)$ определена соотношением (5.3.18). Такой тест асимптотически эквивалентен тесту отношения правдоподобия. При альтернативе $H_{1n}(a)$ и условиях, обеспечивающих свойство ЛАН,

$$T_{n,ML}(\beta_0) \xrightarrow{d_{\beta_0}} N(a (I(g) \mathbf{E}_{\beta_0} u_1^2)^{1/2}, 1) \quad (5.3.19)$$

в силу третьей леммы Ле Кама. При гипотезе H_0 статистика $T_{n,ML}(\beta_0)$ имеет асимптотически стандартное нормальное распределение. Из (5.3.19) и (5.3.12) следует, что АОЭ теста, основанного на $\hat{\beta}_{n,LS}$, относительно теста, основанного на $\hat{\beta}_{n,ML}$, равная $(\delta_{LS}(\beta_0)/\delta_{ML}(\beta_0))^2$, есть $e_{LS,ML}$. Здесь $\delta_{LS}(\beta_0)$ и $\delta_{ML}(\beta_0)$ — асимптотические сдвиги (средние) $T_{n,LS}(\beta_0)$ и $T_{n,ML}(\beta_0)$ при $H_{1n}(a)$. Таким образом, АОЭ для тестов совпадает с АОЭ для оценок.

Подчеркнем, что все перечисленные для $\hat{\beta}_{n,LS}$ свойства асимптотические. Об оптимальных процедурах, основанных на сериальных корреляциях, и точных распределениях в гауссовских моделях, являющихся модификациями разностного уравнения авторегрессии (не обязательно первого порядка), см., например, [2, 45, гл. 6; 37, 74, § 3] и имеющиеся там ссылки. Мы не будем говорить об этом подробно. Одна из причин этому в том, что постулировать гауссовость в реальных прикладных задачах часто не очень естественно. Лучше строить процедуры, свободные от $G(x)$ при конечных n для широкого класса распределений. Процедуры такого типа знаковые и представлены в следующих двух главах.

Выделим теперь особо значимость предположения $E\varepsilon_1^2 < \infty$ для асимптотической нормальности $\sqrt{n}(\hat{\beta}_{n,LS} - \beta)$. Известно (см. [111]), что если $G(x)$ симметрична относительно нуля и выполнено условие 5.2.3, то $\sqrt{n}(\hat{\beta}_{n,LS} - \beta) = O_p(1)$, однако предельное распределение $\sqrt{n}(\hat{\beta}_{n,LS} - \beta)$ вовсе не обязано быть нормальным. Например, как отмечается в [111], если $G(x)$ симметрична относительно нуля и $\lim_{x \rightarrow +\infty} x^\alpha(1 - G(x)) = k > 0$ для некоторого $\alpha \in (0, 2)$ (в этом случае не существует конечного $E|\varepsilon_1|^\alpha$), то $n^\delta(\hat{\beta}_{n,LS} - \beta) = o_p(1)$ при любом $\delta < 1/\alpha$.

Аномально высокая скорость сходимости $\hat{\beta}_{n,LS}$ к β при $E\varepsilon_1^2 = \infty$ объясняется тем, что в этом случае не выполняются достаточные условия ЛАН и нет стандартной асимптотической границы снизу для рисков регулярных оценок (см., например, [16, гл. 2, § 11]).

Уместно здесь же отметить, что все сказанное выше относительно процедур наименьших квадратов в однопараметрической схеме (5.2.1), остается справедливым и в многопараметрической схеме. Разница между однопараметрической и многопараметрической схемами лишь в технических подробностях. Детальные описания можно найти, например, в [2, 45, гл. 5], где, в частности, рассмотрен случай ненулевого среднего у величин $\{\varepsilon_i\}$. Мы вернемся к оценке наименьших квадратов в многопараметрической авторегрессии в гл. 7.

§ 5.4. Оценка наименьших квадратов в нестационарной авторегрессии

Рассмотрим нестационарное соотношение авторегрессии

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i = 1, 2, \dots, \quad u_0 = 0. \quad (5.4.1)$$

В (5.4.1) $\{\varepsilon_i\}$ — последовательность независимых стандартных гаус-

совских величин, $\beta \in \mathbb{R}^1$. Из (5.4.1) следует, что

$$u_i = \sum_{j=0}^{i-1} \beta^j \varepsilon_{i-j}, \quad i = 1, 2, \dots \quad (5.4.2)$$

В силу (5.4.2) дисперсия u_i равна $1 + \beta^2 + \dots + \beta^{i-1}$ и, следовательно, зависит от i при $\beta \neq 0$. Таким образом, последовательность $\{u_i\}$ нестационарна при $\beta \neq 0$.

Пусть u_1, \dots, u_n — наблюдения, порожденные (5.4.1). Мы не будем рассматривать нестационарную авторегрессию систематически. Цель параграфа более скромная — сформулировать и доказать один интересный результат об асимптотическом поведении оценки наименьших квадратов $\hat{\beta}_{n,LS}$ из (5.3.2) (которая, кстати, является в рассматриваемом случае оценкой максимального правдоподобия). Хотя рассматривается параметрическая ситуация (распределение последовательности u_1, \dots, u_n является гауссовским с единственным неизвестным параметром β), нам будет полезен и сам результат, и его доказательство. А именно, далее в § 6.3 мы сформулируем результаты о точном и асимптотическом распределении знаковых тестовых статистик в схеме (5.4.1) в непараметрической ситуации, когда функция распределения величин ε_i неизвестна. Будет полезно сравнить эти результаты с результатами настоящего параграфа.

Итак, обозначим $J_n(\beta)$ информацию Фишера о параметре β , содержащуюся в u_1, \dots, u_n . Легко убедиться, что

$$J_n(\beta) = \begin{cases} \frac{1}{1-\beta^2} \left(n - \frac{1-\beta^{2n}}{1-\beta^2} \right), & |\beta| \neq 1, \\ \frac{1}{2} n(n-1), & |\beta| = 1. \end{cases} \quad (5.4.3)$$

Действительно, совместная плотность величин u_1, \dots, u_n равна

$$g(x_1, \dots, x_n, \beta) = (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (x_i - \beta x_{i-1})^2 \right\}, \quad x_0 = 0.$$

Следовательно,

$$J_n(\beta) = \mathbf{E}_\beta \left(-\frac{\partial^2 \log g(u_1, \dots, u_n, \beta)}{\partial \beta^2} \right) = \sum_{i=1}^n \mathbf{E}_\beta u_i^2.$$

Отсюда и из равенства

$$\mathbf{E}_\beta u_i^2 = \mathbf{E}_\beta \left(\sum_{j=0}^{i-1} \beta^j \varepsilon_{i-j} \right)^2 = \sum_{j=0}^{i-1} \beta^{2j} = \begin{cases} \frac{1-\beta^{2i}}{1-\beta^2}, & |\beta| \neq 1, \\ i, & |\beta| = 1, \end{cases}$$

следует (5.4.3). Из (5.4.3) в свою очередь следует, что

$$J_n(\beta) \sim d_n^2(\beta) := \begin{cases} n/(1-\beta^2), & |\beta| < 1, \\ n^2/2, & |\beta| = 1, \\ \beta^{2n}/(1-\beta^2)^2, & |\beta| > 1. \end{cases} \quad (5.4.4)$$

Обозначим $K(0, 1)$ распределение Коши с параметрами сдвига и масштаба 0 и 1, пусть $H(\beta)$ означает распределение величины

$$\beta \frac{\nu^2(1) - 1}{2^{2/3} \int_0^1 \nu^2(t) dt}, \quad (5.4.5)$$

$\nu(t)$, $t \in [0, 1]$, — стандартный винеровский процесс. Далее будет особо выделен случай $|\beta| = 1$, поэтому стоит отметить, что при $\beta = \pm 1$ плотность величины (5.4.5) (обозначим ее $h_\beta(x)$, а соответствующую функцию распределения $H_\beta(x)$) не является четной. В частности, (см. [64])

$$\begin{aligned} h_1(-0.2) &\approx 0.316, & H_1(-0.2) &\approx 0.617, \\ h_1(0) &\approx 0.341, & H_1(0) &\approx 0.683, \\ h_1(0.2) &\approx 0.357, & H_1(0.2) &\approx 0.753. \end{aligned}$$

Следующая теорема и ее доказательство взяты из [106, § 1 гл. 5].

Теорема 5.4.1. Пусть в схеме (5.4.1) $\{\varepsilon_i, i = 1, 2, \dots\}$ — независимые гауссовские стандартные величины.

Тогда

$$d_n(\beta)(\hat{\beta}_{n,LS} - \beta) \xrightarrow{d_\beta} \begin{cases} N(0, 1), & |\beta| < 1, \\ H(\beta), & |\beta| = 1, \\ K(0, 1), & |\beta| > 1. \end{cases}$$

Доказательство. Пусть

$$U_n := d_n^{-1}(\beta) \sum_{i=1}^n \varepsilon_i u_{i-1}, \quad V_n := d_n^{-2}(\beta) \sum_{i=1}^n u_{i-1}^2.$$

Тогда

$$d_n(\beta)(\hat{\beta}_{n,LS} - \beta) = U_n V_n^{-1} \quad (5.4.6)$$

в силу (5.3.4). Обозначим $f_n(s, t)$ совместную характеристическую

функцию U_n, V_n . Известно (см. [104]), что

$$f(s, t) := \lim_{n \rightarrow \infty} f_n(s, t) = \begin{cases} \exp\{it - s^2/2\}, & |\beta| < 1, \\ \exp\left\{-\frac{is\beta}{\sqrt{2}}\right\} \left[\cos(2\sqrt{it}) - \frac{\beta it}{\sqrt{2is}} \sin(2\sqrt{is})\right]^{-1/2}, & |\beta| = 1, \\ (1 + s^2 - 2it)^{-1/2}, & |\beta| > 1. \end{cases}$$

В случае $|\beta| < 1$ $f(s, t)$ есть характеристическая функция пары $(\xi, 1)$, $\xi \sim N(0, 1)$. Следовательно, $(U_n, V_n) \xrightarrow{d_p} (\xi, 1)$, что влечет в силу (5.4.6)

$$d_n(\beta)(\hat{\beta}_{n,LS} - \beta) = U_n V_n^{-1} \xrightarrow{d_p} \xi \sim N(0, 1).$$

При $|\beta| > 1$ $f(s, t)$ является характеристической функцией пары $(\xi\eta, \eta^2)$, где ξ и η есть независимые стандартные гауссовские величины. Поэтому $(U_n, V_n) \xrightarrow{d_p} (\xi\eta, \eta^2)$ и $U_n V_n^{-1} \xrightarrow{d_p} \xi/\eta \sim K(0, 1)$. Следовательно, $d_n(\beta)(\hat{\beta}_{n,LS} - \beta) \xrightarrow{d_p} K(0, 1)$.

Пусть теперь $\beta = 1$, случай $\beta = -1$ рассматривается аналогично. Тогда в силу (5.4.2)

$$U_n = \frac{\sqrt{2}}{n} \sum_{i=1}^n \varepsilon_i u_{i-1} = \frac{\sqrt{2}}{n} \sum_{i=2}^n \sum_{j=1}^{i-1} \varepsilon_j \varepsilon_i, \quad (5.4.7)$$

$$V_n = \frac{2}{n^2} \sum_{i=1}^n u_{i-1}^2 = \frac{2}{n^2} \sum_{i=2}^n \left(\sum_{j=1}^{i-1} \varepsilon_j\right)^2. \quad (5.4.8)$$

Введем случайный процесс $\nu_n(t) = n^{-1/2} \sum_{i \leq nt} \varepsilon_i$, $t \in [0, 1]$, пусть

$\Delta \nu_n\left(\frac{i}{n}\right) := \nu_n\left(\frac{i}{n}\right) - \nu_n\left(\frac{i-1}{n}\right)$. Тогда из (5.4.7)–(5.4.8) следует

$$U_n = \sqrt{2} \sum_{i=1}^n \nu_n\left(\frac{i-1}{n}\right) \Delta \nu_n\left(\frac{i}{n}\right),$$

$$V_n = 2 \sum_{i=1}^n \nu_n^2\left(\frac{i-1}{n}\right) \frac{1}{n}.$$

В силу принципа инвариантности (см., например, [4, 49, гл. 5]) $\nu_n(t)$ слабо сходится к $\nu(t)$ в пространстве Скорохода $\mathbb{D}[0, 1]$. Поэтому в

силу общих результатов о слабой сходимости (см., например, [82]),

$$(U_n, V_n) \xrightarrow{d_1} \left(\sqrt{2} \int_0^1 \nu(t) d\nu(t), 2 \int_0^1 \nu^2(t) dt \right).$$

Следовательно,

$$d_n(1)(\hat{\beta}_{n,LS} - 1) = U_n V_n^{-1} \xrightarrow{d_1} \frac{\int_0^1 \nu(t) d\nu(t)}{\sqrt{2} \int_0^1 \nu^2(t) dt} = \frac{\nu^2(1) - 1}{2\sqrt{2} \int_0^1 \nu^2(t) dt},$$

поскольку $\int_0^1 \nu(t) d\nu(t) = 1/2(\nu^2(1) - 1)$. Теорема полностью доказана. \square

Следует отметить, что в случае $|\beta| \leq 1$ утверждение теоремы 5.4.1 остается справедливым для схемы (5.4.1), в которой u_0 — любая случайная величина, не зависящая от $\{\varepsilon_i, i = 1, 2, \dots\}$ и имеющая второй момент, а $\{\varepsilon_i\}$ — н.о.р. случайные величины с нулевым средним и конечной дисперсией. При $|\beta| < 1$ этот факт устанавливается вполне аналогично теореме 5.3.1, а при $|\beta| = 1$ аналогично теореме 5.4.1. Если же $|\beta| > 1$, то предельное распределение $d_n(\beta)(\hat{\beta}_{n,LS} - \beta)$ зависит от распределения величин ε_i (см. [89]).

В силу теоремы 5.4.1 статистику $d_n(\beta_0)(\hat{\beta}_{n,LS} - \beta_0)$ при любом $\beta_0 \in \mathbb{R}^1$ можно использовать для проверки гипотезы $H_0: \beta = \beta_0$ в качестве тестовой. Ее распределение при близких альтернативах и $|\beta| < 1$ такое же, как в § 5.3. При $|\beta| \geq 1$ его также можно найти, но мы не будем этим здесь заниматься.

§ 5.5. Процедуры наименьших модулей

Продолжим рассмотрение стационарной схемы (5.2.1). В этом параграфе мы изучим оценки наименьших модулей для β и соответствующие тесты.

5.5.1. Оценка наименьших модулей. Будем считать выполненными

Условие 5.5.1. $E\varepsilon_1 = 0, E|\varepsilon_1|^{2+\Delta} < \infty$ при некотором $\Delta > 0$.

Условие 5.5.2. Существует $G'(x) = g(x), \sup_x g(x) < \infty, g(0) > 0, g(x)$ удовлетворяет условию Липшица в нуле.

При условии 5.5.1 в силу леммы 5.2.1 существует стационарное решение уравнения (5.2.1) вида (5.2.2), и ряд в (5.2.2) сходится в $L^{2+\Delta}$.

Пусть u_0, u_1, \dots, u_n — наблюдения, удовлетворяющие (5.2.2). Поставим задачей исследовать асимптотическое поведение оценки наименьших модулей $\hat{\beta}_{n,LD}$ параметра β , получаемой решением задачи

$$L_n^{LD}(\theta) := \sum_{k=1}^n |u_k - \theta u_{k-1}| \Rightarrow \inf_{\theta \in \mathbb{R}^1} \cdot \quad (5.5.1)$$

Свойства целевой функции $L_n^{LD}(\theta)$ очевидны: это выпуклая вниз функция параметра θ ; на каждом из отрезков $[z_{(k-1)}, z_{(k)}]$ (где $z_k := u_k/u_{k-1}$, а $z_{(k)}$ — k -й по величине член последовательности z_1, \dots, z_n) она линейна.

В любой точке θ , отличной от точек z_k , $k = 1, \dots, n$, производная функции $L_n^{LD}(\theta)$ равна

$$l_n^{LD}(\theta) = \sum_{k=1}^n u_{k-1} \text{sign}(u_k - \theta u_{k-1}). \quad (5.5.2)$$

Так как $G(x)$ непрерывна, то на любом из интервалов $(z_{(k-1)}, z_{(k)})$ $l_n^{LD}(\theta) \neq 0$ с вероятностью единица. Следовательно, решение задачи (5.5.1) п.н. единственно и совпадает с одной из точек z_k . Оно может быть найдено перебором. Его также можно указать явно. А

именно, поместим в каждую точку z_k вероятность $p_k = |u_k| / \sum_{i=1}^n |u_i|$.

Тогда решение задачи (5.5.1) совпадает с медианой распределения $(z_k, p_k, k = 1, \dots, n)$.

Перепишывая $l_n^{LD}(\theta)$ в виде

$$l_n^{LD}(\theta) = \sum_{k=1}^n |u_{k-1}| \text{sign} \left(\frac{u_k}{u_{k-1}} - \theta \right)$$

видим, что траектории $l_n^{LD}(\theta)$ — невозрастающие кусочно постоянные функции со скачками в точках z_k . Для непрерывной $G(x)$ решение $\hat{\beta}_{n,LD}$ задачи (5.5.1) совпадает с п.н. единственным решением уравнения

$$l_n^{LD}(\theta) \div 0 \quad (5.5.3)$$

(символ \div в (5.5.3) означает переход через нуль), которое и будем дальше изучать. При условиях 5.5.1, 5.5.2

$$\mathbf{E}_\beta n^{-1} l_n^{LD}(\theta) = \Lambda_{LD}(\theta),$$

где $\Lambda_{LD}(\theta) = -2\mathbf{E}_\beta u_1 G((\theta - \beta)u_1)$, и существует производная по θ , равная при $\theta = \beta$

$$\lambda_{LD}(\beta) := \Lambda'_{LD}(\beta) = -2g(0)\mathbf{E}_\beta u_1^2 = -2g(0)\mathbf{E}_\beta \varepsilon_1^2 (1 - \beta^2)^{-1}.$$

Действительно, в силу независимости ε_2 и u_1

$$\begin{aligned} \Lambda_{LD}(\theta) &= \mathbf{E}_\beta u_1 \operatorname{sign}(u_2 - \theta u_1) = \mathbf{E}_\beta u_1 \{1 - 2I(\varepsilon_2 < (\theta - \beta)u_1)\} = \\ &= -2\mathbf{E}_\beta \mathbf{E}_\beta \{u_1 I(\varepsilon_2 < (\theta - \beta)u_1)/u_1\} = -2\mathbf{E}_\beta u_1 G((\theta - \beta)u_1). \end{aligned}$$

Далее, в силу формулы Тейлора

$$\begin{aligned} \Lambda_{LD}(\theta) &= -2\mathbf{E}_\beta u_1 (G(0) + g(\theta_1)u_1(\theta - \beta)) = \\ &= -2\mathbf{E}_\beta u_1^2 g(\theta_1)(\theta - \beta), \quad |\theta_1| \leq |(\theta - \beta)u_1|. \end{aligned}$$

Функция $g(x)$ ограничена и удовлетворяет условию Липшица в нуле, а потому удовлетворяет и условию Гёльдера любого порядка $0 < \delta \leq 1$. Возьмем $\delta = \min(\Delta, 1)$. Тогда

$$\Lambda_{LD}(\theta) = -2g(0)\mathbf{E}_\beta u_1^2(\theta - \beta) + O(|\theta - \beta|^{1+\delta}),$$

что влечет выражение для $\Lambda'_{LD}(\beta)$.

Следующая теорема 5.5.1 описывает линейное стохастическое разложение процесса $n^{-1/2}I_n^{LD}(\theta)$ в окрестности точки $\theta = \beta$ размера $O(n^{-1/2})$, равномерное по θ .

Т е о р е м а 5.5.1. Пусть выполнены условия 5.5.1, 5.5.2. Тогда при любом $0 < \Theta < \infty$

$$\sup_{|\theta| \leq \Theta} |n^{-1/2}I_n^{LD}(\beta + n^{-1/2}\theta) - n^{-1/2}I_n^{LD}(\beta) - \lambda_{LD}(\beta)\theta| = o_p(1).$$

Мы вынесли доказательство теоремы 5.5.1 в приложение, § 5.8.

Из теоремы 5.5.1 следует теорема 5.5.2, доказательство которой мы поместили вслед за доказательством теоремы 5.5.1 также в § 5.8.

Т е о р е м а 5.5.2. Пусть выполнены условия 5.5.1, 5.5.2. Тогда:

- 1) $\sqrt{n}(\hat{\beta}_{n,LD} - \beta) = -\lambda_{LD}^{-1}(\beta)n^{-1/2}I_n^{LD}(\beta) + o_p(1)$;
- 2) $\sqrt{n}(\hat{\beta}_{n,LD} - \beta) \xrightarrow{d} N(0, \sigma_{LD}^2(\beta))$, где

$$\sigma_{LD}^2(\beta) = (1 - \beta^2) \left((2g(0))^2 \mathbf{E} \varepsilon_1^2 \right)^{-1}.$$

Отметим, что утверждение теоремы 5.5.2 останется справедливым, если в условии 5.5.1 заменить требование $\mathbf{E} \varepsilon_1 = 0$ на $G(0) = 1/2$. В этом случае $\sigma_{LD}^2(\beta) = \left((2g(0))^2 \mathbf{E}_\beta u_1^2 \right)^{-1}$.

По поводу теоремы 5.5.2 уместно отметить, что строгое обоснование асимптотической нормальности оценки наименьших модулей в авторегрессионной схеме дано относительно недавно. В частности,

в [102] асимптотическая нормальность оценки наименьших модулей установлена единым методом сразу для нескольких схем наблюдений: линейной регрессии с независимыми погрешностями, линейной регрессии со случайными регрессорами, авторегрессии произвольного порядка. Более того, в случае авторегрессии асимптотическая нормальность $\hat{\beta}_{n,LD}$ установлена даже при более слабых условиях, чем условия 5.5.1, 5.5.2, в частности при моментных ограничениях $E\varepsilon_1 = 0$, $0 < E\varepsilon_1^2 < \infty$. Рассуждение в [102] основано на выпуклости целевой функции, минимизацией которой получается оценка наименьших модулей (в однопараметрической авторегрессии это $L_n^{LD}(\theta)$): устанавливается поточечное разложение этой функции, которое в силу выпуклости влечет равномерное разложение, а с ним и утверждение об асимптотической нормальности. Наш подход иной, мы рассматриваем оценку $\hat{\beta}_{n,LD}$ как корень уравнения (5.5.3) и анализируем производную целевой функции $l_n^{LD}(\theta)$, которая никакой выпуклостью не обладает. Тем не менее, удается установить равномерное стохастическое разложение для $l_n^{LD}(\theta)$, а отсюда, как простое следствие, получить теорему 5.5.2. В гл. 6 и 7 мы увидим, что наш подход применим к исследованию и таких оценок (знаковых, ранговых), с которыми не удастся связать никакой выпуклой задачи минимизации, их определяющей. Более того, в § 5.7 методом, аналогичным используемому в этом параграфе, мы исследуем и нестационарную схему авторегрессии. Наконец, в § 5.6, пользуясь определением оценки $\hat{\beta}_{n,LD}$ как корня уравнения (5.5.3), найдем функционал влияния для $\hat{\beta}_{n,LD}$.

Завершим комментарий к теореме 5.5.2 замечанием о поведении оценки $\hat{\beta}_{n,LD}$ в случае бесконечной дисперсии $E\varepsilon_1^2 = \infty$. Результаты здесь вполне аналогичны приведенным в п. 5.3.2 для оценки наименьших квадратов. Например [102], если ε_i имеют распределение Коши, то

$$n(\hat{\beta}_{n,LD} - \beta) = O_p(1).$$

Другие результаты и ссылки можно найти в [51, § 3.2].

Вернемся теперь к рассмотрению схемы (5.2.1) при условиях 5.5.1, 5.5.2. Незвестную предельную дисперсию $\sigma_{LD}^2(\beta)$ легко оценить. Для этого полезно

С л е д с т в и е 5.5.1. Если $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$, то при условиях 5.5.1, 5.5.2

$$n^{-1/2}l_n^{LD}(\hat{\beta}_n) = n^{-1/2}l_n^{LD}(\beta) + \lambda_{LD}(\beta)\sqrt{n}(\hat{\beta}_n - \beta) + o_p(1). \quad (5.5.4)$$

Доказательство. $\forall \varepsilon > 0$ и $0 < \Theta < \infty$

$$\begin{aligned} P_\beta \{ |n^{-1/2}l_n^{LD}(\hat{\beta}_n) - n^{-1/2}l_n^{LD}(\beta) - \lambda_{LD}(\beta)\sqrt{n}(\hat{\beta}_n - \beta)| > \varepsilon \} &\leq \\ &\leq P_\beta \{ |n^{-1/2}l_n^{LD}(\hat{\beta}_n) - n^{-1/2}l_n^{LD}(\beta) - \lambda_{LD}(\beta)\sqrt{n}(\hat{\beta}_n - \beta)| > \varepsilon, \\ &|\sqrt{n}(\hat{\beta}_n - \beta)| \leq \Theta \} + P_\beta \{ |\sqrt{n}(\hat{\beta}_n - \beta)| > \Theta \} \leq \end{aligned}$$

$$\leq P_{\beta} \left\{ \sup_{|\theta| \leq \Theta} |n^{-1/2} l_n^{LD}(\beta + n^{-1/2} \theta) - n^{-1/2} l_n^{LD}(\beta) - \lambda_{LD}(\beta) \theta| > \varepsilon \right\} + P_{\beta} \left\{ |\sqrt{n}(\hat{\beta}_n - \beta)| > \Theta \right\}.$$

В силу теоремы 5.5.1 и условия $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$ оба последних слагаемых сколь угодно малы при достаточно больших n и Θ . \square

Полагая в (5.5.4) $\hat{\beta}_n = \hat{\beta}_{n,LD}$, получаем для любой постоянной $h \neq 0$

$$\hat{\lambda}_n = (l_n^{LD}(\hat{\beta}_{n,LD} + hn^{-1/2}) - l_n^{LD}(\hat{\beta}_{n,LD}))(\sqrt{nh})^{-1} = \lambda_{LD}(\beta) + o_p(1).$$

Поскольку $\sigma_{LD}^2(\beta) = E_{\beta} u_1^2 / \lambda_{LD}^2(\beta)$, то состоятельной оценкой $\sigma_{LD}^2(\beta)$ будет

$$\hat{\sigma}_{n,LD}^2 = s_n^2 / \hat{\lambda}_n^2,$$

где $s_n^2 = n^{-1} \sum_{k=1}^n u_k^2$ — состоятельная оценка $E_{\beta} u_1^2$. Следовательно, множество

$$\left\{ \theta: \hat{\beta}_{n,LD} - \frac{\xi_{1-\alpha/2} \hat{\sigma}_{n,LD}}{\sqrt{n}} < \theta < \hat{\beta}_{n,LD} + \frac{\xi_{1-\alpha/2} \hat{\sigma}_{n,LD}}{\sqrt{n}} \right\}$$

будет асимптотическим доверительным интервалом для β надежности $1 - \alpha$.

5.5.2. Тесты наименьших модулей. Рассмотрим гипотезу $H_0: \beta = \beta_0$. Статистику $T_{n,LD}(\beta_0) := s_n^{-1} n^{-1/2} l_n^{LD}(\beta_0)$ возьмем в качестве тестовой. Вычислять ее проще, чем тестовую статистику, основанную на $\hat{\beta}_{n,LD}$. Условия 5.5.1, 5.5.2 обеспечивают выполнение условия сильного перемешивания (с.п.) для процесса $\{u_i\}$ с экспоненциальным коэффициентом перемешивания [101]. Процесс $\{u_{i-1} \text{sign } \varepsilon_i, i \in \mathbb{Z}\}$, очевидно, также удовлетворяет условию с.п. с экспоненциальным коэффициентом, а потому в силу центральной предельной теоремы для таких процессов (см., например, [15, теорема 18.5.3]) при гипотезе H_0

$$n^{-1/2} l_n^{LD}(\beta_0) = n^{-1/2} \sum_{k=1}^n u_{k-1} \text{sign } \varepsilon_k \xrightarrow{d} N(0, E_{\beta} u_1^2).$$

Следовательно, тестовая статистика $T_{n,LD}(\beta_0)$ при H_0 асимптотически распределена как стандартная гауссовская величина.

Рассмотрим теперь поведение тестовой статистики при альтернативах. Самая простая ситуация в случае фиксированной альтернативы $H_1: \beta \neq \beta_0$. Поскольку $\{u_{i-1} \text{sign}(u_i - \beta_0 u_{i-1}), i \in \mathbb{Z}\}$ — строго стационарная последовательность с с.п., в силу закона больших чисел для таких последовательностей (см., например, [37, 74, гл. 4, § 2,

теорема 2])

$$n^{-1}I_n^{LD}(\beta_0) \xrightarrow{P_\beta} \Lambda_{LD}(\beta_0) = -2E_\beta u_1 G((\beta_0 - \beta)u_1). \quad (5.5.5)$$

Функция $\Lambda_{LD}(\beta_0)$ при условиях 5.5.1, 5.5.2 дифференцируема по β_0 в любой точке β_0 , ее производная не больше нуля, причем $\Lambda'_{LD}(\beta) = = \lambda_{LD}(\beta) = -2g(0)E_\beta u_1^2 < 0$ и $\Lambda_{LD}(\beta) = 0$.

Перечисленные свойства позволяют утверждать, что $\Lambda_{LD}(\beta_0)$ обращается в нуль лишь в точке $\beta_0 = \beta$. Отсюда и из (5.5.5) теперь следует, что при H_1 $n^{-1/2}I_n^{LD}(\beta_0)$ расходится к бесконечности. Следовательно, тест, основанный на $T_{n,LD}(\beta_0)$, состоятелен против альтернативы H_1 .

Рассмотрим теперь случай сближающихся с H_0 альтернатив

$$H_{1n}: \beta = \beta_n, \quad \beta_n \rightarrow \beta_0.$$

Легко убедиться, что при H_{1n}

$$\sqrt{n}[n^{-1/2}I_n^{LD}(\beta_0) - \Lambda_{LD}(\beta_0)] = O_p(1). \quad (5.5.6)$$

Действительно, случайная величина в левой части (5.5.6) п.н. равна

$$n^{-1/2} \sum_{i=1}^n u_{i-1} + n^{-1/2} \sum_{i=1}^n [-2u_{i-1}I(\varepsilon_i < (\beta_0 - \beta_n)u_{i-1}) - \Lambda_{LD}(\beta_0)]. \quad (5.5.7)$$

Из (5.2.1) при $\beta = \beta_n$ следует

$$n^{-1/2} \sum_{i=1}^n u_i = \beta_n n^{-1/2} \sum_{i=1}^n u_{i-1} + n^{-1/2} \sum_{i=1}^n \varepsilon_i,$$

откуда

$$n^{-1/2} \sum_{i=1}^n u_{i-1} = (1 - \beta_n)^{-1} n^{-1/2} \sum_{i=1}^n \varepsilon_i + o_p(1) \xrightarrow{d_{\beta_n}} N(0, E\varepsilon_1^2(1 - \beta_0)^{-2}).$$

Следовательно, первая сумма в (5.5.7) есть $O_p(1)$. Вторая сумма в (5.5.7) центрирована, слагаемые в ней некоррелированы (см. похожее рассуждение относительно последовательности ξ_1, \dots, ξ_n при доказательстве леммы 5.8.1), откуда следует, что дисперсия этой суммы равномерно ограничена по n и сама эта сумма есть $O_p(1)$. Итак, (5.5.6) верно.

Теперь

$$n^{-1/2}I_n^{LD}(\beta_0) = \sqrt{n}[n^{-1}I_n^{LD}(\beta_0) - \Lambda_{LD}(\beta_0)] + \sqrt{n}\Lambda_{LD}(\beta_0),$$

и при $\sqrt{n}(\beta_n - \beta_0) \rightarrow \infty$

$$\sqrt{n}\Lambda_{LD}(\beta_0) \sim -2g(0)\sqrt{n}(\beta_0 - \beta_n)\mathbf{E}_{\beta_0}u_1^2 \rightarrow \infty.$$

Следовательно, если $\sqrt{n}(\beta_n - \beta_0) \rightarrow \infty$, статистика $n^{-1/2}I_n^{LD}(\beta_0)$ тоже расходится к бесконечности, и соответствующий тест состоятелен против таких альтернатив H_{1n} .

Рассмотрим теперь случай альтернатив

$$H_{1n}(a): \beta = \beta_n := \beta_0 + a n^{-1/2} + o(n^{-1/2}), \quad a - \text{постоянная.}$$

Анализ тестовой статистики при $H_{1n}(a)$ основан на следующей теореме 5.5.3, справедливой и при более общей, нежели $H_{1n}(a)$, альтернативе H_{1n} .

Т е о р е м а 5.5.3. Пусть выполнены условия 5.5.1, 5.5.2 и верна альтернатива H_{1n} .

Тогда при любом $0 < \Theta < \infty$

$$\sup_{|\theta| \leq \Theta} |n^{-1/2}I_n^{LD}(\beta_n + n^{-1/2}\theta) - n^{-1/2}I_n^{LD}(\beta_n) - \lambda_{LD}(\beta_0)\theta| = o_p(1).$$

Доказательство теоремы 5.5.3 аналогично доказательству теоремы 5.5.1, и мы его опускаем. Теорема 5.5.3 влечет

С л е д с т в и е 5.5.2. Пусть верна альтернатива $H_{1n}(a)$. Если $\sqrt{n}(\hat{\beta}_n - \beta_n) = O_p(1)$, то при условиях 5.5.1, 5.5.2

$$n^{-1/2}I_n^{LD}(\hat{\beta}_n) + \lambda_{LD}(\beta_0)\sqrt{n}(\hat{\beta}_n - \beta_n) + o_p(1).$$

Следствие 5.5.2 доказывается аналогично следствию 5.5.1. Теперь следствие 5.5.2 влечет следующую теорему.

Т е о р е м а 5.5.4. Пусть выполнены условия 5.5.1, 5.5.2 и верна альтернатива $H_{1n}(a)$.

Тогда:

$$\begin{aligned} 1) T_{n,LD}(\beta_0) &= \\ &= (\mathbf{E}_{\beta_0}u_1^2)^{-1/2} \sum_{k=1}^n u_{k-1} \text{sign } \varepsilon_k + 2g(0)(\mathbf{E}\varepsilon_1^2)^{1/2}(1 - \beta_0^2)^{-1/2}a + o_p(1), \end{aligned}$$

$$T_{n,LD}(\beta_0) \xrightarrow{d_{\beta_n}} N(2g(0)(\mathbf{E}\varepsilon_1^2)^{1/2}(1 - \beta_0^2)^{-1/2}a, 1);$$

$$2) \sqrt{n}(\hat{\beta}_{n,LD} - \beta_0) = -\lambda_{LD}^{-1}(\beta_0)n^{-1/2}I_n^{LD}(\beta_n) + a + o_p(1),$$

$$\sqrt{n}(\hat{\beta}_{n,LD} - \beta_0) \xrightarrow{d_{\beta_n}} N(a, \sigma_{LD}^2(\beta_0)),$$

где $\sigma_{LD}^2(\beta) = (1 - \beta^2)((2g(0))^2 E\varepsilon_1^2)^{-1}$.

Доказательство. Утверждение 1) непосредственно следует из (5.3.14), следствия 5.5.2 (с $\hat{\beta}_n = \beta_0$) и того факта, что при H_{1n}

$$(E_{\beta_0} u_1^2)^{-1/2} n^{-1/2} I_n^{LD}(\beta_n) =$$

$$= (E_{\beta_0} u_1^2)^{-1/2} n^{-1/2} \sum_{k=1}^n u_{k-1} \text{sign } \varepsilon_k \xrightarrow{d_{\beta_n}} N(0, 1).$$

Последнее соотношение устанавливается аналогично соотношению (5.3.15). Утверждение 2) получается рассуждениями, аналогичными доказательству теоремы 5.5.2. \square

Из теоремы 5.5.4 следует, в частности, что при $a = 0$, когда $\sqrt{n}(\beta_n - \beta_0) = o(1)$, гипотезы H_0 и $H_{1n}(0)$ асимптотически неразличимы.

Уместно отметить, что утверждения теоремы 5.5.4 можно установить и стандартным способом с помощью третьей леммы Ле Кама и теоремы 5.5.2, поскольку при условии 5.5.1 и некоторых дополнительных условиях на $G(x)$ (см. п. 5.3.2) наблюдения u_0, u_1, \dots, u_n удовлетворяют условию локальной асимптотической нормальности, и гипотезы H_0 и $H_{1n}(a)$ контигуальны.

Найдем АОЭ $e_{LD,LS}$ теста со статистикой $T_{n,LD}(\beta_0)$ относительно теста, основанного на

$$T_{n,LS}(\beta_0) = (1 - \beta_0^2)^{-1/2} \sqrt{n}(\hat{\beta}_{n,LS} - \beta_0).$$

Поскольку при $H_{1n}(a)$ в силу (5.3.12)

$$T_{n,LS}(\beta_0) \xrightarrow{d_{\beta_n}} N(a(1 - \beta_0^2), 1),$$

то в силу 1) теоремы 5.5.4 $e_{LD,LS} = (2g(0))^2 E\varepsilon_1^2$ не зависит от β и не изменяется при масштабных преобразованиях $G(x)$. Эффективность $e_{LD,LS}$ формально совпадает с АОЭ медианы относительно эмпирического среднего в схеме повторной выборки с неизвестным параметром сдвига (см., например, [19, 95, гл. 5, § 3]). Поэтому в силу [19, 95, теорема 3.3, гл. 3, § 5] для класса \mathfrak{G} функций распределения, имеющих четную плотность с максимумом в нуле, задача

$$e_{LD,LS} = (2g(0))^2 E\varepsilon_1^2 \implies \inf_{G \in \mathfrak{G}}$$

имеет решением $G(x) \sim R(-1/2, 1/2)$, и в этом случае $e_{LD,LS} = 1/3$ ($R(a, b)$ — равномерный закон распределения на $[a, b]$). При $G(x) \sim$

$\sim N(0, 1)$ $e_{LD,LS} = 2/\pi \approx 0.64$; для лапласовской $G(x)$ $e_{LD,LS} = 2$; для логистической $G(x)$ $e_{LD,LS} = \pi^2/12 \approx 0.82$. Для $G(x)$ с тяжелыми хвостами $e_{LD,LS}$ может быть сколь угодно большой. Например, если $G(x)$ — функция распределения Тьюки, т.е.

$$G(x) = (1 - \varepsilon)\Phi(x) + \varepsilon\Phi(x/\tau),$$

то

$$e_{LD,LS} = \frac{2}{\pi} \left(1 - \varepsilon + \frac{\varepsilon}{\tau}\right)^2 (1 - \varepsilon + \varepsilon\tau^2) \rightarrow \infty, \quad \tau \rightarrow \infty.$$

Разумеется, АОЭ оценки $\hat{\beta}_{n,LD}$ относительно $\hat{\beta}_{n,LS}$ также равна $e_{LD,LS}$, и для нее справедливо сказанное выше. Если $G(x)$ — лапласовская функция распределения, то АОЭ оценки $\hat{\beta}_{n,LD}$ относительно оценки максимального правдоподобия равна единице. В этом случае $\hat{\beta}_{n,LD}$ асимптотически эффективная оценка.

5.5.3. Взвешенные оценки наименьших модулей. Завершим параграф кратким описанием оценок, которые естественно назвать взвешенными оценками наименьших модулей. Итак, будем брать оценкой β (и обозначим ее $\hat{\beta}_{n,LDW}$) случайную величину, являющуюся решением задачи

$$L_n^{LDW}(\theta) := \sum_{k=1}^n |\varphi(u_{k-1})(u_k - \theta u_{k-1})| \implies \inf_{\theta \in \mathbb{R}^1}. \quad (5.5.8)$$

Целевая функция $L_n^{LDW}(\theta)$ аналогично $L_n^{LD}(\theta)$ выпукла вниз, а потому это решение всегда существует. Его можно изучать разными способами, например, переходя к эквивалентному уравнению

$$l_n^{LDW}(\theta) := \sum_{k=1}^n |\varphi(u_{k-1})| u_{k-1} \operatorname{sign}(u_k - \theta u_{k-1}) \div 0. \quad (5.5.9)$$

Относительно весовой функции $\varphi(u_{k-1})$ и $G(x)$ предположим выполненными следующие условия

У с л о в и е 5.5.3. $E_\beta |u_1|^{2+\Delta} |\varphi(u_1)| < \infty$ при некотором $\Delta > 0$, $0 < E_\beta u_1^2 |\varphi(u_1)| < \infty$, $E_\beta u_1^2 \varphi^2(u_1) < \infty$.

У с л о в и е 5.5.4. $E_\beta u_1 |\varphi(u_1)| = 0$ или $G(0) = 1/2$.

У с л о в и е 5.5.5. $n^{-1/2} \max_{1 \leq k \leq n} |u_k \varphi(u_k)| = o_p(1)$.

Аналогично доказательству теоремы 5.5.2 устанавливается: при условиях 5.2.3, 5.5.2–5.5.5

$$\sqrt{n}(\hat{\beta}_{n,LDW} - \beta) = \frac{1}{2g(0)E_\beta u_1^2 |\varphi(u_1)|} n^{-1/2} l_n^{LDW}(\beta) + o_p(1), \quad (5.5.10)$$

$$\sqrt{n}(\hat{\beta}_{n,LDW} - \beta) \xrightarrow{d_p} N(0, \sigma_{LDW}^2(\beta)), \quad (5.5.11)$$

где

$$\sigma_{LDW}^2(\beta) = E_{\beta}[u_1^2 \varphi^2(u_1)] \{ [2g(0)]^2 [E_{\beta} u_1^2 |\varphi(u_1)|]^2 \}^{-1}.$$

Если $E \varepsilon_1^2 < \infty$, и мы ограничиваемся простейшей задачей построения по наблюдениям u_0, \dots, u_n точечной оценки β , то взвешенные оценки наименьших модулей не нужны. Действительно,

$$[E_{\beta} u_1^2 |\varphi(u_1)|]^2 \leq E_{\beta} u_1^2 E_{\beta} u_1^2 \varphi^2(u_1),$$

откуда

$$\sigma_{LDW}^2(\beta) \geq \{ [2g(0)]^2 E_{\beta} u_1^2 \}^{-1} = \sigma_{LD}^2(\beta).$$

Равенство достигается лишь при $\varphi(u_{k-1})$, равной постоянной, в этом случае $\hat{\beta}_{n,LDW} = \hat{\beta}_{n,LD}$.

Существуют, однако, ситуации, когда использование взвешенных оценок целесообразно. Одна из таких ситуаций возникает в случае $E \varepsilon_1^2 = \infty$. Тогда нарушаются достаточные условия ЛАН, обеспечивающие \sqrt{n} -асимптотическую нормальность общепотребительных оценок, это затрудняет их применение к доверительному оцениванию и проверке гипотез. В то же время построить \sqrt{n} -асимптотическую нормальную $\hat{\beta}_{n,LDW}$ -оценку легко. Например, при $\varphi(u_{k-1}) = 1/u_{k-1}$ уравнение (5.5.9) переписывается в виде

$$\sum_{k=1}^n \text{sign} \left(\frac{u_k}{u_{k-1}} - \theta \right) \div 0, \quad (5.5.12)$$

и одно из решений (5.5.12), обращающее его в точное равенство, есть медиана $\hat{\beta}_{n,M}$ массива u_k/u_{k-1} , $k = 1, \dots, n$. При нечетных n $\hat{\beta}_{n,M}$ является единственным решением (5.5.12), а при четных n существует целый отрезок решений, и $\hat{\beta}_{n,M}$ — его середина.

Условия 5.2.3, 5.5.3–5.5.5 будут выполнены при $G(0) = 1/2$ и $E|\varepsilon_1|^{1+\Delta} < \infty$, так что при дополнительном условии 5.5.2

$$\begin{aligned} \sqrt{n}(\hat{\beta}_{n,M} - \beta) &= \\ &= \frac{1}{2g(0)E_{\beta}|u_1|} n^{-1/2} \sum_{k=1}^n \text{sign } u_{k-1} \text{sign } \varepsilon_k + o_p(1) \xrightarrow{d_p} N(0, \sigma_M^2(\beta)), \end{aligned}$$

где

$$\sigma_M^2(\beta) = (2g(0)E_{\beta}|u_1|)^{-2}.$$

Детальное исследование медианной оценки $\hat{\beta}_{n,M}$ проведено в [52], где отмечается, в частности, что

$$\frac{1}{2} \left(\sum_{k=1}^n \text{sign } u_{k-1} \text{sign } \varepsilon_k + n \right) \sim \text{Bi} \left(n, \frac{1}{2} \right)$$

при ограничениях

$$P\{\varepsilon_1 < 0\} = P\{\varepsilon_1 > 0\} = 1/2, \quad E \log^+ |\varepsilon_1| < \infty.$$

Следовательно, при этих условиях статистика

$$\sum_{k=1}^n \text{sign } u_{k-1} \text{sign}(u_k - \theta u_{k-1})$$

является центральной, и ее можно использовать для доверительного оценивания β при конечных n . Общеупотребительные оценки $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$ позволяют строить лишь асимптотические доверительные интервалы, свободные от $G(x)$, и притом лишь при $E\varepsilon_1^2 < \infty$.

Мы вернемся к обсуждению взвешенных оценок наименьших модулей в следующем параграфе в связи с вопросами устойчивости оценок к выбросам в наблюдениях u_0, \dots, u_n . Будет показано, что оценки $\hat{\beta}_{n,LDW}$ при подходящем выборе φ обладают ограниченными функционалами влияния (чувствительностями к большим ошибкам) в схеме засорения данных независимыми выбросами. Это другая привлекательная и важная особенность взвешенных оценок наименьших модулей. Общеупотребительные оценки $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$ имеют неограниченные функционалы влияния.

§ 5.6. Функционалы влияния оценок наименьших квадратов и наименьших модулей

Цель настоящего параграфа — оценить количественно устойчивость оценок $\hat{\beta}_{n,LS}$, $\hat{\beta}_{n,LD}$, $\hat{\beta}_{n,LDW}$ к выбросам в наблюдениях u_0, u_1, \dots, u_n . Сделать это можно разными способами, например с помощью функционалов влияния. Для схемы повторной выборки родственное понятие кривой влияния было введено Ф.Хампелем в [72] и оказалось, по словам П.Хубера [42, 81, § 1.5], самым полезным эвристическим инструментом теории робастных статистик. Мы далее будем придерживаться подхода, изложенного в [100].

Итак, пусть наблюдается вектор $Y_n = (y_0, y_1, \dots, y_n)$, где

$$y_i = u_i + z_i^\gamma \xi_i, \quad i \in \mathbb{Z}. \quad (5.6.1)$$

В (5.6.1) последовательность $\{u_i\}$ удовлетворяет (5.2.1); $\{z_i^\gamma\}$ — н.о.р.сл.в., $z_i^\gamma \sim \text{Bi}(1, \gamma)$, $0 \leq \gamma \leq 1$, γ — уровень засорения; $\{\xi_i\}$ — н.о.р.сл.в. с распределением μ_ξ из некоторого класса \mathcal{M}_ξ ; последовательности $\{u_i\}$, $\{z_i^\gamma\}$, $\{\xi_i\}$ предполагаются независимыми между собой. Мы рассматриваем, таким образом, простейшую схему засорения данных независимыми (одиночными) выбросами.

При засорении традиционные оценки перестают быть не только асимптотически нормальными (при естественной нормировке), но

даже и состоятельными. Чтобы охарактеризовать качество интересующей нас оценки $\hat{\beta}_n$ параметра β , построенной по наблюдениям Y_n , предположим, что существует предел

$$\hat{\beta}_n \xrightarrow{P_\beta} \theta_\gamma, \quad \theta_0 = \beta.$$

Простейшей инфинитезимальной характеристикой устойчивости оценки к засорению данных может служить производная

$$IF(\theta_\gamma, \mu_\xi) = \lim_{\gamma \rightarrow 0} \frac{\theta_\gamma - \theta_0}{\gamma}, \quad (5.6.2)$$

называемая *функционалом влияния* оценки $\hat{\beta}_n$ (разумеется, если эта производная существует). Она характеризует величину главного линейного члена в разложении асимптотического смещения

$$\theta_\gamma - \theta_0 = IF(\theta_\gamma, \mu_\xi) \gamma + o(\gamma). \quad (5.6.3)$$

Предпочтительнее использовать оценки, для которых величина

$$GES(\mathfrak{M}_\xi, \theta_\gamma) := \sup_{\mu_\xi \in \mathfrak{M}_\xi} |IF(\theta_\gamma, \mu_\xi)|,$$

называемая *чувствительностью к большим ошибкам*, конечна. В этом случае главный член разложения (5.6.3) для асимптотического смещения равномерно мал при всех возможных засорениях и малых γ . Качественно последнее означает, что оценка при малых γ будет слабо зависеть даже от очень больших выбросов в данных.

Для одного класса оценок найти $IF(\theta_\gamma, \mu_\xi)$ легко. А именно, предположим, оценкой β берется решение уравнения

$$\sum_{i=1}^n \psi_i(Y_n, \theta) = 0. \quad (5.6.4)$$

Оценки такого вида рассматривались давно. Выделим несколько частных случаев. Если

$$\psi_i(Y_n, \theta) = \eta((1 - \theta^2)^{1/2} y_{i-1}, y_i - \theta y_{i-1}),$$

где η — некоторая ограниченная функция, то получаются обобщенные M -оценки (GM -оценки) (см. [90, 98]). При

$$\eta(\xi_1, \xi_2) = \xi_1 \psi(\xi_2)$$

получаются обычные M -оценки (см. [25, 98, 99]), включающие в себя оценки наименьших квадратов и наименьших модулей. Если

$$\begin{aligned} \psi_i(Y_n, \theta) &= \eta(y_i - \theta y_{i-1}, y_{i+1} - \theta y_i) + \theta \eta(y_i - \theta y_{i-1}, y_{i+2} - \theta y_{i+1}) + \dots \\ &\dots + \theta^{n-1} \eta(y_i - \theta y_{i-1}, y_n - \theta y_{n-1}), \quad i = 1, 2, \dots, n-1, \quad \psi_n(Y_n, \theta) = 0, \end{aligned}$$

то получаются RA -оценки [58].

Вернемся к уравнению (5.6.4). Пусть для некоторой функции $\Lambda(\gamma, \theta)$ при $0 \leq \gamma \leq \delta$, $|\theta - \beta| \leq \delta$ выполнено

$$\text{У с л о в и е 5.6.1. } n^{-1} \sum_{i=1}^n \psi_i(Y_n, \theta) \xrightarrow{P_n} \Lambda(\gamma, \theta).$$

Пусть выполнены также следующие условия.

$$\text{У с л о в и е 5.6.2. } \Lambda(0, \beta) = 0.$$

У с л о в и е 5.6.3. Частные производные $\frac{\partial}{\partial \gamma} \Lambda(\gamma, \theta)$ и $\frac{\partial}{\partial \theta} \Lambda(\gamma, \theta)$ существуют и непрерывны по (γ, θ) для $0 \leq \gamma \leq \delta$, $|\theta - \beta| \leq \delta$.

$$\text{У с л о в и е 5.6.4. } \lambda(\beta) := \frac{\partial}{\partial \theta} \Lambda(0, \beta) \neq 0.$$

При условиях 5.6.1–5.6.4 в силу теоремы о неявной функции в некоторой окрестности точки $(0, \beta)$ уравнение

$$\Lambda(\gamma, \theta) = 0 \tag{5.6.5}$$

определяет однозначную и непрерывную функцию $\theta_\gamma = \theta(\gamma)$, $\theta_\gamma \rightarrow \theta_0 = \beta$ при $\gamma \rightarrow 0$. Более того, функция θ_γ дифференцируема и, в частности,

$$\left. \frac{d\theta_\gamma}{d\gamma} \right|_{\gamma=0} = -\lambda^{-1}(\beta) \frac{\partial}{\partial \gamma} \Lambda(0, \beta). \tag{5.6.6}$$

При условиях 5.6.1–5.6.4 при малых γ с вероятностью, стремящейся к единице, существует такой корень $\hat{\beta}_n$ уравнения (5.6.4), что $\hat{\beta}_n \xrightarrow{P_n} \theta_\gamma$. Тогда соответствующий $\hat{\beta}_n$ функционал влияния дается в силу (5.6.6) соотношением

$$IF(\theta_\gamma, \mu_\xi) = -\lambda^{-1}(\beta) \frac{\partial}{\partial \gamma} \Lambda(0, \beta). \tag{5.6.7}$$

5.6.1. Функционал влияния оценки наименьших квадратов. Обратимся в качестве примера к оценке наименьших квадратов $\hat{\beta}_{n,LS}$.

Предположения. $E\varepsilon_1 = 0$, $E\varepsilon_1^2 < \infty$, $E\xi_1^2 < \infty$.

Оценка $\hat{\beta}_{n,LS}$ определяется соотношением (5.6.4) с

$$\psi_i^{LS}(Y_n, \theta) = y_{i-1}(y_i - \theta y_{i-1}). \tag{5.6.8}$$

В силу эргодической теоремы при любых γ, θ существует предел

$$n^{-1} \sum_{i=1}^n \psi_i^{LS}(Y_n, \theta) \xrightarrow{P_\theta} \Lambda_{LS}(\gamma, \theta),$$

где

$$\Lambda_{LS}(\gamma, \theta) = \mathbf{E}_\beta y_0 (y_1 - \theta y_0) = \mathbf{E}_\beta u_0 u_1 - \theta (\mathbf{E}_\beta u_0^2 + \gamma \mathbf{E} \xi_0^2) + \gamma^2 (\mathbf{E} \xi_0^2)^2.$$

Легко видеть, что

$$\Lambda_{LS}(0, \beta) = \mathbf{E}_\beta u_0 \varepsilon_1 = 0$$

в силу независимости ε_1 и u_0 . Итак, условия 5.6.1, 5.6.2 выполнены. Условия 5.6.3, 5.6.4 также, очевидно, выполнены при любых γ, θ , и

$$\lambda_{LS}(\beta) := \frac{\partial}{\partial \theta} \Lambda_{LS}(0, \beta) = -\mathbf{E}_\beta u_0^2 = -\mathbf{E} \varepsilon_1^2 (1 - \beta^2)^{-1} \neq 0,$$

$$\frac{\partial}{\partial \gamma} \Lambda_{LS}(0, \beta) = -\beta \mathbf{E} \xi_0^2.$$

Решение уравнения (5.6.5) в рассматриваемом случае при всех γ единственно и равно

$$\theta_\gamma^{LS} = \mathbf{E}_\beta y_0 y_1 / \mathbf{E}_\beta y_0^2.$$

Уравнение (5.6.4) с $\psi_i^{LS}(Y_n, \theta)$ из (5.6.8) с вероятностью, стремящейся к единице, при всех γ имеет единственное решение

$$\hat{\beta}_{n,LS} = \sum_{i=1}^n y_{i-1} y_i / \sum_{i=1}^n y_{i-1}^2, \quad \hat{\beta}_{n,LS} \xrightarrow{P_\theta} \theta_\gamma^{LS}.$$

В силу (5.6.7) получаем

$$IF(\theta_\gamma^{LS}, \mu_\xi) = - \frac{\beta(1-\beta^2)}{\mathbf{E} \varepsilon_1^2} \mathbf{E} \xi_0^2. \quad (5.6.9)$$

Если $\xi_i = \xi$, ξ — постоянная и $IF(\theta_\gamma^{LS}, \xi) := IF(\theta_\gamma^{LS}, \mu_\xi)$, то

$$IF(\theta_\gamma^{LS}, \xi) = - \frac{\beta(1-\beta^2)}{\mathbf{E} \varepsilon_1^2} \xi^2 \quad (5.6.10)$$

— неограниченная непрерывная функция ξ .

Пусть \mathfrak{M}_i , $i = 1, 2$, — класс распределений μ_ξ с конечным i -м абсолютным моментом. Тогда

$$GES(\mathfrak{M}_2, \theta_\gamma^{LS}) = \infty, \quad \beta \neq 0,$$

и, таким образом, даже малая доля больших выбросов в данных может сколь угодно сильно влиять на $\hat{\beta}_{n,LS}$.

5.6.2. Функционал влияния оценки наименьших модулей. Рассмотрим теперь оценку наименьших модулей $\hat{\beta}_{n,LD}$. Пусть выполнены следующие

Предположения. $E\epsilon_1 = 0$, $E\epsilon_1^2 < \infty$, $E\xi_1^2 < \infty$, $g(x)$ непрерывна и ограничена, $g(0) > 0$.

В случае засорения оценка наименьших модулей определяется соотношением (5.6.4), в котором точное равенство следует заменить на символ перехода через нуль, а

$$\psi_i^{LD}(Y_n, \theta) = y_{i-1} \text{sign}(y_i - \theta y_{i-1}).$$

Условие 5.6.1 выполнено в силу эргодической теоремы при любых γ , θ с функцией

$$\Lambda_{LD}(\gamma, \theta) = E_{\beta} y_0 \text{sign}(y_1 - \theta y_0) = E_{\beta} y_0 (1 - 2G((\theta - \beta)u_0 - z_1^{\gamma} \xi_1 + \theta z_0^{\gamma} \xi_0)).$$

Очевидно, $\Lambda_{LD}(0, \beta) = E_{\beta} u_0 \text{sign} \epsilon_1 = 0$, и условие 5.6.2 выполнено также. Чтобы проверить условия 5.6.3, 5.6.4, введем события \mathbb{H}_i , $i = 0, 1, 2$, состоящие в том, что точно i величин из z_0^{γ} , z_1^{γ} отличны от нуля. С помощью формулы полного математического ожидания перепишем $\Lambda_{LD}(\gamma, \theta)$ в виде

$$\begin{aligned} \Lambda_{LD}(\gamma, \theta) &= \\ &= \sum_{i=0}^2 E_{\beta} \{ y_0 (1 - 2G((\theta - \beta)u_0 - z_1^{\gamma} \xi_1 + \theta z_0^{\gamma} \xi_0)) | \mathbb{H}_i \} P\{\mathbb{H}_i\}. \end{aligned} \quad (5.6.11)$$

Условные средние в (5.6.11) не зависят от γ и непрерывно дифференцируемы по θ , а $P\{\mathbb{H}_i\}$ — полиномы от γ . Следовательно, $\Lambda_{LD}(\gamma, \theta)$ имеет непрерывные частные производные по γ и θ . Находим из (5.6.11)

$$\frac{\partial}{\partial \gamma} \Lambda_{LD}(0, \beta) = E_{\beta} \xi_0 (1 - 2G(\beta \xi_0)),$$

$$\frac{\partial}{\partial \theta} \Lambda_{LD}(0, \beta) = \lambda_{LD}(\beta) = -2g(0) E_{\beta} u_0^2 = -2g(0) E \epsilon_1^2 (1 - \beta^2)^{-1} \neq 0.$$

Напомним, что функция $\lambda_{LD}(\beta)$ уже появлялась в п. 5.5.1. Мы проверили все условия 5.6.1–5.6.4. Уравнение (5.6.4) в силу монотонности по θ функции $\psi_i^{LD}(Y_n, \theta)$ имеет п.н. единственное решение $\hat{\beta}_{n,LD}$, а уравнение (5.6.5) при малых γ — единственное решение θ_{γ}^{LD} ,

$\hat{\beta}_{n,LD} \xrightarrow{P_g} \theta_\gamma^{LD}$. В силу (5.6.7) получаем

$$IF(\theta_\gamma^{LD}, \mu_\xi) = \frac{1 - \beta^2}{2g(0)\mathbb{E}\varepsilon_1^2} \mathbb{E}\xi_0(1 - 2G(\beta\xi_0)). \quad (5.6.12)$$

Если $\xi_i = \xi$ и ξ — постоянная, то

$$IF(\theta_\gamma^{LD}, \xi) = \frac{1 - \beta^2}{2g(0)\mathbb{E}\varepsilon_1^2} \xi(1 - 2G(\beta\xi)) \quad (5.6.13)$$

— неограниченная непрерывная функция ξ . Поскольку

$$GES(\mathfrak{M}_2, \theta_\gamma^{LD}) = \infty, \quad \beta \neq 0,$$

то $\hat{\beta}_{n,LD}$ неустойчива к выбросам в наблюдениях u_0, \dots, u_n . Сравнивая соотношение (5.6.13) с соотношением (5.6.10), видим, что вместо квадратичного роста по ξ для $\hat{\beta}_{n,LS}$, в случае $\hat{\beta}_{n,LD}$ имеем линейную зависимость. Качественно последнее означает, что оценка $\hat{\beta}_{n,LD}$ меньше зависит от выбросов, чем $\hat{\beta}_{n,LS}$.

5.6.3. Функционал влияния взвешенной оценки наименьших модулей. Найдем теперь функционал влияния взвешенных оценок наименьших модулей.

Условие 5.6.5. $\mathbb{E} \log^+ |\varepsilon_1| < \infty$, $\mathbb{E}_\beta y_1^2 |\varphi(y_1)| < \infty$ при малых γ , $\mathbb{E}_\beta u_1^2 |\varphi(u_1)| \neq 0$.

Условие 5.6.6. $\mathbb{E}_\beta u_1 |\varphi(u_1)| = 0$ или $G(0) = 1/2$; $g(x)$ непрерывна и ограничена, $g(0) > 0$.

Взвешенная оценка наименьших модулей $\hat{\beta}_{n,LDW}$ в случае засорения получается решением уравнения

$$\sum_{k=1}^n |\varphi(y_{k-1})| y_{k-1} \operatorname{sign}(y_k - \theta y_{k-1}) \div 0. \quad (5.6.14)$$

Рассуждая, как для оценки наименьших модулей, получим, что любое решение последнего уравнения при малых γ сходится по вероятности к θ_γ^{LDW} и

$$\begin{aligned} IF(\theta_\gamma^{LDW}, \mu_\xi) &= \\ &= \frac{\mathbb{E}_\beta [u_0 |\varphi(u_0)| (1 - 2G(-\xi_1)) + (u_0 + \xi_0) |\varphi(u_0 + \xi_0)| (1 - 2G(\beta\xi_0))]}{2g(0)\mathbb{E}_\beta u_1^2 |\varphi(u_1)|}. \end{aligned}$$

Если $\varphi(x) = 1$ и $\mathbb{E}\varepsilon_1 = 0$, то последнее выражение совпадает с функ-

ционалом влияния оценки наименьших модулей (5.6.12). Выражение для $IF(\theta_\gamma^{LDW}, \mu_\xi)$ совместно с требованием $\sup_x |x\varphi(x)| < \infty$ влечет

$$GES(\mathcal{M}_\xi, \theta_\gamma^{LDW}) < \infty,$$

где \mathcal{M}_ξ — класс распределений μ_ξ , удовлетворяющих условию $E_\beta y_1^2 |\varphi(y_1)| < \infty$. В частности, для $\varphi(y_{k-1}) = 1/y_{k-1}$ оценка $\hat{\beta}_{n,LDW}$ есть медианная оценка $\hat{\beta}_{n,M}$, $\theta_\gamma^{LDW} = \theta_\gamma^M$. Тогда при моментных ограничениях $0 < E|\varepsilon_1| < \infty$, $E|\xi_1| < \infty$ (они обеспечивают выполнение условия 5.6.5) и условия 5.6.6

$$IF(\theta_\gamma^M, \mu_\xi) = \frac{E_\beta [\text{sign } u_0(1 - 2G(-\xi_1)) + \text{sign}(u_0 + \xi_0)(1 - 2G(\beta\xi_0))]}{2g(0)E_\beta |u_1|}. \quad (5.6.15)$$

Из (5.6.15) следует, что

$$GES(\mathcal{M}_1, \theta_\gamma^M) \leq [g(0)E_\beta |u_1|]^{-1}.$$

В заключение параграфа отметим, что для схемы независимых выбросов и для более общей, чем рассмотрено здесь, схемы засорения, когда выбросы образуют пачки, для M -, GM -, RA -оценок (но только с гладкими функциями ψ_i в (5.6.4)) функционалы влияния найдены в [100]. В этой же работе установлена простая связь между функционалами влияния и функцией влияния Хампеля, которая может быть введена не только для независимых, одинаково распределенных наблюдений, но и для временных рядов.

§ 5.7. О проверке стационарности авторегрессионного уравнения

Рассмотрим естественное обобщение схемы (5.2.1) — простейшее нестационарное соотношение авторегрессии

$$u_i = \beta_i u_{i-1} + \varepsilon_i, \quad i \in \mathbb{Z}. \quad (5.7.1)$$

Предположим, что верно условие 5.5.1 и $\sup_i |\beta_i| < 1$. Тогда существует решение уравнения (5.7.1) вида

$$u_i = \sum_{j=0}^{\infty} \gamma_{ij} \varepsilon_{i-j}, \quad (5.7.2)$$

где $\gamma_{i0} = 1$, $\gamma_{ij} = \beta_i \beta_{i-1} \dots \beta_{i-j+1}$ при $j \geq 1$. Ряд (5.7.1) сходится в $\mathbb{L}^{2+\Delta}$.

Отметим, что для $\{u_i\}$ из (5.7.2) $\sup_i \mathbf{E} u_i^2 < \infty$, и для любого решения уравнения (5.7.1), удовлетворяющего последнему неравенству, справедливо представление (5.7.2).

Пусть наблюдаются u_0, u_1, \dots, u_n , удовлетворяющие (5.7.2). Цель настоящего параграфа состоит в построении тестов для проверки гипотезы стационарности

$$H_0: \beta_i = \beta, \quad i \in \mathbb{Z}, \quad (5.7.3)$$

β неизвестно, и исследования мощности тестов при альтернативе

$$H_{1n}: \beta_k = \beta + a_{kn} n^{-1/2}, \quad k = 1, \dots, n, \quad \sup_{k,n} |a_{kn}| \leq A < \infty. \quad (5.7.4)$$

Для проверки H_0 будем использовать оценки $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$ параметра β , построенные в предположении справедливости H_0 . Начнем наше рассмотрение с технически более сложного случая, когда используется оценка наименьших модулей $\hat{\beta}_{n,LD}$, получаемая решением уравнения (5.5.3)

$$\sum_{k=1}^n u_{k-1} \text{sign}(u_k - \theta u_{k-1}) \div 0.$$

Введем определяемый этим уравнением случайный процесс

$$w_n^{LD}(t) = \sigma_u^{-1}(\beta) n^{-1/2} \sum_{k \leq nt} u_{k-1} \text{sign}(u_k - \hat{\beta}_{n,LD} u_{k-1}), \quad t \in [0, 1],$$

где $\sigma_u^2(\beta) = \mathbf{E} \varepsilon_1^2 (1 - \beta^2)^{-1}$. Заметим, что при H_0 $\mathbf{E} \beta u_1^2 = \sigma_u^2(\beta)$.

Нашей ближайшей целью будет понять характер асимптотического поведения $w_n^{LD}(t)$ при H_0 . Эмпирически сделать это довольно просто.

Обозначим через $w(t)$, $t \in [0, 1]$, броуновский мост, т.е. $w(t) = \nu(t) - t\nu(1)$, где $\nu(t)$ — стандартный винеровский процесс. Пусть $\mathbb{D}[0, 1]$ — метрическое пространство функций без разрывов второго рода с метрикой Скорохода (определение см., например, [4, 49, гл. 5]).

Обозначим также

$$\nu_n^{LD}(t) = \sigma_u^{-1}(\beta) n^{-1/2} \sum_{k \leq nt} u_{k-1} \text{sign} \varepsilon_k. \quad (5.7.5)$$

Для случайной последовательности $\{u_{i-1} \text{sign} \varepsilon_i, i \in \mathbb{Z}\}$ при условии 5.5.1 и H_0 выполнены предположения теоремы 21.1 [4, 49] (это вытекает из рассуждений примера 1 в [4, 49, § 21 гл. 4]), в силу которой $\nu_n^{LD}(t)$ слабо сходится в $\mathbb{D}[0, 1]$ к $\nu(t)$. Последнее утверждение есть

один из вариантов известного принципа инвариантности Донскера—Прохорова (см. [4, 49, § 16]). Теперь при любом фиксированном $t > 0$ в предположении, что H_0 верна, получаем в силу следствия 5.5.1

$$w_n^{LD}(t) = \sigma_u^{-1}(\beta) n^{-1/2} \sum_{k \leq nt} u_{k-1} \text{sign } \varepsilon_k + \\ + \sigma_u^{-1}(\beta) \lambda_{LD}(\beta) \sqrt{t} \sqrt{nt} (\hat{\beta}_{n,LD} - \beta) + o_p(1). \quad (5.7.6)$$

Функция $\lambda_{LD}(\beta) = -2g(0) \mathbb{E} \varepsilon_1^2 (1 - \beta^2)^{-1}$ была введена в п. 5.5.1.

В силу п. 1 теоремы 5.5.2 при условиях 5.5.1, 5.5.2

$$\sqrt{n}(\hat{\beta}_{n,LD} - \beta) = -\lambda_{LD}^{-1}(\beta) n^{-1/2} \sum_{k=1}^n u_{k-1} \text{sign } \varepsilon_k + o_p(1).$$

Подставляя последнее выражение для $\sqrt{n}(\hat{\beta}_{n,LD} - \beta)$ в (5.7.6), получаем в силу определения (5.7.5)

$$w_n^{LD}(t) = \nu_n^{LD}(t) - t\nu_n^{LD}(1) + o_p(1). \quad (5.7.7)$$

Соотношение (5.7.7) (справедливость его установлена только при фиксированном t) и известная связь $w(t) = \nu(t) - t\nu(1)$ позволяет предположить (и это эмпирическая часть рассуждения), что при гипотезе H_0 $w_n^{LD}(t)$ слабо сходится в $\mathbb{D}[0, 1]$ к броуновскому мосту $w(t)$.

Естественно ожидать, как это обычно бывает, что при альтернативе H_{1n} распределения $w_n^{LD}(t)$ будут отличаться сдвигом. Чтобы охарактеризовать последний, введем функцию

$$a_n(t) = n^{-1} \sum_{k \leq nt} a_{kn},$$

и для некоторой функции $a(t) \in \mathbb{D}[0, 1]$ будем считать выполненным

$$\text{Условие 5.7.1. } \sup_{0 \leq t \leq 1} |a_n(t) - a(t)| \rightarrow 0.$$

Сформулируем теперь точное утверждение о поведении $w_n^{LD}(t)$ при H_0 и H_{1n} . Обозначим для краткости $-\sigma_u^{-1}(\beta) \lambda_{LD}(\beta)$ через $\sigma_{LD}(\beta)$, т.е.

$$\sigma_{LD}^{-1}(\beta) = 2g(0) (\mathbb{E} \varepsilon_1^2)^{1/2} (1 - \beta^2)^{-1/2}. \quad (5.7.8)$$

Функция σ_{LD}^2 есть асимптотическая дисперсия оценки наименьших модулей $\hat{\beta}_{n,LD}$ (см. теорему 5.5.2).

Теорема 5.7.1. Пусть выполнены условия 5.5.1, 5.5.2.

Тогда при гипотезе H_0 из (5.7.3) $w_n^{LD}(t)$ сходится слабо в $\mathbb{D}[0, 1]$ к броуновскому мосту $w(t)$, а при альтернативе H_{1n} из (5.7.4) и

дополнительном условии 5.7.1 — к $w(t) + \sigma_{LD}^{-1}(\beta)\delta(t)$, где $\delta(t) = a(t) - ta(1)$.

Прежде чем доказывать теорему 5.7.1, обсудим ее следствия.

Пусть

$$\hat{w}_n^{LD}(t) = s_n^{-1} n^{-1/2} \sum_{k \leq nt} u_{k-1} \text{sign}(u_k - \hat{\beta}_{n,LD} u_{k-1}),$$

где $s_n^2 = n^{-1} \sum_{k=1}^n u_k^2$.

Легко проверить, что и при H_0 , и при H_{1n} s_n^2 является состоятельной оценкой $\sigma_u^2(\beta)$, а потому для $\hat{w}_n^{LD}(t)$ справедливы те же утверждения, что и для $w_n^{LD}(t)$. Поэтому в условиях теоремы 5.7.1 при H_0 из (5.7.3) статистики

$$\hat{D}_n^{LD} := \sup_{0 \leq t \leq 1} |\hat{w}_n^{LD}(t)| = s_n^{-1} \max_{m \leq n} \left| n^{-1/2} \sum_{k \leq m} u_{k-1} \text{sign}(u_k - \hat{\beta}_{n,LD} u_{k-1}) \right|,$$

$$\begin{aligned} (\hat{w}_n^{LD})^2 &:= s_n^{-2} \int_0^1 (\hat{w}_n^{LD}(t))^2 dt = \\ &= s_n^{-2} n^{-2} \sum_{m \leq n} \left(\sum_{k \leq m} u_{k-1} \text{sign}(u_k - \hat{\beta}_{n,LD} u_{k-1}) \right)^2 \end{aligned}$$

сходятся по распределению соответственно к $\sup_{0 \leq t \leq 1} |w(t)|$ и $\int_0^1 w^2(t) dt$. Распределения последних — известные и табулированные распределения Колмогорова и Смирнова [10, табл. 6.1 и 6.4a].

Статистики \hat{D}_n^{LD} и $(\hat{w}_n^{LD})^2$ — статистики тестов для проверки гипотезы H_0 .

При альтернативе H_{1n} из (5.7.4) и условиях 5.5.1, 5.5.2, 5.7.1 \hat{D}_n^{LD} и $(\hat{w}_n^{LD})^2$ сходятся соответственно к

$$\sup_{0 \leq t \leq 1} |w(t) + \sigma_{LD}^{-1}(\beta)\delta(t)| \quad \text{и} \quad \int_0^1 (w(t) + \sigma_{LD}^{-1}(\beta)\delta(t))^2 dt.$$

Доказательство теоремы 5.7.1. Приведем схему доказательства в предположении справедливости H_{1n} . Подробности можно найти в [9].

Справедливо следующее соотношение, доказываемое аналогично

теореме 5.5.1: при любом $0 < \Theta < \infty$ в условиях теоремы 5.7.1

$$\begin{aligned} \sup_{|\theta| \leq \Theta} \sup_{1 \leq m \leq n} \left| n^{-1/2} \sum_{k=1}^m \text{sign}(u_k - (\beta + n^{-1/2} \theta) u_{k-1}) - \right. \\ \left. - n^{-1/2} \sum_{k=1}^m u_{k-1} \text{sign} \varepsilon_k - \lambda_{LD}(\beta) \theta \frac{m}{n} + \right. \\ \left. + \lambda_{LD}(\beta) n^{-1} \sum_{k=1}^m a_{kn} \right| = o_p(1). \quad (5.7.9) \end{aligned}$$

Из (5.7.9) аналогично следствию 5.5.1 получаем, что для любой последовательности $\hat{\beta}_n$ такой, что $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$, имеет место соотношение

$$\begin{aligned} \sup_{1 \leq m \leq n} \left| n^{-1/2} \sum_{k=1}^m \text{sign}(u_k - \hat{\beta}_n u_{k-1}) - \right. \\ \left. - n^{-1/2} \sum_{k=1}^m u_{k-1} \text{sign} \varepsilon_k - \lambda_{LD}(\beta) \sqrt{n}(\hat{\beta}_n - \beta) \frac{m}{n} + \right. \\ \left. + \lambda_{LD}(\beta) n^{-1} \sum_{k=1}^m a_{kn} \right| = o_p(1). \quad (5.7.10) \end{aligned}$$

Пусть процесс $\nu_n^{LD}(t)$ определен в (5.7.5). При H_{1n} справедливо следующее разложение, устанавливаемое аналогично 2) теоремы 5.5.4:

$$\sqrt{n}(\hat{\beta}_{n,LD} - \beta) = \sigma_{LD}(\beta) \nu_n^{LD}(1) + a_n(1) + o_p(1), \quad (5.7.11)$$

$\sigma_{LD}(\beta)$ определена в (5.7.8).

В силу (5.7.11) $\sqrt{n}(\hat{\beta}_{n,LD} - \beta) = O_p(1)$. Отсюда и из (5.7.10) при $\hat{\beta}_n = \hat{\beta}_{n,LD}$ несложно получить

$$\begin{aligned} \sup_{0 \leq t \leq 1} |w_n^{LD}(t) - \nu_n^{LD}(t) + \\ + \sigma_{LD}^{-1}(\beta) \sqrt{n}(\hat{\beta}_{n,LD} - \beta)t - \sigma_{LD}^{-1}(\beta) a_n(t)| = o_p(1). \quad (5.7.12) \end{aligned}$$

Подставляя (5.7.11) в (5.7.12), получаем

$$\begin{aligned} \sup_{0 \leq t \leq 1} |w_n^{LD}(t) - (\nu_n^{LD}(t) - t \nu_n^{LD}(1)) - \\ - \sigma_{LD}^{-1}(\beta)(a_n(t) - t a_n(1))| = o_p(1), \quad (5.7.13) \end{aligned}$$

В силу принципа инвариантности для процессов с φ -перемешиванием [4, 49, теорема 21.1] $\nu_n^{LD}(t)$ слабо сходится к стандартному винне-

ровскому процессу $\nu(t)$. Следовательно, $\nu_n^{LD}(t) - t\nu_n^{LD}(1)$ сходится к броуновскому мосту $w(t) = \nu(t) - t\nu(1)$, что совместно с (5.7.13) доказывает теорему 5.7.1. \square

Рассмотрим теперь оценку наименьших квадратов $\hat{\beta}_{n,LS}$. Она определяется уравнением (5.3.1)

$$\sum_{k=1}^n u_{k-1}(u_k - \theta u_{k-1}) = 0,$$

которое в свою очередь определяет процесс

$$w_n^{LS}(t) = (\mathbb{E}\varepsilon_1^2)^{-1/2} \sigma_u^{-1}(\beta) n^{-1/2} \sum_{k \leq nt} u_{k-1}(u_k - \hat{\beta}_{n,LS} u_{k-1}), \quad t \in [0, 1].$$

Аналогично теореме 5.7.1 получаем, что при условии 5.2.1 и H_0 $w_n^{LS}(t)$ слабо сходится в $\mathbb{D}[0, 1]$ к броуновскому мосту $w(t)$, а при H_{1n} и дополнительном условии 5.7.1 — к $w(t) + \sigma_{LS}^{-1}(\beta) \delta(t)$, где $\sigma_{LS}^2(\beta) = 1 - \beta^2$ — асимптотическая дисперсия $\hat{\beta}_{n,LS}$.

Понятно, как строить дальше процесс $\hat{w}_n^{LS}(t)$ и статистики \hat{D}_n^{LS} , $(\hat{\omega}_n^{LS})^2$.

Разумеется, вместо $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$ можно использовать и иные оценки β , лишь бы соответствующий процесс был согласован с уравнением, определяющим оценку. Например, в [52] оценкой β бралась медиана $\hat{\beta}_{n,M}$ массива $\{u_k/u_{k-1}, k = 1, \dots, n\}$, тогда

$$w_n^M(t) = n^{-1/2} \sum_{k \leq nt} \text{sign } u_{k-1} \text{sign}(u_k - \hat{\beta}_{n,M} u_{k-1}).$$

§ 5.8. Приложение. Доказательства теорем

Доказательство теоремы 5.5.1. Пусть

$$\Delta_k(x) = \begin{cases} 1, & \varepsilon_k < x, \\ 1/2, & \varepsilon_k = x, \\ 0, & \varepsilon_k > x, \end{cases} \quad k = 1, \dots, n.$$

Легко проверить, что для $I_n^{LD}(\theta)$ из (5.5.2) справедливо представление

$$n^{-1/2} I_n^{LD}(\beta + n^{-1/2} \theta) - n^{-1/2} I_n^{LD}(\beta) = -2z_{1n}(\theta) - 2z_{2n}(\theta), \quad (5.8.1)$$

где

$$z_{1n}(\theta) = n^{-1/2} \sum_{k=1}^n u_{k-1} \left(\Delta_k(n^{-1/2}\theta u_{k-1}) - G(n^{-1/2}\theta u_{k-1}) - \Delta_k(0) + G(0) \right),$$

$$z_{2n}(\theta) = n^{-1/2} \sum_{k=1}^n u_{k-1} \left(G(n^{-1/2}\theta u_{k-1}) - G(0) \right).$$

Лемма 5.8.1. При условиях 5.5.1, 5.5.2 $\sup_{|\theta| \leq \Theta} |z_{1n}(\theta)| = o_p(1)$.

Доказательство. Представим $z_{1n}(\theta)$ в виде

$$z_{1n}(\theta) = p_{1n}(\theta) + p_{2n}(\theta),$$

где

$$p_{1n}(\theta) = n^{-1/2} \sum_{k=1}^n u_{k-1} I(u_{k-1} > 0) \times$$

$$\times \left(\Delta_k(n^{-1/2}\theta u_{k-1}) - G(n^{-1/2}\theta u_{k-1}) - \Delta_k(0) + G(0) \right),$$

$I(\cdot)$ — индикатор события, а $p_{2n}(\theta)$ определяется аналогичным выражением, в котором $I(u_{k-1} > 0)$ заменен на $I(u_{k-1} \leq 0)$. Разделим отрезок $[-\Theta, \Theta]$ на 3^{m_n} частей точками

$$\eta_{sn} = -\Theta + 2\Theta 3^{-m_n} s, \quad s = 0, 1, \dots, 3^{m_n},$$

и пусть $3^{m_n} \sim n^{\delta/4}$, где $\delta = \min(\Delta, 1)$. Пусть η_{jn} — та точка среди $\{\eta_{sn}\}$, для которой

$$0 \leq \eta_{jn} - \theta \leq 2\Theta 3^{-m_n}. \quad (5.8.2)$$

Обозначим

$$\hat{u}_{k-1,s} = u_{k-1} (1 - 2\Theta 3^{-m_n} \eta_{sn}^{-1} I(u_{k-1} \leq 0)), \quad (5.8.3)$$

$$\tilde{u}_{k-1,s} = u_{k-1} (1 - 2\Theta 3^{-m_n} \eta_{sn}^{-1} I(u_{k-1} \geq 0)), \quad k = 1, \dots, n. \quad (5.8.4)$$

Для построенных величин в силу (5.8.2) и определений (5.8.3), (5.8.4) справедливы неравенства

$$\dot{\eta}_{jn} \tilde{u}_{k-1,j} \leq \theta u_{k-1} \leq \eta_{jn} \hat{u}_{k-1,j}, \quad k = 1, \dots, n. \quad (5.8.5)$$

Обозначим

$$\hat{U}_{sn} = (\hat{u}_{0s}, \dots, \hat{u}_{n-1,s}), \quad \tilde{U}_{sn} = (\tilde{u}_{0s}, \dots, \tilde{u}_{n-1,s}).$$

Обозначим также

$$q_n(\eta_{jn}, \hat{U}_{jn}) = n^{-1/2} \sum_{k=1}^n u_{k-1} I(u_{k-1} > 0) \times \\ \times \left(\Delta_k(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - G(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - \Delta_k(0) + G(0) \right).$$

В силу монотонности по y функций $\Delta_k(y)$, $G(y)$ и (5.8.5) справедливы неравенства

$$p_{1n}(\theta) \leq q_n(\eta_{jn}, \hat{U}_{jn}) + n^{-1/2} \sum_{k=1}^n u_{k-1} I(u_{k-1} > 0) \times \\ \times \left(G(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - G(n^{-1/2} \eta_{jn} \tilde{u}_{k-1,j}) \right), \\ p_{1n}(\theta) \geq q_n(\eta_{jn}, \tilde{U}_{jn}) - n^{-1/2} \sum_{k=1}^n u_{k-1} I(u_{k-1} > 0) \times \\ \times \left(G(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - G(n^{-1/2} \eta_{jn} \tilde{u}_{k-1,j}) \right).$$

Следовательно,

$$\sup_{|\theta| \leq \Theta} |p_{1n}(\theta)| \leq \\ \leq \sup_{0 \leq j \leq 3^{m_n}} \{ |q_n(\eta_{jn}, \hat{U}_{jn})| + |q_n(\eta_{jn}, \tilde{U}_{jn})| \} + \quad (5.8.6) \\ + \sup_{0 \leq j \leq 3^{m_n}} n^{-1/2} \sum_{k=1}^n u_{k-1} I(u_{k-1} > 0) \left(G(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - \right. \\ \left. - G(n^{-1/2} \eta_{jn} \tilde{u}_{k-1,j}) \right). \quad (5.8.7)$$

С помощью формулы Тейлора, условий 5.5.1, 5.5.2 и определений (5.8.3), (5.8.4) убеждаемся, что (5.8.7) есть $O_p(3^{-m_n}) = o_p(1)$. Рассмотрим в (5.8.6) слагаемое $q_n(\eta_{jn}, \hat{U}_{jn})$. Положим для краткости

$$\xi_k = \xi_k(j, n) = u_{k-1} I(u_{k-1} > 0) \left(\Delta_k(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - \right. \\ \left. - G(n^{-1/2} \eta_{jn} \hat{u}_{k-1,j}) - \Delta_k(0) + G(0) \right).$$

Пусть $\Omega_{\leq i}$ — сигма-алгебра, порожденная $\{\varepsilon_s, s \leq i\}$. Очевидно,

$$\mathbf{E}_\beta \xi_k = \mathbf{E} \{ \mathbf{E}(\xi_k | \Omega_{\leq k-1}) \} = 0,$$

так как $\mathbf{E}_\beta \{ \xi_k | \Omega_{\leq k-1} \} = 0$ п.н.

Аналогично, для $k < j$

$$\mathbf{E}_\beta \xi_k \xi_j = \mathbf{E}_\beta \{ \mathbf{E}_\beta(\xi_k \xi_j | \Omega_{\leq j-1}) \} = \mathbf{E}_\beta \{ \xi_k \mathbf{E}_\beta(\xi_j | \Omega_{\leq j-1}) \} = 0.$$

Кроме того, $\forall x_1, x_2 \in R^1$ и $\delta = \min(\Delta, 1)$

$$\mathbf{E} \{ \Delta_k(x_1) - G(x_1) - \Delta_k(x_2) + G(x_2) \}^2 \leq |G(x_1) - G(x_2)| \leq |G(x_1) - G(x_2)|^\delta.$$

Отсюда

$$\mathbf{E}_\beta \xi_k^2 \leq \mathbf{E}_\beta u_{k-1}^2 |G(n^{-1/2} \eta_{jn} \hat{u}_{k-1,n}) - G(0)|^\delta = O(n^{-\delta/2})$$

равномерно по k, j в силу условий 5.5.1, 5.5.2. Отсюда, так как

$$q_n(\eta_{jn}, \hat{U}_{jn}) = n^{-1/2} \sum_{k=1}^n \xi_k,$$

$$\sup_{0 \leq j \leq 3^{m_n}} \mathbf{E}_\beta q_n^2(\eta_{jn}, \hat{U}_{jn}) = O(n^{-\delta/2}). \quad (5.8.8)$$

Теперь из (5.8.8) $\forall \varepsilon > 0$ в силу неравенства Чебышева получаем

$$\begin{aligned} \mathbf{P}_\beta \left\{ \sup_{0 \leq j \leq 3^{m_n}} |q_n(\eta_{jn}, \hat{U}_{jn})| > \varepsilon \right\} &\leq \sum_{j=0}^{3^{m_n}} \mathbf{P}_\beta \{ |q_n(\eta_{jn}, \hat{U}_{jn})| > \varepsilon \} \leq \\ &\leq \varepsilon^{-2} \sum_{j=0}^{3^{m_n}} \mathbf{E}_\beta q_n^2(\eta_{jn}, \hat{U}_{jn}) = O(3^{m_n} n^{-\delta/2}) = O(n^{-\delta/4}) = o(1) \end{aligned}$$

в силу выбора m_n . Аналогично показывается, что

$$\sup_{0 \leq j \leq 3^{m_n}} |q_n(\eta_{jn}, \tilde{U}_{jn})| = o_p(1).$$

Следовательно, (5.8.6) есть $o_p(1)$. Значит,

$$\sup_{|\theta| \leq \Theta} |p_{1n}(\theta)| = o_p(1).$$

Аналогично показывается, что

$$\sup_{|\theta| \leq \Theta} |p_{2n}(\theta)| = o_p(1).$$

Лемма 5.8.1 доказана. \square

Л е м м а 5.8.2. При условиях 5.5.1, 5.5.2

$$z_{1n}(\theta) = g(0) \theta E \varepsilon_1^2 (1 - \beta^2)^{-1} + \varepsilon_n(\theta),$$

где $\sup_{|\theta| \leq \theta} |\varepsilon_n(\theta)| = o_p(1)$.

Утверждение леммы 5.8.2 следует из формулы Тейлора и закона больших чисел. Утверждение теоремы 5.8.1 следует теперь из (5.8.1) и лемм 5.8.1 и 5.8.2. \square

Д о к а з а т е л ь с т в о т е о р е м ы 5.5.2. В силу теоремы 5.5.1 при достаточно большом $A > 0$ $I_n^{LD}(\beta - n^{-1/2}A) > 0$ с вероятностью, сколь угодно близкой к единице, равномерно по $n > n_0$ и, аналогично, $I_n^{LD}(\beta + n^{-1/2}A) < 0$. Это влечет в силу монотонности $I_n^{LD}(\theta)$: $\hat{\beta}_{n,LD} \in (\beta - n^{-1/2}A, \beta + n^{-1/2}A)$ с вероятностью, сколь угодно близкой к единице, равномерно по $n > n_0$. Следовательно,

$$\sqrt{n}(\hat{\beta}_{n,LD} - \beta) = O_p(1).$$

Теперь отметим, что $n^{-1/2} I_n^{LD}(\hat{\beta}_{n,LD}) = o_p(1)$. Действительно, скачки функции $I_n^{LD}(\theta)$ не больше

$$\max_{1 \leq k \leq n} |u_{k-1}| \leq \max_{1 \leq k \leq n-1} |\varepsilon_k| (1 - |\beta|)^{-1} + \beta^k |u_0|.$$

Так как $n^{-1/2} \max_{1 \leq k \leq n} |\varepsilon_k| = o_p(1)$ при $E \varepsilon_1^2 < \infty$, то $n^{-1/2} \max_{1 \leq k \leq n} |u_{k-1}| = o_p(1)$.

Теперь остается положить в (5.5.4) $\hat{\beta}_n = \hat{\beta}_{n,LD}$, разрешить (5.5.4) относительно $\sqrt{n}(\hat{\beta}_{n,LD} - \beta)$ и воспользоваться тем, что $n^{-1/2} I_n^{LD}(\beta) \xrightarrow{d} N(0, E \varepsilon_1^2)$. Теорема 5.5.2 доказана. \square

Глава 6

ЗНАКОВЫЙ АНАЛИЗ ОДНОПАРАМЕТРИЧЕСКОЙ АВТОРЕГРЕССИИ

§ 6.1. Введение в знаковый авторегрессионный анализ

В этой главе развит непараметрический знаковый анализ простейшей однопараметрической схемы авторегрессии

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i \in \mathbb{Z}, \quad (6.1.1)$$

в которой $|\beta| < 1$, β — неизвестный параметр; $\{\varepsilon_i\}$ — н.о.р.сл.в. с неизвестной невырожденной функцией распределения $G(x)$. Относительно $G(x)$ на разных этапах исследования предполагаются выполненными некоторые из следующих условий.

Условие 6.1.1. $P\{\varepsilon_1 > 0\} = P\{\varepsilon_1 < 0\} = 1/2$.

Условие 6.1.2. $E\varepsilon_1 = 0$, $E|\varepsilon_1|^{1+\Delta} < \infty$ при некотором $\Delta > 0$.

Условие 6.1.3. Существует плотность $g(x) = G'(x)$ и $\sup_x g(x) < \infty$, $g(0) > 0$, $g(x)$ удовлетворяет условию Липшица в нуле.

Как мы уже отмечали в § 5.2 (лемма 5.2.2), при условии $E \log^+ |\varepsilon_1| < \infty$ существует п.н. единственное строго стационарное решение уравнения (6.1.1) вида

$$u_i = \sum_{j=0}^{\infty} \beta^j \varepsilon_{i-j}, \quad (6.1.2)$$

ряд в (6.1.2) сходится абсолютно с вероятностью 1. Если $E|\varepsilon_1|^{1+\Delta} < \infty$, то (6.1.2) сходится и в $L^{1+\Delta}$ (лемма 5.2.1).

Пусть наблюдаются u_0, u_1, \dots, u_n , удовлетворяющие (6.1.2). Наша цель — построить по этим наблюдениям непараметрические знаковые тесты для проверки гипотез относительно параметра β и непараметрические знаковые точечные и доверительные оценки β .

Мы начинаем в § 6.2 с проверки гипотезы

$$H_0: \beta = \beta_0 \quad (6.1.3)$$

против односторонних и двусторонних альтернатив. Чтобы построить непараметрический тест, мы сначала строим новую последовательность наблюдений

$$S_k(\theta) = \text{sign}(u_k - \theta u_{k-1}), \quad \theta \in \mathbb{R}^1, \quad k = 1, \dots, n, \quad (6.1.4)$$

которые объединяем в вектор

$$S^n(\theta) = (S_1(\theta), \dots, S_n(\theta)). \quad (6.1.5)$$

При гипотезе H_0 и условии 6.1.1 $\{S_k(\beta_0) = \text{sign } \varepsilon_k, k = 1, \dots, n\}$ — н.о.р.сл.в.,

$$P_{\beta_0}\{S_k(\beta_0) = -1\} = P_{\beta_0}\{S_k(\beta_0) = 1\} = 1/2,$$

т.е. статистика $S^n(\beta_0)$ свободна от $G(x)$, и гипотеза H_0 для вектора $S^n(\beta_0)$ является простой. Мы находим разложение вероятности $P_{\beta_0}\{S^n(\beta_0) = s_n\}$ при $\beta \rightarrow \beta_0$, где s_n — реализация $S^n(\beta_0)$. С помощью этого разложения мы строим локально наиболее мощные (ЛНМ) против односторонних альтернатив тесты, основанные на $S^n(\beta_0)$.

Статистика этих тестов имеет вид

$$I_n^S(\beta_0) = \sum_{t=1}^{n-1} \beta_0^{t-1} \sum_{k=t+1}^n S_{k-t}(\beta_0) S_k(\beta_0)$$

и свободна от $G(x)$ при H_0 и условии 6.1.1; кроме того,

$$n^{-1/2} I_n^S(\beta_0) \xrightarrow{d_{\beta_0}} N(0, (1 - \beta_0^2)^{-1}).$$

Здесь и далее, как и в гл. 5, ради краткости символ $n \rightarrow \infty$ опущен.

В § 6.3 мы строим знаковые тесты в нестационарной авторегрессии

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i = 1, 2, \dots, \quad u_0 = 0,$$

в которой неизвестный параметр $\beta \in \mathbb{R}^1$. Для такой схемы при конечных n удается получить результаты, аналогичные стационарному случаю из § 6.2. Асимптотические же результаты существенно зависят от величины β и схожи с изложенными в § 5.4.

Далее и до конца главы мы возвращаемся к рассмотрению стационарной схемы (6.1.1).

Факт построения ЛНМ знаковых тестов, основанных на $S^n(\beta_0)$, интересен, но более важен вопрос об эффективности построенных

тестов относительно известных. В § 6.4 мы рассматриваем последовательность близких альтернатив

$$H_{1n}(a): \beta = \beta_n := \beta_0 + an^{-1/2} + o(n^{-1/2}), \quad a - \text{постоянная. (6.1.6)}$$

В силу условия 6.1.2 дисперсия наблюдений u_0, \dots, u_n , вообще говоря, бесконечна, и для этой последовательности не выполнены достаточные условия ЛАН (см. п. 5.3.2). Поэтому мы не можем использовать для отыскания предельного распределения статистики $n^{-1/2} l_n^S(\beta_0)$ при $H_{1n}(a)$ стандартный путь — третью лемму Ле Кама. Конечно, можно пытаться установить справедливость условия ЛАН непосредственно для последовательности новых наблюдений $S^n(\beta_0)$ в точке $\beta = \beta_0$ и дальше использовать третью лемму Ле Кама. Мы, однако, отказываемся от такого способа исследования. Во-первых, условие ЛАН для $S^n(\beta_0)$ в точке $\beta = \beta_0$ удастся установить лишь при более жестких моментных условиях, чем условие 6.1.2. И второе, более существенное обстоятельство: мы хотим искать оценки параметра, минимизируя тестовую статистику $l_n^S(\theta)$, для чего нам нужны равномерные результаты о поведении $l_n^S(\theta)$ вблизи точки $\theta = \beta$.

Поэтому мы устанавливаем для $n^{-1/2} l_n^S(\theta)$ равномерное линейное стохастическое разложение при $H_{1n}(a)$ и условиях 6.1.1–6.1.3:

$$\sup_{|\theta| \leq \Theta_n} |n^{-1/2} l_n^S(\beta_n + n^{-1/2} \theta) - n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0) \theta| = o_p(1),$$

где $\Theta_n \rightarrow \infty$ со степенной скоростью, зависящей от Δ , а

$$\lambda_S(\beta_0) = -2g(0) \mathbf{E}|\varepsilon_1|(1 - \beta_0^2)^{-1}.$$

Это один из основных результатов главы. Из этого равномерного разложения следует, что при $H_{1n}(a)$

$$n^{-1/2} l_n^S(\beta_0) \xrightarrow{d_{\beta_n}} N(-\lambda_S(\beta_0)a, (1 - \beta_0^2)^{-1}),$$

где асимптотический сдвиг (среднее) равен

$$-\lambda_S(\beta_0)a = 2g(0) \mathbf{E}|\varepsilon_1|(1 - \beta_0^2)^{-1}a.$$

Здесь уместно отметить, что при $\mathbf{E}\varepsilon_1^2 < \infty$ и иных условиях, обеспечивающих ЛАН для последовательности u_0, \dots, u_n , гипотезы H_0 и $H_{1n}(a)$ контигуальны. В этом (частном) случае асимптотический сдвиг можно вычислить с помощью третьей леммы Ле Кама.

Зная асимптотический сдвиг легко найти асимптотическую относительную эффективность (АОЭ) знакового теста относительно других известных тестов (см. § 6.5), в частности, относительно теста, основанного на $\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0)$, где $\hat{\beta}_{n,LS}$ — оценка наименьших квадратов. Эта эффективность $e_{S,LS}$ оказывается равной

$$e_{S,LS} = (2g(0) \mathbf{E}|\varepsilon_1|)^2.$$

Как уже говорилось, найденное для $n^{-1/2} l_n^S(\beta_n + n^{-1/2}\theta)$ стохастическое разложение — это не только способ отыскания распределения тестов при близких альтернативах. С его помощью мы исследуем непараметрические знаковые оценки β (§ 6.6). Мы строим несколько таких оценок. Все они имеют одно и то же предельное распределение. Остановимся на одной оценке, $\hat{\beta}_{n,S}$. Сначала мы устанавливаем равномерную по $|\theta| \leq 1 - \delta$ сходимость $n^{-1} l_n^S(\theta)$ по вероятности к функции $\Lambda_S(\theta)$ такой, что $\Lambda_S(\beta) = 0$, $\Lambda'(\beta) = \lambda_S(\beta)$. Функция $\lambda_S(\beta)$ уже встречалась в стохастическом разложении. Это делает естественным определение оценки β как корня уравнения

$$l_n^S(\theta) \div 0.$$

Стохастическое разложение для $n^{-1/2} l_n^S(\beta + n^{-1/2}\theta)$ позволяет установить существование для этого уравнения такого решения $\hat{\beta}_{n,S}$, для которого

$$\sqrt{n}(\hat{\beta}_{n,S} - \beta) \xrightarrow{d} N(0, (1 - \beta^2)(2g(0)E|\epsilon_1|)^{-2}).$$

АОЭ знаковой оценки $\hat{\beta}_{n,S}$ относительно $\hat{\beta}_{n,LS}$ оказывается равной $e_{S,LS}$, т.е. такой, как у соответствующих тестов.

В § 6.7 мы находим функционал влияния $\hat{\beta}_{n,S}$ в случае независимых выбросов и показываем, что чувствительность оценки к большим выбросам конечна на некотором естественном классе засорений. Качественно это означает, что знаковая оценка устойчива к одиночным выбросам в наблюдениях u_0, \dots, u_n .

§ 6.8 мы посвящаем обсуждению численных результатов.

Выделим теперь причины, в силу которых использование анонсированных знаковых процедур целесообразно. Некоторые из них мы уже упоминали выше. Будем для определенности говорить о стационарной схеме. Итак, первая причина связана с условием 6.1.2, в силу которого условие ЛАН для u_0, \dots, u_n , вообще говоря, не выполнено и стандартные процедуры не имеют известного предельного распределения (см. п. 5.3.2, 5.5.1). В то же время наши знаковые процедуры асимптотически нормальны при $E|\epsilon_1|^{1+\Delta} < \infty$. Далее, знаковые тесты свободны от $G(x)$ при конечных n , что позволяет проверять гипотезы и строить доверительные интервалы для умеренных выборок. Это важнейшее свойство знаковых процедур. Классические процедуры свободны от $G(x)$ лишь асимптотически. В третьих, для $G(x)$ с тяжелыми хвостами АОЭ $e_{S,LS}$ может быть больше единицы и вообще сколь угодно большой. Мы покажем также, что она ограничена снизу на естественном классе распределений $G(x)$. В четвертых, функционал влияния оценки $\hat{\beta}_{n,S}$ ограничен, в отличие от общепотребительных оценок $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$. Удобно также, что функция $l_n^S(\theta)$ инвариантна относительно параметра масштаба $G(x)$, и параметр β можно оценивать независимо от этого параметра масштаба. В сущности единственная серьезная претензия к знаковым про-

процедурам — относительно невысокая АОЭ для $G(x)$ со стандартными, типа гауссовских, хвостами. Хорошей альтернативой знаковым процедурам служат ранговые. Мы говорим об этом далее кратко, подробное обсуждение ранговых процедур в настоящей книге не предусмотрено. Впрочем, у знаковых процедур перед ранговыми есть серьезное преимущество: во многих случаях они применимы и для независимых, но разнораспределенных величин $\{\varepsilon_i\}$, удовлетворяющих условию $P\{\varepsilon_i < 0\} = P\{\varepsilon_i > 0\} = 1/2$. Некоторые из перечисленных свойств остаются верными и для нестационарной авторегрессии, например свобода знаковых тестов от $G(x)$ при конечных n .

Завершим параграф замечанием по поводу предположения $E\varepsilon_1 = 0$ из условия 6.1.2. В приложениях условие $E\varepsilon_1 = 0$ может быть слишком ограничительным. В качестве простейшей альтернативы рассмотрим авторегрессионную модель

$$v_i = \beta v_{i-1} + \delta_i, \quad i \in \mathbb{Z}, \quad |\beta| < 1,$$

в которой $\delta_i = \varepsilon_i + \nu$, постоянная ν неизвестна. Такую модель удобно переписать в виде

$$v_i = \mu + u_i, \quad i \in \mathbb{Z}, \quad (6.1.7)$$

где $\mu = (1 - \beta)^{-1}\nu$, а последовательность $\{u_i\}$ удовлетворяет (6.1.1). Пусть v_0, \dots, v_n — наблюдения из строго стационарного решения (6.1.7), а $\hat{\mu}_n$ — построенная по этим наблюдениям оценка μ . Найдем оценки ненаблюдаемых величин u_0, \dots, u_n : $\hat{u}_k = v_k - \hat{\mu}_n$, $k = 0, \dots, n$. По этим оценкам естественно строить знаковые статистики аналогично тому, как мы строим их в схеме (6.1.1) по u_0, \dots, u_n и использовать их для проверки H_0 и оценивания β . Каковы свойства таких процедур? Оказывается, если $\sqrt{n}(\hat{\mu}_n - \mu) = O_p(1)$ и выполнены условия 6.1.1–6.1.3, то асимптотические свойства те же самые, что и описанные в § 6.2, § 6.4–6.6 свойства знаковых процедур в стационарной авторегрессии (6.1.1) с нулевым средним. Это вполне ожидаемый результат, аналогичное свойство процедур наименьших квадратов хорошо известно (см. [2, 45, п. 5.5.5]). К сожалению, важные свойства локальной оптимальности знаковых тестов и их свободы при гипотезе для конечных объемов наблюдений теряются. В случае засоренных выборок устойчивость знаковых оценок также зависит от устойчивости $\hat{\mu}_n$.

Мы опускаем детальное рассмотрение схемы авторегрессии с ненулевым средним, так как идейно оно мало чем отличается от приведенного далее подробного исследования авторегрессии с нулевым средним.

§ 6.2. Знаковые тесты

Начнем с проверки гипотезы (6.1.3) $H_0: \beta = \beta_0$ в схеме (6.1.1) по наблюдениям u_0, u_1, \dots, u_n , удовлетворяющим (6.1.2). Мы построим локально наиболее мощный тест для проверки H_0 против односторонних, скажем, правосторонних альтернатив

$$H_1^+: \beta > \beta_0, \quad (6.2.1)$$

основанный на векторе знаков $S^n(\beta_0)$, определенном в (6.1.4)–(6.1.5).

С вероятностью 1 реализациями $S^n(\beta_0)$ служат последовательности длины n , составленные из $+1$ и -1 . Пусть $s^n = (s_1, \dots, s_n)$ — произвольная последовательность такого вида. Положим

$$\gamma_{tn}(s^n) = \sum_{k=t+1}^n s_{k-t} s_k, \quad t = 1, \dots, n-1. \quad (6.2.2)$$

Теорема 6.2.1. При условиях 6.1.1–6.1.3 и $n \geq 2$

$$\begin{aligned} P_{\beta}\{S^n(\beta_0) = s^n\} &= \\ &= \left(\frac{1}{2}\right)^n \left(1 + 2g(0)E|\varepsilon_1| \sum_{t=1}^{n-1} \beta_0^{t-1} \gamma_{tn}(s^n)(\beta - \beta_0)\right) + o(\beta - \beta_0), \quad \beta \rightarrow \beta_0. \end{aligned}$$

Доказательство. Обозначим через $\Omega_{\leq 0}$ сигма-алгебру, порожденную случайными величинами $\{\varepsilon_j, j \leq i\}$, $I(\cdot)$ — индикатор события. Тогда для $n = 1$

$$\begin{aligned} P_{\beta}\{S_1(\beta_0) = s_1\} &= \\ &= E_{\beta}\left\{I(S_1(\beta_0) = 1) \frac{1+s_1}{2} + I(S_1(\beta_0) = -1) \frac{1-s_1}{2}\right\} = \\ &= \frac{1+s_1}{2} - s_1 E_{\beta} I(S_1(\beta_0) = -1) = \\ &= \frac{1+s_1}{2} - s_1 E_{\beta} I(u_1 - \beta_0 u_0 < 0) = \\ &= \frac{1+s_1}{2} - s_1 E_{\beta} I(\varepsilon_1 < (\beta_0 - \beta)u_0) = \\ &= \frac{1+s_1}{2} - s_1 E_{\beta} \left\{E_{\beta} I(\varepsilon_1 < (\beta_0 - \beta)u_0) \Big| \Omega_{\leq 0}\right\} = \\ &= \frac{1+s_1}{2} - s_1 E_{\beta} G((\beta_0 - \beta)u_0), \end{aligned}$$

так как ε_1 и $\Omega_{\leq 0}$ независимы. Итак,

$$P_{\beta}\{S_1(\beta_0) = s_1\} = \frac{1+s_1}{2} - s_1 E_{\beta} G((\beta_0 - \beta)u_0). \quad (6.2.3)$$

В силу условия 6.1.3 $g(x)$ удовлетворяет в нуле условию Гёльдера произвольного порядка $0 < \delta \leq 1$ с некоторой постоянной $h = h(\delta)$. Положим $\delta = \min(\Delta, 1)$. Тогда

$$G(x) = G(0) + g(0)x + \varepsilon(x),$$

где $|\varepsilon(x)| \leq h|x|^{1+\delta}$, h не зависит от x . Отсюда

$$\begin{aligned} \mathbf{E}_\beta G((\beta_0 - \beta)u_0) &= \\ &= \mathbf{E} \left\{ \frac{1}{2} - g(0)(\beta - \beta_0)u_0 + \varepsilon((\beta_0 - \beta)u_0) \right\} = \frac{1}{2} + o(\beta - \beta_0), \end{aligned} \quad (6.2.4)$$

так как $G(0) = 1/2$, $\mathbf{E}_\beta u_0 = 0$, $\mathbf{E}_\beta |u_0|^{1+\delta} < \infty$ в силу условий 6.1.1, 6.1.2. Из (6.2.3)–(6.2.4) получаем

$$\mathbf{P}_\beta \{S_1(\beta_0) = s_1\} = \frac{1}{2} + o(\beta - \beta_0). \quad (6.2.5)$$

Аналогичные, хотя и более громоздкие вычисления приводят к соотношению для $n \geq 2$:

$$\begin{aligned} \mathbf{P}_\beta \{S^n(\beta_0) = s^n\} &= \frac{1}{2} \mathbf{P}_\beta \{S^{n-1}(\beta_0) = s^{n-1}\} + \\ &+ \left(\frac{1}{2}\right)^n 2g(0) \mathbf{E}|\varepsilon_1| \sum_{t=1}^{n-1} \beta_0^{t-1} s_n s_{n-t} (\beta - \beta_0) + o(\beta - \beta_0). \end{aligned} \quad (6.2.6)$$

Рекуррентное соотношение (6.2.6) при начальном условии (6.2.5) влечет утверждение теоремы 6.2.1. \square

Следующая теорема дает статистический тест для проверки (6.1.3) против (6.2.1). Она непосредственно вытекает из теоремы 6.2.1.

Теорема 6.2.2. Пусть выполнены условия 6.1.1–6.1.3.

Тогда локально наиболее мощный тест для проверки гипотезы $H_0: \beta = \beta_0$ против альтернативы $H_1^+: \beta > \beta_0$, основанный на векторе знаков $S^n(\beta_0)$, имеет при $n \geq 2$ критическую область

$$\mathbb{Q}_n^+ = \left\{ s^n: \sum_{t=1}^{n-1} \beta_0^{t-1} \gamma_{tn}(s^n) > \text{const} \right\}. \quad (6.2.7)$$

Для левосторонних альтернатив локально наиболее мощный тест для проверки H_0 выглядит аналогично и отличается от (6.2.7) лишь знаком неравенства. В случае двусторонних альтернатив $H_1: \beta \neq \beta_0$ мы предлагаем использовать двустороннюю версию (6.2.7) с критической областью

$$\mathbb{Q}_n = \left\{ s^n: \left| \sum_{t=1}^{n-1} \beta_0^{t-1} \gamma_{tn}(s^n) \right| > \text{const} \right\}. \quad (6.2.8)$$

Для $S_k(\theta)$, $k = 1, \dots, n$, из (6.1.4) положим

$$\Gamma_{tn}(\theta) = \sum_{k=t+1}^n S_{k-t}(\theta) S_k(\theta), \quad t = 1, \dots, n-1, \quad (6.2.9)$$

$$l_n^S(\theta) = \sum_{t=1}^{n-1} \theta^{t-1} \Gamma_{tn}(\theta). \quad (6.2.10)$$

Тогда $l_n^S(\beta_0)$ есть статистика тестов (6.2.7) и (6.2.8). При гипотезе H_0 статистики $\Gamma_{tn}(\beta_0)$ превращаются в статистики

$$\Gamma_{tn}(\beta_0) = \sum_{k=t+1}^n \text{sign}(\varepsilon_{k-t} \varepsilon_k), \quad (6.2.11)$$

распределенные для любого n свободно от $G(x)$ при условии 6.1.1. Следовательно, в этом случае и тестовая статистика $l_n^S(\beta_0)$ также свободна от $G(x)$ при любом n . Хотя распределение $l_n^S(\beta_0)$ зависит от β_0 , оно может быть оценено методом статистических испытаний. Таким образом, критические значения в (6.2.7) и (6.2.8) могут быть вычислены.

Мы приводим некоторые результаты таких вычислений в § 6.8. Там, в частности, для $n = 50(50), 200$, $\beta_0 = 0, \pm 0.1, \pm 0.2, \dots, \pm 0.9$ и $\alpha = 0.05, 0.95$ приведена таблица квантилей $c_n^\alpha(\beta_0)$ уровня α нормированной статистики

$$T_{n,S}(\beta_0) = n^{-1/2} \sqrt{1 - \beta_0^2} l_n^S(\beta_0).$$

(Статистика $T_{n,S}(\beta_0)$, как следует из приведенных далее соотношений (6.2.14), имеет при H_0 асимптотическую единичную дисперсию и нулевое среднее.) Кроме того, множество

$$\mathbb{D}_{n\alpha} = \left\{ \theta: c_n^\alpha(\theta) \leq T_{n,S}(\theta) \leq c_n^{1-\alpha}(\theta) \right\} \quad (6.2.12)$$

является при каждом n доверительным множеством для β надежности не меньше $1 - 2\alpha$. Функции квантилей $c_n^\alpha(\theta)$ и $c_n^{1-\alpha}(\theta)$ могут быть оценены на сетке значений θ . Строя одновременно кривые $c_n^\alpha(\theta)$, $c_n^{1-\alpha}(\theta)$ и $T_{n,S}(\theta)$, можно построить явно множество $\mathbb{D}_{n\alpha}$, и мы делаем это в примерах в § 6.8.

В важном частном случае гипотезы $H_0: \beta = 0$, соответствующем предположению о независимости наблюдений, тестовая статистика $l_n^S(0)$ есть

$$\Gamma_{1n}(0) = \sum_{k=2}^n \text{sign}(u_{k-1} u_k). \quad (6.2.13)$$

При $H_0: \beta = 0$ и условии 6.1.1 распределение статистики (6.2.13) сводится к биномиальному, ибо по биномиальному закону $\text{Bi}(n-1, 1/2)$ распределена случайная переменная

$$\frac{1}{2} \sum_{k=2}^n (1 + \text{sign}(\varepsilon_{k-1}\varepsilon_k)) = \frac{1}{2}(n-1 + \Gamma_{1n}(0)).$$

Чтобы в этом убедиться, достаточно проверить (индукцией по n), что случайные величины $\text{sign}(\varepsilon_1\varepsilon_2), \text{sign}(\varepsilon_2\varepsilon_3), \dots, \text{sign}(\varepsilon_{n-1}\varepsilon_n)$ при условии 6.1.1 независимы и

$$\mathbf{P}\{\text{sign}(\varepsilon_{k-1}\varepsilon_k) = -1\} = \mathbf{P}\{\text{sign}(\varepsilon_{k-1}\varepsilon_k) = 1\} = \frac{1}{2}.$$

Упоминания о знаковом тесте со статистикой (6.2.13) для проверки независимости уже встречались в литературе. В частности, такой тест был предложен в [61], где было указано его распределение при гипотезе.

При H_0 и условии 6.1.1 распределения случайных переменных $\Gamma_{tn}(\beta_0)$ из (6.2.11) при любом $t = 1, 2, \dots, n-1$ тоже сводятся к биномиальным, а именно,

$$\frac{1}{2}(n-t + \Gamma_{tn}(\beta_0)) \sim \text{Bi}\left(n-t, \frac{1}{2}\right).$$

Кроме того, легко проверить, что при гипотезе H_0 и условии 6.1.1

$$\mathbf{E}\beta_0 \Gamma_{tn}(\beta_0) = 0, \quad \mathbf{E}\beta_0 \Gamma_{tn}(\beta_0) \Gamma_{rn}(\beta_0) = (n-t)\delta_{tr}, \quad (6.2.14)$$

где δ_{tr} — символ Кронекера.

Из этих свойств вытекает следующая теорема, которая описывает поведение тестовой статистики при гипотезе.

Теорема 6.2.3. Пусть выполнено условие 6.1.1.

Тогда при гипотезе $H_0: \beta = \beta_0$

$$n^{-1/2} l_n^S(\beta_0) \xrightarrow{d_{\beta_0}} N(0, (1 - \beta_0^2)^{-1}).$$

Доказательство. Для $1 \leq k < n-1$ имеем

$$l_n^S(\beta_0) = n^{-1/2} \sum_{t=1}^{n-1} \beta_0^{t-1} \Gamma_{tn}(\beta_0) = Z_{kn} + X_{kn},$$

где

$$Z_{kn} = n^{-1/2} \sum_{t=1}^k \beta_0^{t-1} \Gamma_{tn}(\beta_0),$$

$$X_{kn} = n^{-1/2} \sum_{t=k+1}^{n-1} \beta_0^{t-1} \Gamma_{tn}(\beta_0).$$

Перепишем Z_{kn} в виде

$$Z_{kn} = n^{-1/2} \sum_{j=1}^n a_{jk} + \delta_{kn},$$

где $a_{jk} = \text{sign}(\varepsilon_j \varepsilon_{j+1}) + \beta_0 \text{sign}(\varepsilon_j \varepsilon_{j+2}) + \dots + \beta_0^{k-1} \text{sign}(\varepsilon_j \varepsilon_{j+k})$, а $|\delta_{kn}| \leq (1 + 2 + \dots + k)n^{-1/2}$.

Величины $\{a_{jk}, j = 1, \dots, n\}$ образуют выборку из строго стационарного k -зависимого процесса, что влечет с учетом (6.2.14)

$$Z_{kn} \xrightarrow{d_{\beta_0}} Z_k \sim N\left(0, \frac{1 - \beta_0^{2k}}{1 - \beta_0^2}\right)$$

в силу [2, 45, теорема 7.7.5]. Очевидно,

$$Z_k \xrightarrow{d_{\beta_0}} N(0, (1 - \beta_0^2)^{-1}), \quad k \rightarrow \infty.$$

Осталось заметить, что $\forall \delta, \varepsilon > 0$ и $k > k_0$ в силу (6.2.14)

$$\sup_n \mathbf{P}_{\beta_0} \{|X_{kn}| \geq \delta\} \leq \sup_n \delta^{-2} \mathbf{E}_{\beta_0} X_{kn}^2 \leq \delta^{-2} \sum_{t \geq k+1} \beta_0^{2(t-1)} < \varepsilon,$$

и применить лемму 5.3.1. Теорема доказана. \square

В силу этой теоремы при больших n для вычисления критических значений в (6.2.7)–(6.2.8) можно использовать нормальную аппроксимацию. Кроме того, множество

$$\mathbb{A}_{n\alpha} = \left\{ \theta: |n^{-1/2}(1 - \theta^2)^{1/2} l_n^S(\theta)| < \xi_{1-\alpha} \right\} \quad (6.2.15)$$

является асимптотическим доверительным множеством для неизвестного параметра β уровня $1 - 2\alpha$.

В заключение параграфа отметим простой, но важный факт: если $\{\varepsilon_i\}$ — разнораспределенные случайные величины, но

$$\mathbf{P}\{\varepsilon_i < 0\} = \mathbf{P}\{\varepsilon_i > 0\} = 1/2, \quad i \in \mathbb{Z}, \quad (6.2.16)$$

то распределение вектора $S^n(\beta_0)$, а значит, и распределения статистик $\Gamma_{tn}(\beta_0)$, $l_n^S(\beta_0)$ при H_0 останутся свободными от $G(x)$ и таки-

ми же, как в случае одинаково распределенных $\{\varepsilon_i\}$. Отсюда следует, что при H_0 и условия (6.2.16):

множество $\mathbb{D}_{n\alpha}$ из (6.1.12) останется доверительным множеством для β надежности не меньше $1 - 2\alpha$;

утверждение теоремы 6.2.3 о предельном распределении статистики $n^{-1/2} t_n^S(\beta_0)$ будет по-прежнему справедливым;

множество $\mathbb{A}_{n\alpha}$ из (6.2.15) останется асимптотическим доверительным множеством для β надежности $1 - 2\alpha$.

Результаты, представленные в этом параграфе, были опубликованы в [56]. Теорема 6.2.1 иным, чем здесь методом, была получена в [34].

§ 6.3. Знаковые тесты в нестационарной авторегрессии

В этом параграфе будем рассматривать нестационарную схему авторегрессии из § 5.4

$$u_i = \beta u_{i-1} + \varepsilon_i, \quad i = 1, 2, \dots, \quad u_0 = 0, \quad (6.3.1)$$

в которой $\{\varepsilon_i\}$ — н.о.р.сл.в. с неизвестной функцией распределения $G(x)$; $\beta \in \mathbb{R}^1$ — неизвестный параметр.

Из (6.3.1) следует, что

$$u_i = \sum_{j=0}^{i-1} \beta^j \varepsilon_{i-j}, \quad i = 1, 2, \dots \quad (6.3.2)$$

Пусть наблюдения u_1, \dots, u_n порождаются соотношением (6.3.1). Цель параграфа — построить и изучить в схеме (6.3.1) локально оптимальные знаковые тесты для проверки гипотезы (6.1.3) $H_0: \beta = \beta_0$.

Замечательно, что нам удастся сделать это при всех возможных β_0 . При конечных n результаты будут вполне аналогичны изложенным в § 6.2. Асимптотические результаты будут схожи с результатами, изложенными в § 5.4.

Итак, пусть вектор знаков $S^n(\beta_0)$ определен в (6.1.4)–(6.1.5); $s^n = (s_1, \dots, s_n)$, $s_n = \pm 1$, есть вектор реализаций $S^n(\beta_0)$, величины $\gamma_{in}(s^n)$ определены в (6.2.2). Следующая теорема показывает, что разложения правдоподобия вектора $S^n(\beta_0)$ в схемах (6.1.1) при $|\beta_0| < 1$ и (6.3.1) при любом β_0 одинаковы.

Теорема 6.3.1. Пусть выполнены условия 6.1.1–6.1.3. Тогда в схеме (6.3.1) при $n \geq 2$ для любого $\beta_0 \in \mathbb{R}^1$

$$\begin{aligned} P_\beta \{ S^n(\beta_0) = s^n \} = \\ = \left(\frac{1}{2}\right)^n \left(1 + 2g(0)E|\varepsilon_1| \sum_{i=1}^{n-1} \beta_0^{i-1} \gamma_{in}(s^n)(\beta - \beta_0) \right) + o(\beta - \beta_0), \quad \beta \rightarrow \beta_0. \end{aligned}$$

Доказательство теоремы 6.3.1 получается с помощью представления (6.3.2) тем же методом, каким доказательство теоремы 6.2.1 получается из представления (6.1.2). Поэтому мы его опускаем.

Непосредственно из теоремы 6.3.1 следует утверждение теоремы 6.3.2.

Теорема 6.3.2. Пусть выполнены условия 6.1.1–6.1.3.

Тогда для любого $\beta_0 \in \mathbb{R}^1$ в схеме (6.3.1) при $n \geq 2$ тест χ критической областью

$$Q_n^+ = \left\{ s^n : \sum_{t=1}^{n-1} \beta_0^{t-1} \gamma_{tn}(s^n) > \text{const} \right\} \quad (6.3.3)$$

является локально наиболее мощным для проверки гипотезы $H_0: \beta = \beta_0$ против альтернативы $H_1^+: \beta > \beta_0$ среди всех тестов, основанных на знаках $S^n(\beta_0)$.

Разумеется, если в (6.3.3) поменять знак неравенства на противоположный, то получим критическую область локально наиболее мощного (ЛНМ) знакового теста для проверки H_0 против левосторонней альтернативы $H_0^-: \beta < \beta_0$.

Пусть статистика $l_n^S(\theta)$, $\theta \in \mathbb{R}^1$, определена соотношениями (6.2.10) и (6.2.9). Тогда $l_n^S(\beta_0)$ есть тестовая статистика для (6.3.3). Как и в предыдущем параграфе, она свободна от $G(x)$ при гипотезе H_0 и единственном условии 6.1.1 при любом $\beta_0 \in \mathbb{R}^1$.

Квантили $l_n^S(\beta_0)$ обозначим $\tilde{c}_n^\alpha(\beta_0)$, их можно оценить методом Монте-Карло при любом $\beta_0 \in \mathbb{R}^1$. Множество

$$\tilde{D}_{n\alpha} = \{ \theta : \tilde{c}_n^\alpha(\beta_0) \leq l_n^S(\theta) \leq \tilde{c}_n^{1-\alpha}(\beta_0) \}$$

есть доверительное множество для β надежности не меньше $1 - 2\alpha$. Построить $\tilde{D}_{n\alpha}$ явно можно так же, как в стационарном случае (см. § 6.8).

Итак, ситуация при конечных n (применительно к знаковым тестам) вполне аналогична стационарному случаю. Интереснее вопрос об асимптотическом распределении $l_n^S(\beta_0)$ при H_0 и условии 6.1.1. Прежде всего заметим, что если $I_n(\beta, \theta)$ — информация Фишера о параметре β , содержащаяся в $S^n(\theta)$, то в силу теоремы 6.3.1

$$I_n(\beta, \beta) = (2g(0)E|\varepsilon_1|)^2 E_\beta (l_n^S(\beta))^2.$$

С помощью (6.2.14) находим

$$E_\beta (l_n^S(\beta))^2 = \sum_{t=1}^{n-1} \beta^{2(t-1)} (n-t) = \begin{cases} \frac{1}{1-\beta^2} \left(n - \frac{1-\beta^{2n}}{1-\beta^2} \right), & |\beta| \neq 1, \\ \frac{1}{2} n(n-1), & |\beta| = 1. \end{cases}$$

Следовательно, $I_n(\beta, \beta) \sim (2g(0)E|\varepsilon_1|)^2 d_n^2(\beta)$, где $d_n^2(\beta)$ определено соотношением (5.4.4), т.е.

$$d_n^2(\beta) = \begin{cases} n/(1-\beta^2), & |\beta| < 1, \\ n^2/2, & |\beta| = 1, \\ \beta^{2n}/(1-\beta^2)^2, & |\beta| > 1. \end{cases}$$

Сравнивая знаковую информацию $I_n(\beta, \beta)$ с информацией $J_n(\beta)$ из § 5.4 (см. соотношения (5.4.3)–(5.4.4)), видим, что они различаются при каждом n постоянной $(2g(0)E|\varepsilon_1|)^2$.

Обозначим $R(\beta)$ распределение случайной величины $\beta \frac{1}{\sqrt{2}}(\zeta^2 - 1)$, где $\zeta \sim N(0, 1)$. При $|\beta| > 1$ обозначим также $L(\beta)$ распределение случайной величины $\frac{\beta^2 - 1}{\beta^2} \xi \eta$, где ξ и η — н.о.р.сл.в.,

$$\xi = \sum_{i \geq 1} \beta^{-(i-1)} a_i,$$

$\{a_i, i = 1, 2, \dots\}$ — последовательность н.о.р.сл.в., $P\{a_i = \pm 1\} = 1/2$. Следующая теорема описывает асимптотическое нулевое распределение тестовой статистики.

Теорема 6.3.3. Пусть верна гипотеза $H_0: \beta = \beta_0$.

Тогда при условии 6.1.1 в схеме (6.3.1)

$$d_n^{-1}(\beta_0) I_n^S(\beta_0) \xrightarrow{d_{\beta_0}} \begin{cases} N(0, 1), & |\beta_0| < 1, \\ R(\beta_0), & |\beta_0| = 1, \\ L(\beta_0), & |\beta_0| > 1. \end{cases}$$

Теорема 6.3.3 аналогична теореме 5.4.1. В частности, при $|\beta_0| > 1$ слабый предел $d_n^{-1}(\beta_0) I_n^S(\beta_0)$ есть произведение двух независимых, одинаково распределенных (н.о.р.) финитных сомножителей. Аналогично числитель U_n оценки наименьших квадратов (см. доказательство теоремы 5.4.1) представлялся произведением двух н.о.р. гауссовских величин. Разумеется, утверждение теоремы 6.3.3 останется справедливым и для разнораспределенных $\{\varepsilon_i\}$, удовлетворяющих (6.2.16).

Доказательство. Для случая $|\beta_0| < 1$ утверждение теоремы 6.3.3, очевидно, следует из теоремы 6.2.3.

Пусть $\beta_0 = 1$ (случай $\beta_0 = -1$ аналогичен). Обозначим ради

краткости $a_i = \text{sign } \varepsilon_i$. Тогда при H_0

$$l_n^S(1) = \sum_{i=1}^{n-1} \Gamma_{in}(1) = \sum_{i=2}^n \sum_{j=1}^{i-1} a_j a_i.$$

Поскольку при условии 6.1.1 п.н.

$$\left(n^{-1/2} \sum_{i=1}^n a_i \right)^2 = 1 + 2n^{-1} \sum_{i=1}^n \sum_{j=1}^{i-1} a_j a_i,$$

то с вероятностью 1

$$\frac{1}{\sqrt{2}} \left[\left(n^{-1/2} \sum_{i=1}^n a_i \right)^2 - 1 \right] = \frac{\sqrt{2}}{n} l_n^S(1) = d_n^{-1}(1) l_n^S(1).$$

Очевидно, $n^{-1/2} \sum_{i=1}^n a_i \xrightarrow{d} N(0, 1)$, откуда $d_n^{-1}(1) l_n^S(1) \xrightarrow{d} R(1)$.

Пусть теперь $\beta_0 > 1$ (случай $\beta_0 < -1$ аналогичен). Тогда

$$\begin{aligned} d_n^{-1}(\beta_0) l_n^S(\beta_0) &= \frac{\beta_0^2 - 1}{\beta_0^n} \sum_{i=1}^{n-1} \beta_0^{i-1} \Gamma_{in} = \\ &= \frac{\beta_0^2 - 1}{\beta_0^2} (\Gamma_{n-1,n}(\beta_0) + \alpha \Gamma_{n-2,n}(\beta_0) + \dots + \alpha^{n-2} \Gamma_{1n}(\beta_0)), \end{aligned}$$

где $\alpha = \beta_0^{-1}$. Пусть $n+1 > 2k$, k фиксировано. Тогда

$$S_n := \Gamma_{n-1,n}(\beta_0) + \alpha \Gamma_{n-2,n}(\beta_0) + \dots + \alpha^{n-2} \Gamma_{1n}(\beta_0) = Z_{kn} + X_{kn},$$

где

$$\begin{aligned} Z_{kn} &= \Gamma_{n-1,n}(\beta_0) + \dots + \alpha^{k-1} \Gamma_{n-k,n}(\beta_0), \\ X_{kn} &= \alpha^k \Gamma_{n-k-1,n}(\beta_0) + \dots + \alpha^{n-2} \Gamma_{1n}(\beta_0). \end{aligned}$$

В силу (6.2.14) при H_0 и условии 6.1.1

$$\sup_n \mathbf{E}_{\beta_0} X_{kn}^2 \leq \sum_{i \geq k} (i+1) \alpha^{2i} \rightarrow 0, \quad k \rightarrow \infty. \quad (6.3.4)$$

Рассмотрим Z_{kn} . Имеем при H_0

$$\begin{aligned} Z_{kn} &= a_1 a_n + \alpha(a_1 a_{n-1} + a_2 a_n) + \dots \\ &\quad \dots + \alpha^{k-1} (a_1 a_{n-k+1} + a_2 a_{n-k+2} + \dots + a_k a_n). \end{aligned}$$

Введем последовательность н.о.р.сл.в. $\{b_i, i = 1, 2, \dots\}$, которая не зависит от $\{a_i\}$, $P\{b_i = \pm 1\} = 1/2$. Поскольку a_1, \dots, a_k и a_{n-k+1}, \dots, a_n независимы между собой при $n+1 > 2k$, то Z_{kn} совпадает по распределению с величиной

$$Z_k := a_1 b_1 + \alpha(a_1 b_2 + a_2 b_1) + \dots + \alpha^{k-1}(a_1 b_k + a_2 b_{k-1} + \dots + a_k b_1).$$

Следовательно, при фиксированном k

$$Z_{kn} \xrightarrow{d_{\beta_0}} Z_k. \quad (6.3.5)$$

Очевидно,

$$Z_k \xrightarrow{d_{\beta_0}} (a_1 + \alpha a_2 + \dots)(b_1 + \alpha b_2 + \dots), \quad k \rightarrow \infty. \quad (6.3.6)$$

Соотношения (6.3.4)–(6.3.6) обеспечивают выполнение условий леммы 5.3.1, в силу которой

$$S_n \xrightarrow{d_{\beta_0}} (a_1 + \alpha a_2 + \dots)(b_1 + \alpha b_2 + \dots) = \xi \eta. \quad (6.3.7)$$

Поскольку $d_n^{-1}(\beta_0) l_n^S(\beta_0) = \frac{\beta_0^2 - 1}{\beta_0^2} S_n$, то утверждение теоремы следует из (6.3.7) и определения $L(\beta_0)$. \square

§ 6.4. Теорема о равномерном стохастическом разложении. Мощность знаковых тестов при близких альтернативах

Продолжим рассмотрение стационарной схемы (6.1.1) и перейдем к исследованию асимптотической мощности тестов со статистикой $n^{-1/2} l_n^S(\beta_0)$ (см. определение в (6.2.10) и (6.2.9)) при альтернативах (6.1.6)

$$H_{1n}(a): \beta = \beta_n := \beta_0 + a n^{-1/2} + o(n^{-1/2}),$$

a — постоянная. Стандартный путь для отыскания предельного распределения тестовых статистик при контигуальных альтернативах состоит в использовании третьей леммы Ле Кама. Мы получим соответствующее утверждение для статистики $n^{-1/2} l_n^S(\beta_0)$ иначе — как следствие формулируемой ниже теоремы 6.4.1, которая понадобится нам также в § 6.6 для исследования знаковых оценок. Эта теорема в предположении, что альтернатива $H_{1n}(a)$ верна, описывает линейное стохастическое разложение процесса $n^{-1/2} l_n^S(\beta_n + n^{-1/2} \theta)$, равномерное для $|\theta| \leq \Theta_n$, где $\Theta_n \rightarrow \infty$ со степенной скоростью.

Обозначим

$$\lambda_S(\beta) = -2g(0) E|\varepsilon_1| (1 - \beta^2)^{-1}. \quad (6.4.1)$$

Теорема 6.4.1. Пусть выполнены условия 6.1.1–6.1.3. Тогда при альтернативе $H_{1n}(a)$

$$\sup_{|\theta| \leq \Theta_n} |n^{-1/2} l_n^S(\beta_n + n^{-1/2}\theta) - n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0)\theta| = o_p(1), \quad (6.4.2)$$

где $\Theta_n = \Theta n^\alpha$, $\alpha < \frac{\delta}{2(1+2\delta)}$, $\delta = \min(\Delta, 1)$, $0 < \Theta < \infty$.

Доказательство теоремы 6.4.1 весьма кропотливо, и мы вынесли его в приложение, § 6.9. Для случая $\alpha = 0$, т.е. $\Theta_n = \Theta$, эта теорема была сформулирована и доказана в [56]. Усиление результата для случая $\alpha > 0$ окажется существенным в дальнейшем, в частности в § 6.6.

Из этого доказательства теоремы 6.4.1 в § 6.9 (а именно, из соотношения (6.9.4), леммы 6.9.2 и доказательства леммы 6.9.6) следует, что при фиксированном $t = 1, 2, \dots$ в условиях теоремы 6.4.1

$$\sup_{|\theta| \leq \Theta_n} |n^{-1/2} \Gamma_{tn}(\beta_n + n^{-1/2}\theta) - n^{-1/2} \Gamma_{tn}(\beta_n) - \lambda_S(0)\beta_0^{t-1}\theta| = o_p(1). \quad (6.4.3)$$

Сопоставление (6.4.3) с определениями $l_n^S(\theta)$ и $\lambda_S(\beta)$ соответственно в (6.2.10) и (6.4.1) объясняет вид второго члена разложения (6.4.2).

Приведем здесь же полезное следствие из теоремы 6.4.1.

С л е д с т в и е 6.4.1. Пусть выполнены условия 6.1.1–6.1.3 и $\hat{\beta}_n$ — любая последовательность, для которой $n^{1/2-\alpha}(\hat{\beta}_n - \beta_n) = O_p(1)$, $\alpha < \frac{\delta}{2(1+2\delta)}$, $\delta = \min(\Delta, 1)$.

Тогда при $H_{1n}(a)$

$$n^{-1/2} l_n^S(\hat{\beta}_n) = n^{-1/2} l_n^S(\beta_n) + \lambda_S(\beta_0)\sqrt{n}(\hat{\beta}_n - \beta_n) + o_p(1).$$

Д о к а з а т е л ь с т в о. При любых $\varepsilon > 0$ и $0 < \Theta < \infty$, $\Theta_n = \Theta n^\alpha$ имеем

$$\begin{aligned} & \mathbf{P}_{\beta_n} \left\{ |n^{-1/2} l_n^S(\hat{\beta}_n) - n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0)\sqrt{n}(\hat{\beta}_n - \beta_n)| > \varepsilon \right\} \leq \\ & \leq \mathbf{P}_{\beta_n} \left\{ |n^{-1/2} l_n^S(\beta_n + \sqrt{n}(\hat{\beta}_n - \beta_n)n^{-1/2}) - n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0)\sqrt{n}(\hat{\beta}_n - \beta_n)| > \varepsilon, |\sqrt{n}(\hat{\beta}_n - \beta_n)| \leq \Theta_n \right\} + \\ & \quad + \mathbf{P}_{\beta_n} \left\{ |\sqrt{n}(\hat{\beta}_n - \beta_n)| > \Theta_n \right\} \leq \\ & \leq \mathbf{P}_{\beta_n} \left\{ \sup_{|\theta| \leq \Theta_n} |n^{-1/2} l_n^S(\beta_n + n^{-1/2}\theta) - n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0)\theta| > \varepsilon \right\} + \\ & \quad + \mathbf{P}_{\beta_n} \left\{ |\sqrt{n}(\hat{\beta}_n - \beta_n)| > \Theta_n \right\}. \end{aligned}$$

Обе последние вероятности сколь угодно малы при всех достаточно больших n и достаточно большом Θ : первая в силу теоремы 6.4.1, а вторая в силу условия $n^{1/2-\alpha}(\hat{\beta}_n - \beta_n) = O_p(1)$. \square

Легко проверить, что при условии 6.1.1 асимптотическое распределение статистики $n^{-1/2} l_n^S(\beta_n)$ при альтернативе $H_{1n}(a)$ из (6.1.6) то же самое, что и статистики $n^{-1/2} l_n^S(\beta_0)$ при гипотезе H_0 из (6.1.3). С учетом этого замечания следствие 6.4.1 (с $\hat{\beta}_n = \beta_0$ и $\alpha = 0$) и теорема 6.2.3 влекут следующую теорему.

Т е о р е м а 6.4.2. Пусть выполнены условия 6.1.1–6.1.3.

Тогда при альтернативе $H_{1n}(a)$

$$n^{-1/2} l_n^S(\beta_0) = n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0) a + o_p(1)$$

и, следовательно,

$$n^{-1/2} l_n^S(\beta_0) \xrightarrow{d_{p_n}} N(-\lambda_S(\beta_0) a, (1 - \beta_0^2)^{-1}).$$

Отсюда асимптотический сдвиг нормированной статистики

$$T_{n,S}(\beta_0) = n^{-1/2} \sqrt{1 - \beta_0^2} l_n^S(\beta_0)$$

при $H_{1n}(a)$ равен

$$\delta_S(\beta_0) = 2g(0) \mathbb{E}|\varepsilon_1| (1 - \beta_0^2)^{-1/2} a. \quad (6.4.4)$$

Отсюда асимптотическая мощность соответствующего теста уровня α , например, при альтернативе

$$H_{1n}^+(a): \beta = \beta_0 + n^{-1/2} a + o(n^{-1/2}), \quad a > 0,$$

равна

$$\Phi(\delta_S(\beta_0) - \xi_{1-\alpha}).$$

Разумеется, из (6.4.3) следует, что при фиксированном t в условиях теоремы 6.4.2 статистика $n^{-1/2} \Gamma_{tn}(\beta_0)$ асимптотически нормальна $N(-\lambda_S(0) \beta_0^{t-1} a, 1)$. Таким образом, асимптотический сдвиг статистики $n^{-1/2} \Gamma_{tn}(\beta_0)$ при $H_{1n}(a)$ равен

$$\delta_{tS} = -\lambda_S(0) \beta_0^{t-1} a = 2g(0) \mathbb{E}|\varepsilon_1| \beta_0^{t-1} a.$$

Обозначим $I_n(\beta, \theta)$ информацию Фишера о параметре β , содержа-

щуются в $S^n(\theta)$. В силу теоремы 6.2.1

$$\left. \frac{\partial \log \mathbf{P}_\beta \{S^n(\beta_0) = s^n\}}{\partial \beta} \right|_{\beta=\beta_0} = 2g(0) \mathbf{E}|\varepsilon_1| \sum_{t=1}^{n-1} \beta_0^{t-1} \gamma_{tn}(s^n). \quad (6.4.5)$$

Из (6.4.5) и определения $I_n^S(\beta_0)$ (6.2.10) следует, что

$$I_n(\beta_0, \beta_0) = (2g(0) \mathbf{E}|\varepsilon_1|)^2 \mathbf{E}_{\beta_0} (I_n^S(\beta_0))^2.$$

Заметим, что знаковая информация $I_n(\beta, \beta)$ в стационарной схеме (6.1.1) оказалась равной (в случае $|\beta| < 1$) знаковой информации в нестационарной схеме (6.3.1) (см. § 6.3).

Пользуясь (6.2.14), аналогично § 6.3 получаем

$$n^{-1} I_n(\beta_0, \beta_0) \sim (2g(0) \mathbf{E}|\varepsilon_1|)^2 (1 - \beta_0^2)^{-1}. \quad (6.4.6)$$

В силу (6.4.6) для асимптотического сдвига из (6.4.4) получаем представление

$$\delta_S(\beta_0) = \lim_{n \rightarrow \infty} (n^{-1} I_n(\beta_0, \beta_0))^{1/2} a.$$

Последнее представление аналогично хорошо известному соотношению для сдвига статистики локально наиболее мощного теста в случае независимых наблюдений при стандартных условиях регулярности (см., например, [18, 59, § 9.3]).

В следующем параграфе мы выпишем АОЭ построенного знакового теста со статистикой $T_{n,S}(\beta_0)$ относительно тестов, основанных на оценках наименьших квадратов, модулей, а также ранговых и медианных тестов.

§ 6.5. Сравнение знаковых тестов

с другими непараметрическими тестами

Сравним построенные и изученные в § 6.2 и 6.4 (т.е. в стационарном случае) знаковые тесты с тестами, основанными на общепотребительных оценках наименьших квадратов $\hat{\beta}_{n,LS}$, наименьших модулей $\hat{\beta}_{n,LD}$, а также с непараметрическими ранговыми и медианными тестами, краткие описания которых дадим ниже.

Начнем с оценки наименьших квадратов $\hat{\beta}_{n,LS}$. Как уже отмечалось в п. 5.3.2 (см. (5.3.12)), если $\mathbf{E}\varepsilon_1 = 0$, $\mathbf{E}\varepsilon_1^2 < \infty$ и верна альтернатива

$$H_{1n}(a): \beta = \beta_n := \beta_0 + an^{-1/2} + o(n^{-1/2}),$$

то

$$T_{n,LS}(\beta_0) = (1 - \beta_0^2)^{-1/2} \sqrt{n}(\hat{\beta}_{n,LS} - \beta_0) \xrightarrow{d_{\beta_0}} N((1 - \beta_0^2)^{-1/2} a, 1). \quad (6.5.1)$$

В силу (6.5.1) и (6.4.4) асимптотическая относительная эффективность (АОЭ) теста со статистикой $T_{n,S}(\beta_0) = n^{-1/2} \sqrt{1 - \beta_0^2} l_n^S(\beta_0)$ относительно теста, основанного на статистике $T_{n,LS}(\beta_0)$, равна

$$e_{S,LS} = (2g(0) \mathbf{E}|\varepsilon_1|)^2. \quad (6.5.2)$$

Эффективность $e_{S,LS}$ не зависит от β_0 и не изменяется при изменении параметра масштаба $G(x)$. В частности, для $G(x) \sim N(0, 1)$ $e_{S,LS} = (2/\pi)^2 \approx 0.41$; для функции распределения Лапласа $e_{S,LS} = 1$; для логистической функции распределения $G(x)$ $e_{S,LS} = (\log 2)^2 \approx 0.48$; для $G(x)$ с тяжелыми хвостами $e_{S,LS}$ может быть и больше единицы, и вообще сколь угодно большой. Скажем, для смеси Тьюки, когда $G(x) = (1 - \varepsilon)\Phi(x) + \varepsilon\Phi(x/\tau)$, находим

$$e_{S,LS} = \left\{ \frac{2}{\pi} \left(1 - \frac{\tau - 1}{\tau} \varepsilon\right) (1 + (\tau - 1)\varepsilon) \right\}^2 \rightarrow \infty, \quad \tau \rightarrow \infty.$$

Если \mathfrak{G} — класс функций распределения, имеющих четную плотность с максимумом в нуле, то задача

$$e_{S,LS} = (2g(0) \mathbf{E}|\varepsilon_1|)^2 \Rightarrow \inf_{G \in \mathfrak{G}} \quad (6.5.3)$$

имеет решением $G(x) \sim R(-0.5, 0.5)$, и в этом случае $e_{S,LS} = 1/4$. Таким образом, на естественном классе \mathfrak{G} АОЭ ограничена снизу. Доказательство этого факта вполне аналогично доказательству теоремы 3.3 из гл. 5 в [19, 95].

Для оценки наименьших модулей $\hat{\beta}_{n,LD}$ при $H_{1n}(a)$ в силу теоремы 5.5.4 имеем

$$\sqrt{n}(\hat{\beta}_{n,LD} - \beta_0) \xrightarrow{d_{\beta_0}} N\left(a, (1 - \beta_0^2)((2g(0))^2 \mathbf{E}\varepsilon_1^2)^{-1}\right). \quad (6.5.4)$$

В силу (6.4.4) и (6.5.4) АОЭ $e_{S,LD}$ теста со статистикой $T_{n,S}(\beta_0)$ относительно теста, основанного на $\sqrt{n}(\hat{\beta}_{n,LD} - \beta_0)$, равна

$$e_{S,LD} = (\mathbf{E}|\varepsilon_1|)^2 / \mathbf{E}\varepsilon_1^2. \quad (6.5.5)$$

Тест, основанный на статистике $\sqrt{n}(\hat{\beta}_{n,LD} - \beta_0)$, можно применять, если известна предельная дисперсия оценки $\hat{\beta}_{n,LD}$, равная в силу теоремы 5.5.4 (и при H_0 , и при $H_{1n}(a)$)

$$\sigma_{LD}^2(\beta_0) = (1 - \beta_0^2)((2g(0))^2 \mathbf{E}\varepsilon_1^2)^{-1},$$

или ее состоятельная оценка. Удобнее поэтому пользоваться тестовой статистикой $T_{n,LD}(\beta_0)$ из п. 5.5.2:

$$T_{n,LD}(\beta_0) = s_n^{-1} n^{-1/2} \sum_{k=1}^n u_{k-1} \text{sign}(u_k - \beta_0 u_{k-1}), \quad s_n^2 = n^{-1} \sum_{k=1}^n u_k^2.$$

В силу теоремы 5.5.4 при $H_{1n}(a)$

$$T_{n,LD}(\beta_0) \xrightarrow{d_{\beta_n}} N(2g(0)(E\varepsilon_1^2)^{1/2}(1 - \beta_0^2)^{-1/2}a, 1),$$

следовательно, АОЭ знакового теста относительно теста со статистикой $T_{n,LD}(\beta_0)$ также равна $e_{S,LD}$ из (6.5.5). Эта эффективность не зависит от β_0 и параметра масштаба $G(x)$. Хотя $e_{S,LS} < 1$, напомним, что $\hat{\beta}_{n,LD}$ асимптотически нормальна при $H_{1n}(a)$, в частности, при условии $E|\varepsilon_1|^{2+\Delta} < \infty$, в то время как для знаковой статистики достаточно условия $E|\varepsilon_1|^{1+\Delta} < \infty$.

Перейдем теперь к ранговым тестам. Проверка гипотез с помощью локально оптимальных ранговых и знако-ранговых тестов в схемах авторегрессии-скользящего среднего и линейной регрессии с авторегрелированными остатками имеет давнюю историю и обширную библиографию. Как правило, рассматривались схемы с конечной дисперсией, удовлетворяющие условию ЛАН, асимптотическое исследование проводилось с использованием U -статистик, мощности тестов при континуальных альтернативах находились с помощью третьей леммы Ле Кама (см., например, [34, 68–71, 92] и имеющуюся там библиографию).

Иной подход (и родственные представленному в настоящей книге) использован в [55], где строятся и исследуются ранговые оценки параметра в однопараметрической авторегрессионной схеме с, вообще говоря, бесконечной дисперсией. Этот подход применим и для проверки гипотез.

А именно, пусть

$$\varepsilon_k(\theta) = u_k - \theta u_{k-1}, \quad k = 1, \dots, n,$$

и пусть $R_k(\theta)$ — ранг $\varepsilon_k(\theta)$ в последовательности $\varepsilon_1(\theta), \dots, \varepsilon_n(\theta)$.

Известно (см., например, [34]), что при естественных условиях регулярности ЛНМ тест для проверки $H_0: \beta = \beta_0$ против односторонних альтернатив, основанный на рангах $R_1(\beta_0), \dots, R_n(\beta_0)$, имеет статистикой

$$\sum_{t=1}^{n-1} \beta_0^{t-1} \sum_{k=t+1}^n a_n(R_{k-t}(\beta_0), R_k(\beta_0)), \quad (6.5.6)$$

где двуместные метки $a_n(t, s) = E\varepsilon_{(t)}l(\varepsilon_{(s)})$, $l(x) = g'(x)/g(x)$. Тестовая статистика (6.5.6) зависит от $g(x)$, и ЛНМ ранговый тест, как обычно, является параметрическим. Легко видеть, асимптотически для меток $a_n(t, s)$ будут метки $G^{-1}(t/(n+1))l(G^{-1}(s/(n+1)))$.

Структура последних определяет естественный класс непараметрических тестовых статистик

$$\sum_{i=1}^{n-1} \beta_0^{i-1} \sum_{k=i+1}^n \varphi_1\left(\frac{R_{k-i}(\beta_0)}{n+1}\right) \varphi_2\left(\frac{R_k(\beta_0)}{n+1}\right), \quad (6.5.7)$$

функции $\varphi_i(u)$, $u \in [0, 1]$, теперь не обязаны зависеть от $G(x)$.

Для исследования (6.5.7) рассмотрим процесс

$$I_n^R(\theta) = \sum_{t=1}^{n-1} \theta^{t-1} \Gamma_{tn}^R(\theta), \quad (6.5.8)$$

где

$$\Gamma_{tn}^R(\theta) = \sum_{k=t+1}^n \varphi_1\left(\frac{R_{k-t}(\theta)}{n+1}\right) \varphi_2\left(\frac{R_k(\theta)}{n+1}\right).$$

Относительно функций $\varphi_i(u)$, $i = 1, 2$, и $G(x)$ сделаем следующие предположения.

Условие 6.5.1. $\mathbb{E}\varphi_i(G(\varepsilon_1)) = 0$, $\varphi_i(u)$ дифференцируемы и $\varphi_i'(u)$ липшицевы.

Условие 6.5.2. $\mathbb{E}\varepsilon_1 = 0$, $\mathbb{E}|\varepsilon_1|^{1+\Delta} < \infty$ для некоторого $\Delta > 0$.

Условие 6.5.3. $\sup_x g(x) < \infty$, $g(x)$ удовлетворяет условию Липшица.

Условие 6.5.4.

$$\lambda_R(\beta) := -\mathbb{E}\{\varepsilon_1 \varphi_1(G(\varepsilon_1))\} \mathbb{E}\{\varphi_2'(G(\varepsilon_1))g(\varepsilon_1)\} (1 - \beta^2)^{-1} \neq 0.$$

Используя метод из [55], можно показать, что если выполнены условия 6.5.1–6.5.4, верна $H_{1n}(a)$, $0 < \Theta < \infty$, то

$$\sup_{|\theta| \leq \Theta \log n} |n^{-1/2} I_n^R(\beta_n + n^{-1/2}\theta) - n^{-1/2} I_n^R(\beta_n) - \lambda_R(\beta_0)\theta| = o_p(1).$$

Это стохастическое разложение влечет аналогично теореме 6.4.2, что при $H_{1n}(a)$ и условиях 6.5.1–6.5.4

$$n^{-1/2} I_n^R(\beta_0) = n^{-1/2} I_n^R(\beta_n) - \lambda_R(\beta_0)a + o_p(1),$$

откуда

$$n^{-1/2} I_n^R(\beta_0) \xrightarrow{d\beta_n} N(-\lambda_R(\beta_0)a, \mathbb{E}\varphi_1^2(G(\varepsilon_1))\mathbb{E}\varphi_2^2(G(\varepsilon_2))(1 - \beta_0^2)^{-1}). \quad (6.5.9)$$

При непрерывной $G(x)$ и гипотезе H_0 статистика $I_n^R(\beta_0)$ свободна от $G(x)$ при конечных n .

Рассмотрим случай спирменовских меток $\varphi_i(u) = u - 1/2$ и пусть

$$T_{n,R}(\beta_0) = 12\sqrt{1 - \beta_0^2} n^{-1/2} I_n^R(\beta_0).$$

В силу (6.5.9) при гипотезе H_0 статистика $T_{n,R}(\beta_0)$ распределена асимптотически как стандартная гауссовская величина. При $H_{1n}(a)$

$$T_{n,R}(\beta_0) \xrightarrow{d_{\beta_n}} N\left(12\left(\int g^2(x)dx \int xG(x)g(x)dx\right)(1 - \beta_0^2)^{-1/2}a; 1\right). \quad (6.5.10)$$

Из (6.4.4) и (6.5.10) следует, что АОЭ $e_{S,R}$ ЛНМ знакового теста со статистикой $T_{n,S}(\beta_0)$ относительно теста со статистикой $T_{n,R}(\beta_0)$ в случае спирменовских меток есть

$$e_{S,R} = \frac{(2g(0)\mathbf{E}|\varepsilon_1|)^2}{144\left(\int g^2(x)dx \int xG(x)g(x)dx\right)^2} \quad (6.5.11)$$

и не зависит от β_0 и параметра масштаба $G(x)$. В частности, для $G(x) \sim N(0, 1)$ $e_{S,R} \approx 0.44$; для логистической $G(x)$ $e_{S,R} \approx 0.48$; для ланпассовской $G(x)$ $e_{S,R} \approx 0.82$.

В приведенных примерах $e_{S,R} < 1$. Это свидетельствует, как, впрочем, и следовало ожидать, о предпочтении ранговых тестов перед знаковыми. Напомним, однако, что статистика $n^{-1/2} I_n^S(\beta_0)$ при гипотезе $H_0: \beta = \beta_0$ сохраняет предельное гауссовское распределение (и распределение при конечных n) и для разнораспределенных ε_i , удовлетворяющих условию (6.2.16). Ранговые статистики требуют одинаковой распределенности величин ε_i .

В заключение параграфа рассмотрим еще медианные тесты, введенные в [52], и связанные с медианной оценкой из п. 5.5.3. Они получаются следующим образом.

Пусть $\hat{\beta}_{n,M}$ — медиана массива

$$\left\{ \frac{u_k}{u_{k-1}}, k = 1, \dots, n \right\}.$$

Тогда (см. § 5.5) при условиях 6.1.1–6.1.3

$$\sqrt{n}(\hat{\beta}_{n,M} - \beta) = -\lambda_M^{-1}(\beta) n^{-1/2} I_n^M(\beta) + o_p(1), \quad (6.5.12)$$

где

$$\lambda_M(\beta) = -2g(0)\mathbf{E}_\beta|u_1|, \quad I_n^M(\theta) = \sum_{k=1}^n \text{sign } u_{k-1} \text{sign}(u_k - \theta u_{k-1}).$$

Тестовая статистика для проверки H_0 определяется главной частью разложения (6.5.12) и равна $T_{n,M}(\beta_0) = n^{-1/2} I_n^M(\beta_0)$. При условиях 6.1.1-6.1.3 и $H_{1n}(a)$

$$T_{n,M}(\beta_0) \xrightarrow{d} N(-\lambda_M(\beta_0)a, 1). \quad (6.5.13)$$

Напомним (см. п. 5.5.3), что при гипотезе H_0 и предположениях $E \log^+ |\varepsilon_1| < \infty$ и $P\{\varepsilon_1 < 0\} = P\{\varepsilon_1 > 0\} = 1/2$ статистика $I_n^M(\beta_0)$ свободна от $G(x)$, а именно

$$(I_n^M(\beta_0) + n)/2 \sim \text{Bi}(n, 1/2).$$

Из соотношений (6.5.13) и (6.4.4) следует, что АОЭ $e_{S,M}(\beta_0)$ знакового теста со статистикой $T_{n,S}(\beta_0)$ относительно теста со статистикой $T_{n,M}(\beta_0)$ равна

$$e_{S,M}(\beta_0) = (1 - \beta_0^2)^{-1} (E|\varepsilon_1|/E_{\beta_0}|u_1|)^2. \quad (6.5.14)$$

Из (6.1.1) следует

$$E|\varepsilon_1|(1 + |\beta|)^{-1} \leq E_{\beta}|u_1| \leq E|\varepsilon_1|(1 - |\beta|)^{-1},$$

что влечет совместно с (6.5.14)

$$\frac{1 - |\beta_0|}{1 + |\beta_0|} \leq e_{S,M}(\beta_0) \leq \frac{1 + |\beta_0|}{1 - |\beta_0|}.$$

При $G(x) \sim N(0, 1)$ $e_{S,M} = 1$ при всех β_0 , а при $\beta_0 = 0$ $e_{S,M} = 1$ при всех $G(x)$, удовлетворяющих условиям 6.1.1-6.1.3.

В заключение параграфа подчеркнем особо: если знаковые тесты (а также ранговые и медианные) свободны от $G(x)$ при конечных n , то общепотребительные тесты, основанные на оценках $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$, свободны от $G(x)$ лишь асимптотически.

§ 6.6. Знаковые оценки параметров

В этом параграфе мы строим непараметрические знаковые оценки параметра β в схеме (6.1.1) по наблюдениям u_0, u_1, \dots, u_n , удовлетворяющим (6.1.2). Метод оценивания основан на результатах § 6.2 и § 6.4.

6.6.1. Знаковая оценка $\hat{\beta}_{n,S}$. Введем величины

$$p_t(\theta) = 2P_{\beta}\{(u_1 - \theta u_0)(u_{1+t} - \theta u_t) > 0\} - 1, \quad t = 1, 2, \dots \quad (6.6.1)$$

При фиксированном θ для $\Gamma_{tn}(\theta)$, определенных соотношением (6.2.9),

$$E_{\beta} n^{-1} \Gamma_{tn}(\theta) = (1 - t/n) p_t(\theta), \quad t = 1, 2, \dots, n-1. \quad (6.6.2)$$

Из (6.6.2), пользуясь определением $l_n^S(\theta)$ (6.2.10) и неравенством $|p_t(\theta)| \leq 1$, получаем для $|\theta| < 1$

$$E_{\beta} n^{-1} l_n^S(\theta) = \sum_{t=1}^{n-1} \theta^{t-1} \left(1 - \frac{t}{n}\right) p_t(\theta) \rightarrow \Lambda_S(\theta),$$

где

$$\Lambda_S(\theta) = \sum_{t=1}^{\infty} \theta^{t-1} p_t(\theta). \quad (6.6.3)$$

Равномерную сходимость $n^{-1} l_n^S(\theta)$ к $\Lambda_S(\theta)$ по вероятности устанавливает следующая теорема.

Теорема 6.6.1. Пусть $G(x)$ удовлетворяет условию Липшица и $E|\varepsilon_1|^{\Delta} < \infty$ при некотором $\Delta > 0$.

Тогда при любом $0 < \delta < 1$

$$\sup_{|\theta| \leq 1-\delta} |n^{-1} l_n^S(\theta) - \Lambda_S(\theta)| = o_p(1)$$

и функция $\Lambda_S(\theta)$ равномерно непрерывна для $|\theta| \leq 1 - \delta$.

Доказательство теоремы 6.6.1 мы не приводим, оно идейно и технически близко к доказательству теоремы 6.4.1. Заинтересованный читатель может найти его в [56].

При условиях 6.1.1–6.1.3 легко вычислить производную функции $\Lambda_S(\theta)$ в точке $\theta = \beta$, которая оказывается равной $\lambda_S(\beta)$ из (6.4.1). Теоремы 6.6.1 и 6.4.1 утверждают, таким образом, что

$$n^{-1} l_n^S(\theta) = \Lambda_S(\theta) + o_p(1), \quad (6.6.4)$$

$$n^{-1/2} l_n^S(\beta + n^{-1/2}\theta) = n^{-1/2} l_n^S(\beta) + \lambda_S(\beta)\theta + o_p(1), \quad (6.6.5)$$

причем бесконечно малые в (6.6.4) и (6.6.5) равномерно малы на компактах $|\theta| \leq 1 - \delta$ — в первом и $|\theta| \leq \Theta_n$ — во втором случае.

Очевидно, $\Lambda_S(\beta) = 0$ при условии 6.1.1. Это замечание вместе с теоремой 6.6.1 позволяет утверждать, что если функция $\Lambda_S(\theta)$ имеет единственный корень на некотором отрезке, содержащем β , то уравнение

$$l_n^S(\theta) \div 0 \quad (6.6.6)$$

с вероятностью, стремящейся к единице, будет иметь корень $\hat{\beta}_{n,S}$, и этот корень будет состоятельной оценкой для β . Символ \div в (6.6.6) означает, как и раньше, что под решением понимаются те точки θ , в которых функция $l_n^S(\theta)$, имеющая только разрывы первого рода, переходит через нуль.

Таким образом, естественно искать оценку неизвестного параметра β как решение уравнения (6.6.6), по модулю меньше единицы.

Мы уже упоминали в § 5.6 класс RA -оценок, определяемых как решение уравнения

$$\sum_{i=1}^{n-1} \psi_i(U_n, \theta) \div 0, \quad (6.6.7)$$

в котором $U_n = (u_0, \dots, u_n)$ и

$$\begin{aligned} \psi_i(U_n, \theta) = & \eta(u_i - \theta u_{i-1}, u_{i+1} - \theta u_i) + \theta \eta(u_i - \theta u_{i-1}, u_{i+2} - \theta u_{i+1}) + \dots \\ & \dots + \theta^{n-1} \eta(u_i - \theta u_{i-1}, u_n - \theta u_{n-1}). \end{aligned} \quad (6.6.8)$$

Такой класс оценок был введен в [58] для общей схемы авторегрессии—скользящего среднего.

Из определения $l_n^S(\theta)$ (6.2.10) следует, что и знаковая оценка $\hat{\beta}_{n,S}$ является RA -оценкой с

$$\eta(\xi_1, \xi_2) = \text{sign}(\xi_1, \xi_2).$$

К сожалению, RA -оценки были исследованы лишь для гладких функций η , поэтому мы не можем воспользоваться имеющимися результатами и будем исследовать знаковые оценки непосредственно. Основным результатом в этом исследовании — теорема 6.4.1.

Любопытно подчеркнуть, что RA -оценки были введены без всякой связи с тестами, в то время как наши непараметрические знаковые оценки естественно порождаются локально оптимальными непараметрическими знаковыми тестами.

Вернемся к уравнению (6.6.6). К сожалению, установить отдельность корня $\theta = \beta$ функции $\Lambda_S(\theta)$ при весьма скромных предположениях теоремы 6.6.1 не удается. Далее до конца этого параграфа при исследовании решения (6.6.6) мы предполагаем выполненными более сильные условия 6.1.1–6.1.3. При этих условиях мы устанавливаем не только существование, но и асимптотическую нормальность решения $\hat{\beta}_{n,S}$ уравнения (6.6.6).

Следующая теорема 6.6.2 была опубликована в [56].

Теорема 6.6.2. Пусть выполнены условия 6.1.1–6.1.3.

Тогда с вероятностью, стремящейся к единице, существует решение $\hat{\beta}_{n,S}$ уравнения (6.6.6) такое, что

$$\sqrt{n}(\hat{\beta}_{n,S} - \beta) = -\lambda_S^{-1}(\beta) n^{-1/2} l_n^S(\beta) + o_p(1),$$

$$\sqrt{n}(\hat{\beta}_{n,S} - \beta) \xrightarrow{d} N(0, \sigma_S^2(\beta)),$$

где

$$\sigma_S^2(\beta) = (1 - \beta^2)(2g(0)E|\varepsilon_1|)^{-2}.$$

Доказательство. В силу теоремы 6.4.1 (с $\beta_n = \beta_0 = \beta$) при достаточно большом $A > 0$

$$n^{-1/2} l_n^S(\beta - An^{-1/2}) = n^{-1/2} l_n^S(\beta) - \lambda_S(\beta)A + o_p(1) > 0$$

с вероятностью, сколь угодно близкой к единице, равномерно по $n > n_0$, так как

$$n^{-1/2} l_n^S(\beta) \xrightarrow{d_p} N(0, (1 - \beta^2))$$

в силу теоремы 6.2.3 и $\lambda_S(\beta) < 0$. Выписывая аналогичное неравенство

$$l_n^S(\beta + An^{-1/2}) < 0,$$

убеждаемся, что в интервале $(\beta - An^{-1/2}, \beta + An^{-1/2})$ с вероятностью, сколь угодно близкой к единице, равномерно по $n > n_0$ существует корень $\hat{\beta}_{n,S}$ уравнения (6.6.6) и $\sqrt{n}(\hat{\beta}_{n,S} - \beta) = O_p(1)$. В точках разрыва функция $l_n^S(\theta)$ имеет скачки не больше $2(1 - \theta^2)^{-1}$, поэтому $n^{-1/2} l_n^S(\hat{\beta}_{n,S}) = o_p(1)$. Применяя следствие 6.4.1 (с $\hat{\beta}_n = \hat{\beta}_{n,S}$, $\beta_n = \beta_0 = \beta$, $\alpha = 0$), получаем представление

$$\sqrt{n}(\hat{\beta}_{n,S} - \beta) = -\lambda_S^{-1}(\beta)n^{-1/2} l_n^S(\beta) + o_p(1),$$

которое влечет в силу теоремы 6.2.3 сходимость

$$\sqrt{n}(\hat{\beta}_{n,S} - \beta) \xrightarrow{d_p} N(0, \sigma_S^2(\beta)),$$

где

$$\sigma_S^2(\beta) = \lambda_S^{-2}(\beta)(1 - \beta^2)^{-1} = (1 - \beta^2)(2g(0) E|\epsilon_1|)^{-2}.$$

Теорема 6.6.2 доказана. \square

Теорема 6.6.2 позволяет вычислить АОЭ знаковой оценки $\hat{\beta}_{n,S}$ в сопоставлении с другими. Например, с оценкой наименьших квадратов $\hat{\beta}_{n,LS}$, оценкой наименьших модулей $\hat{\beta}_{n,LD}$ и медианной оценкой $\hat{\beta}_{n,M}$. Эти эффективности оказываются равными $e_{S,LS}$, $e_{S,LD}$ и $e_{S,M}$ соответственно и определены соотношениями (6.5.2), (6.5.5), (6.5.14) соответственно. Относительно них можно повторить все сказанное в § 6.5.

Сравним знаковую оценку $\hat{\beta}_{n,S}$ с ранговой, которая получается решением уравнения

$$l_n^R(\theta) \div 0;$$

функция $l_n^R(\theta)$ определена в (6.5.8). Последнее уравнение, как показано в [55], при условиях 6.5.1–6.5.4 с вероятностью, стремящейся к единице, имеет такое решение $\hat{\beta}_{n,R}$, что

$$\sqrt{n}(\hat{\beta}_{n,R} - \beta) \xrightarrow{d_p} N(0, \sigma_R^2(\beta)),$$

где

$$\sigma_R^2(\beta) = (1 - \beta^2) \frac{\mathbf{E}\varphi_1^2(G(\varepsilon_1))\mathbf{E}\varphi_2^2(G(\varepsilon_1))}{\left(\mathbf{E}\left\{\varepsilon_1\varphi_1(G(\varepsilon_1))\right\}\mathbf{E}\left\{\varphi_2'(G(\varepsilon_1))g(\varepsilon_1)\right\}\right)^2}.$$

С помощью последнего соотношения легко найти АОЭ знаковой оценки $\hat{\beta}_{n,S}$ относительно ранговой $\hat{\beta}_{n,R}$, и в случае спирменовских меток она равна эффективности $e_{S,R}$ из (6.5.11). Отметим, что ранговые оценки типа $\hat{\beta}_{n,R}$ в многопараметрической авторегрессии с конечной дисперсией, удовлетворяющей условиям ЛАН, исследовались в [65], но иным, чем в [55], методом.

6.6.2. Знаковая оценка $\beta_{n,S}^*$. Если корень уравнения (6.6.6) при $|\theta| < 1$ единственный, то его можно найти с нужной степенью точности, например методом деления отрезка пополам. К сожалению, функция $l_n^S(\theta)$ не является монотонной (в отличие, например, от $l_n^{LD}(\theta)$ и $l_n^M(\theta)$ из § 5.5), а потому гарантировать единственность решения уравнения (6.6.6) нельзя. Если же уравнение имеет несколько решений, то необходим правильный выбор одного из этих решений. Разумеется, в качестве $\hat{\beta}_{n,S}$ всегда можно брать корень уравнения (6.6.6), ближайший к какой-нибудь \sqrt{n} -состоятельной оценке $\hat{\beta}_n$ параметра β . Например, если $G(x)$ симметрична относительно нуля и $\mathbf{E} \log^+ |\varepsilon_1| < \infty$, то годится $\hat{\beta}_{n,LS}$ (см. п. 5.3.2), а если выполнены условия 6.1.1–6.1.3, то годится медианная оценка $\hat{\beta}_{n,M}$ (см. п. 5.5.3). Это простейший выход из затруднения.

Второй способ состоит в построении оценки, асимптотически эквивалентной $\hat{\beta}_{n,S}$.

Теорема 6.6.3. Пусть выполнены условия 6.1.1–6.1.3, $\hat{\beta}_n$ — предварительная \sqrt{n} -состоятельная оценка β , \hat{e}_n — состоятельная оценка $\lambda_S(\beta)$,

$$\beta_{n,S}^* := \hat{\beta}_n - (n\hat{e}_n)^{-1} l_n^S(\hat{\beta}_n).$$

Тогда

$$\begin{aligned} \sqrt{n}(\beta_{n,S}^* - \beta) &= -\lambda_S^{-1}(\beta)n^{-1/2} l_n^S(\beta) + o_p(1), \\ \sqrt{n}(\beta_{n,S}^* - \beta) &\xrightarrow{d} N\left(0, \sigma_S^2(\beta)\right), \quad \sigma_S^2(\beta) = (1 - \beta^2)(2g(0)\mathbf{E}|\varepsilon_1|)^{-2}. \end{aligned}$$

Доказательство теоремы 6.6.3 получается применением следствия 6.4.1 (с $\beta_n = \beta_0 = \beta$, $\alpha = 0$) и теоремы 6.2.3.

Оценки $\hat{\beta}_{n,S}$ из теоремы 6.6.2 и $\beta_{n,S}^*$ из теоремы 6.6.3 эквивалентны в смысле предельного распределения.

О подходящих оценках $\hat{\beta}_n$ мы уже говорили. Подходящих оценок $\hat{\varepsilon}_n$ также несколько. Например, в силу следствия 6.4.1 (при $\beta_n = \beta_0 = \beta$) можно взять

$$\hat{\varepsilon}_n = \left(l_n^S(\hat{\beta}_n + hn^{-1/2}) - l_n^S(\hat{\beta}_n) \right) (n^{1/2}h)^{-1}$$

для фиксированного $h \neq 0$. Также можно состоятельно оценить $\lambda_S(\beta)$, оценивая состоятельно по отдельности $g(0)$ и $E|\varepsilon_1|$. Например, легко проверить, что состоятельной оценкой среднего будет

$$\hat{m}_n = n^{-1} \sum_{k=1}^n |u_k - \hat{\beta}_n u_{k-1}|, \quad (6.6.9)$$

а состоятельной оценкой плотности $g(0)$ — статистика типа Парзена—Розенблатта

$$\hat{g}_n = (nh_n)^{-1} \sum_{k=1}^n \varphi \left(\frac{u_k - \hat{\beta}_n u_{k-1}}{h_n} \right), \quad (6.6.10)$$

где $\varphi(x)$ — стандартная гауссовская плотность вероятности, а $h_n = n^{-\alpha}$, $0 < \alpha < 1/4$. Тогда $\hat{\varepsilon}_n = -2\hat{g}_n \hat{m}_n (1 - \hat{\beta}_n^2)^{-1}$.

6.6.3. Знаковая оценка $\tilde{\beta}_{n,S}^*$. В заключение параграфа рассмотрим еще один способ построения знаковой оценки. Его идея состоит в том, чтобы от решения уравнения (6.6.6) перейти к задаче минимизации функции $l_n^S(\theta)$, но не по всем возможным θ , как это обычно делается, а по θ из «асимптотически малого компакта»

$$\mathbb{Q}(\hat{\beta}_n) := \{ \theta: |\theta - \hat{\beta}_n| \leq n^{-1/2} \log n \};$$

где $\hat{\beta}_n$ — предварительная \sqrt{n} -состоятельная оценка β . При этом, чтобы задача минимизации всегда имела решение, удобно в качестве целевой функции взять не $l_n^S(\theta)$, а асимптотически равномерно близкую ей кусочно постоянной функцию

$$l_n^S(\hat{\beta}_n, \theta) = \sum_{t=1}^{n-1} \hat{\beta}_n^{t-1} \Gamma_{tn}(\theta),$$

имеющую разрывы в точках u_k/u_{k-1} , $k = 1, \dots, n$. Итак, в качестве оценки β будем брать (и обозначать $\tilde{\beta}_{n,S}$) любую случайную величину, являющуюся решением задачи

$$|l_n^S(\hat{\beta}_n, \theta)| \Rightarrow \inf_{\theta \in \mathbb{Q}(\hat{\beta}_n)}. \quad (6.6.11)$$

В качестве $\tilde{\beta}_{n,S}$ можно брать, например, середину самого левого отрезка, на котором достигается минимум целевой функции в (6.6.11).

Можно показать (см. [9]), что при условиях 6.1.1–6.1.3 и $0 < \Theta < \infty$

$$\sup_{|\theta| \leq \Theta \log n} |n^{-1/2} l_n^S(\hat{\beta}_n, \beta + n^{-1/2}\theta) - n^{-1/2} l_n^S(\beta + n^{-1/2}\theta)| = o_p(1).$$

Последнее соотношение и теорема 6.4.1 (с $\beta_n = \beta_0 = \beta$) влекут стохастическое разложение

$$\sup_{|\theta| \leq \Theta \log n} |n^{-1/2} l_n^S(\hat{\beta}_n, \beta + n^{-1/2}\theta) - n^{-1/2} l_n^S(\beta) - \lambda_S(\beta)\theta| = o_p(1).$$

Из этого разложения в свою очередь следует (см. доказательство следствия 6.4.1): если оценка β_n^* такова, что

$$\sqrt{n}(\beta_n^* - \beta) / \log n = O_p(1),$$

то

$$n^{-1/2} l_n^S(\hat{\beta}_n, \beta_n^*) = n^{-1/2} l_n^S(\beta) + \lambda_S(\beta)\sqrt{n}(\beta_n^* - \beta) + o_p(1). \quad (6.6.12)$$

Рассмотрим теперь величину

$$\tilde{\beta}_n = -\lambda_S^{-1}(\beta)n^{-1/2} l_n^S(\beta),$$

для которой $\sqrt{n}(\tilde{\beta}_n - \beta) = O_p(1)$ в силу теоремы 6.2.3. Подставляя $\tilde{\beta}_n$ вместо β_n^* в (6.6.12), убеждаемся, что

$$n^{-1/2} l_n^S(\hat{\beta}_n, \tilde{\beta}_n) = o_p(1). \quad (6.6.13)$$

Поскольку с вероятностью, стремящейся к единице, $\tilde{\beta}_n \in \mathbb{Q}(\hat{\beta}_n)$, то в силу (6.6.13) любое решение $\tilde{\beta}_{n,S}$ задачи (6.6.11) и по-прежнему удовлетворяет условию

$$n^{-1/2} l_n^S(\hat{\beta}_n, \tilde{\beta}_{n,S}) = o_p(1). \quad (6.6.14)$$

Теперь (6.6.12) (с $\beta_n^* = \tilde{\beta}_{n,S}$), (6.6.14) и теорема 6.2.3 влекут следующую теорему.

Т е о р е м а 6.6.4. Пусть выполнены условия 6.1.1–6.1.3 и $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$.

Тогда для любого решения $\tilde{\beta}_{n,S}$ задачи (6.6.11) имеют место соотношения

$$\sqrt{n}(\tilde{\beta}_{n,S} - \beta) = \lambda_S^{-1}(\beta)l_n^S(\beta) + o_p(1),$$

$$\sqrt{n}(\tilde{\beta}_{n,S} - \beta) \xrightarrow{d_p} N(0, \sigma_S^2(\beta)), \quad \sigma_S^2(\beta) = (1 - \beta^2)(2g(0)\mathbb{E}|\varepsilon_1|)^{-2}.$$

В следующем параграфе мы изучим устойчивость построенных оценок к выбросам в наблюдениях u_0, \dots, u_n .

§ 6.7. Функционалы влияния знаковых оценок

Пусть при наличии засорения наблюдается вектор $Y_n = (y_0, \dots, y_n)$, где

$$y_i = u_i + z_i^\gamma \xi_i, \quad i \in \mathbb{Z}. \quad (6.7.1)$$

В (6.7.1) величины $\{u_i\}$ подчиняются модели (6.1.1); $\{z_i^\gamma\}$ — н.о.р.сл.в., $z_i^\gamma \sim \text{Bi}(1, \gamma)$, $0 \leq \gamma \leq 1$; $\{\xi_i\}$ — н.о.р.сл.в. с распределением μ_ξ ; последовательности $\{u_i\}$, $\{z_i^\gamma\}$, $\{\xi_i\}$ независимы между собой. Мы рассматриваем, таким образом, схему независимых (одиночных) выбросов из § 5.6.

6.7.1. Функционал влияния знаковой оценки $\hat{\beta}_{n,S}$. Знаковая оценка $\hat{\beta}_{n,S}$, определенная в случае незасоренной выборки уравнением (6.6.6), в случае схемы (6.7.1) удовлетворяет уравнению

$$l_n^S(\theta) = \sum_{i=1}^{n-1} \psi_i^S(Y_n, \theta) \div 0, \quad (6.7.2)$$

в котором

$$\begin{aligned} \psi_i^S(Y_n, \theta) = & \text{sign}((y_i - \theta y_{i-1})(y_{i+1} - \theta y_i)) + \\ & + \theta \text{sign}((y_i - \theta y_{i-1})(y_{i+2} - \theta y_{i+1})) + \dots \\ & \dots + \theta^{n-1} \text{sign}((y_i - \theta y_{i-1})(y_n - \theta y_{n-1})). \end{aligned} \quad (6.7.3)$$

Как мы уже отмечали в предыдущем параграфе, знаковая оценка $\hat{\beta}_{n,S}$ является *RA*-оценкой. Функционалы влияния таких оценок найдены в [100], но, к сожалению, только для гладких функций η , определяющих оценку (см. (6.6.7)–(6.6.8)). Мы найдем функционалы влияния знаковых оценок непосредственно, без привлечения упомянутых результатов.

Аналогично теореме 6.6.1 показывается: в условиях формулируемой ниже леммы 6.7.1 равномерно по $|\theta| \leq 1 - \delta$, $0 < \delta < 1$, для любого γ , $0 \leq \gamma \leq 1$,

$$n^{-1} l_n^S(\theta) = n^{-1} \sum_{i=1}^{n-1} \psi_i^S(Y_n, \theta) \xrightarrow{P_\theta} \Lambda_S(\gamma, \theta), \quad (6.7.4)$$

где

$$\Lambda_S(\gamma, \theta) = \sum_{t=1}^{\infty} \theta^{t-1} \left(2 P_\beta \{ (y_1 - \theta y_0)(y_{1+t} - \theta y_t) > 0 \} - 1 \right). \quad (6.7.5)$$

В действительности (6.7.4) верно и при условиях, более слабых, чем в лемме 6.7.1. Однако мы не будем заниматься подобными уточнениями, так как основная теорема 6.7.1 будет доказана именно в предположениях леммы 6.7.1.

Легко видеть, что при условии 6.1.1 $\Lambda_S(0, \beta) = \Lambda_S(\beta) = 0$, где функция $\Lambda_S(\theta)$ определена соотношением (6.6.3).

Лемма 6.7.1. Пусть $E\varepsilon_1 = 0$, $G(0) = 1/2$, $E|\xi_1| < \infty$, $g(x)$ ограничена, непрерывна, $g(0) > 0$.

Тогда существуют непрерывные по (γ, θ) , $0 \leq \gamma \leq 1$, $|\theta| < 1$, производные $\frac{\partial}{\partial \gamma} \Lambda_S(\gamma, \theta)$, $\frac{\partial}{\partial \theta} \Lambda_S(\gamma, \theta)$ и

$$\frac{\partial}{\partial \theta} \Lambda_S(0, \beta) = \lambda_S(\beta) \hat{=} -2g(0) E|\varepsilon_1| (1 - \beta^2)^{-1} < 0, \quad (6.7.6)$$

$$\frac{\partial}{\partial \gamma} \Lambda_S(0, \beta) = E_\beta (1 - 2G(-\xi_1)) (1 - 2G(\beta\xi_1)). \quad (6.7.7)$$

Напомним, что функция $\lambda_S(\beta)$ уже появлялась в (6.4.1).

Доказательство. Обозначим

$$p_t(\gamma, \theta) = P_\beta \{ (y_1 - \theta y_0)(y_{1+t} - \theta y_t) > 0 \} - 1/2,$$

тогда в силу (6.7.5)

$$\Lambda_S(\gamma, \theta) = 2 \sum_{t=1}^{\infty} \theta^{t-1} p_t(\gamma, \theta). \quad (6.7.8)$$

Перепишем $p_t(\gamma, \theta)$ в виде $p_t(\gamma, \theta) = p_{t1}(\gamma, \theta) + p_{t2}(\gamma, \theta) - 1/2$, где

$$p_{t1}(\gamma, \theta) = P_\beta \{ y_1 - \theta y_0 < 0, y_{1+t} - \theta y_t < 0 \},$$

$$p_{t2}(\gamma, \theta) = P_\beta \{ y_1 - \theta y_0 > 0, y_{1+t} - \theta y_t > 0 \}.$$

Покажем, что величины $p_{ti}(\gamma, \theta)$ непрерывно дифференцируемы, и найдем их производные. Рассмотрим $p_{t1}(\gamma, \theta)$, и пусть дальше $t = 1$. Рассуждения в других случаях аналогичны. Пусть Ω — сигма-алгебра, порожденная $\{\varepsilon_i, i \leq 1, z_0^\gamma, z_1^\gamma, z_2^\gamma, \xi_0, \xi_1, \xi_2\}$. Тогда

$$\begin{aligned} p_{11}(\gamma, \theta) &= E_\beta \{ I(y_1 - \theta y_0 < 0) I(y_2 - \theta y_1 < 0) \} = \\ &= E_\beta E_\beta \{ I(\varepsilon_1 - (\theta - \beta)u_0 + z_1^\gamma \xi_1 - \theta z_0^\gamma \xi_0 < 0) \times \\ &\quad \times I(\varepsilon_2 - (\theta - \beta)u_1 + z_2^\gamma \xi_2 - \theta z_1^\gamma \xi_1 < 0) | \Omega \} = \end{aligned}$$

$$= \mathbf{E}_\beta \left\{ I(\varepsilon_1 - (\theta - \beta)u_0 + z_1^\gamma \xi_1 - \theta z_0^\gamma \xi_0 < 0) \times \right. \\ \left. \times G((\theta - \beta)u_1 - z_2^\gamma \xi_2 + \theta z_1^\gamma \xi_1) \right\}. \quad (6.7.9)$$

Пусть событие \mathbb{H}_i , $i = 0, 1, 2, 3$, состоит в том, что i величин из z_1^γ , z_2^γ , z_3^γ отличны от нуля. Перепишем (6.7.9) по формуле полного математического ожидания в виде

$$\sum_{i=0}^3 \mathbf{E}_\beta \left\{ I(\varepsilon_1 < (\theta - \beta)u_0 - z_1^\gamma \xi_1 + \theta z_0^\gamma \xi_0) \times \right. \\ \left. \times G((\theta - \beta)u_1 - z_2^\gamma \xi_2 + \theta z_1^\gamma \xi_1) | \mathbb{H}_i \right\} \mathbf{P}\{\mathbb{H}_i\}. \quad (6.7.10)$$

В (6.7.10) $P\{\mathbb{H}_i\}$ — полиномы по γ , а условные средние от γ не зависят и в условиях леммы непрерывно дифференцируемы по θ . Следовательно, функция $p_{11}(\gamma, \theta)$ непрерывно дифференцируема по (γ, θ) , $0 \leq \gamma \leq 1$, $\theta \in \mathbb{R}^1$, причем производные равномерно ограничены. Тем же свойством обладают функции $p_{t1}(\gamma, \theta)$ при $t > 1$ и $p_{t2}(\gamma, \theta)$ при $t \geq 1$, а значит, и $p_t(\gamma, \theta)$ при $t \geq 1$. Отсюда следует равномерная при $|\theta| \leq 1 - \delta$, $0 \leq \gamma \leq 1$, сходимость рядов, получаемых формальным дифференцированием ряда (6.7.8) по θ и γ . Следовательно, $\Lambda_S(\gamma, \theta)$ непрерывно дифференцируема по (γ, θ) , $0 \leq \gamma \leq 1$, $|\theta| < 1$.

Соотношение (6.7.6) легко установить, если воспользоваться (6.7.8) и переписать каждое из $p_t(\gamma, \theta)$ в виде, аналогичном (6.7.9). Наконец, для доказательства (6.7.7) заметим, что $p_1(0, \beta) = 0$, $p_t(\gamma, \beta) = o(\gamma)$ при $t \geq 2$, а

$$p_1(\gamma, \beta) = (1/2) \gamma \mathbf{E}_\beta \operatorname{sign}(\varepsilon_1 + \xi_1)(\varepsilon_2 - \beta \xi_1) + o(\gamma) = \\ = (1/2) \gamma \mathbf{E}_\beta (1 - 2G(-\xi_1))(1 - 2G(\beta \xi_1)) + o(\gamma).$$

Лемма 6.7.1 доказана. \square

Из леммы 6.7.1 и условия $\Lambda_S(0, \beta) = 0$ следует, что в некоторой окрестности точки $(0, \beta)$ уравнение

$$\Lambda_S^\gamma(\gamma, \theta) = 0$$

определяет дифференцируемую функцию $\theta_\gamma^S = \theta(\gamma)$, причем $\theta_0^S = \beta$,

$$\left. \frac{d\theta_\gamma^S}{d\gamma} \right|_{\gamma=0} = -\lambda_S^{-1}(\beta) \frac{\partial}{\partial \gamma} \Lambda_S(0, \beta), \quad (6.7.11)$$

где $\lambda_S(\beta)$ и $\frac{\partial}{\partial \gamma} \Lambda_S(0, \beta)$ выписаны в (6.7.6) и (6.7.7).

При малых γ функция $\Lambda_S(\gamma, \theta)$ строго убывает по θ в окрест-

ности точки β , так как в силу (6.7.6) $\frac{\partial}{\partial \theta} \Lambda_S(0, \beta) < 0$ и, следовательно, $\frac{\partial}{\partial \theta} \Lambda_S(\gamma, \theta) < 0$ в окрестности $(0, \beta)$. Этот факт и соотношение (6.7.4) влекут: с вероятностью, стремящейся к единице, существует такое решение $\hat{\beta}_{n,S}$ уравнения (6.7.2), что

$$\hat{\beta}_{n,S} \xrightarrow{P_\beta} \theta_\gamma^S. \quad (6.7.12)$$

Теперь соотношения (6.7.11), (6.7.12), лемма 6.7.1 и определение функционала влияния (5.6.2) влекут следующую теорему.

Теорема 6.7.1. Пусть выполнены условия леммы 6.7.1.

Тогда при достаточно малом γ с вероятностью, стремящейся к единице, существует такое решение $\hat{\beta}_{n,S}$ уравнения (6.7.2), что

$$\hat{\beta}_{n,S} \xrightarrow{P_\beta} \theta_\gamma^S.$$

Функционал влияния оценки $\hat{\beta}_{n,S}$ равен

$$IF(\theta_\gamma^S, \mu_\xi) = \frac{1 - \beta^2}{2g(0) \mathbb{E}|\varepsilon_1|} \mathbb{E}_\beta(1 - 2G(-\xi_1))(1 - 2G(\beta\xi_1)). \quad (6.7.13)$$

При $\beta = 0$ функционал влияния $IF(\theta_\gamma^S, \mu_\xi) = 0$. Если $\xi_i = \xi$ и ξ — постоянная, то

$$\begin{aligned} IF(\theta_\gamma^S, \mu_\xi) &= IF(\theta_\gamma^S, \xi) = \\ &= \frac{1 - \beta^2}{2g(0) \mathbb{E}|\varepsilon_1|} (1 - 2G(-\xi))(1 - 2G(\beta\xi)), \end{aligned} \quad (6.7.14)$$

это ограниченная непрерывная функция ξ .

Пусть \mathfrak{M}_i , $i = 1, 2$, — класс распределений μ_ξ с конечным i -м абсолютным моментом. Модуль среднего в (6.7.13) не больше единицы, причем при $\xi_1 = \xi$, ξ — постоянная, он стремится к единице, если $\xi \rightarrow \infty$ и $\beta \neq 0$. Следовательно,

$$GES(\mathfrak{M}_1, \theta_\gamma^S) = \frac{1 - \beta^2}{2g(0) \mathbb{E}|\varepsilon_1|} < \infty, \quad \beta \neq 0, \quad (6.7.15)$$

и знаковая оценка слабо зависит от выбросов в данных. В этом ее преимущество перед оценками наименьших квадратов и наименьших модулей, которые имеют неограниченные чувствительности даже на более узком, чем \mathfrak{M}_1 , классе \mathfrak{M}_2 (см. соотношения (5.6.9)–(5.6.10) и (5.6.12)–(5.6.13)).

Сравним чувствительность (6.7.15) с чувствительностью медианной оценки. Последняя конечна на \mathfrak{M}_1 . В частности, для гауссовской

$G(x)$ находим из (5.6.15), что $GES(\mathfrak{M}_1, \theta_\gamma^M) = \frac{\pi}{2}(1 - \beta^2)^{1/2}$. Следовательно, в этом случае

$$\frac{GES(\mathfrak{M}_1, \theta_\gamma^S)}{GES(\mathfrak{M}_1, \theta_\gamma^M)} = (1 - \beta^2)^{1/2} < 1,$$

т.е. знаковая оценка предпочтительнее медианной.

6.7.2. Функционал влияния знаковой оценки $\tilde{\beta}_{n,S}$. Рассмотрим теперь знаковую оценку $\tilde{\beta}_{n,S}$, определяемую в случае незасоренной выборки соотношением (6.6.11). При наличии засорения $\tilde{\beta}_{n,S}$ определяется как решение задачи

$$|l_n^S(\hat{\beta}_n, \theta)| \Rightarrow \inf_{\theta \in \mathbb{Q}(\hat{\beta}_n)}, \quad (6.7.16)$$

где $\mathbb{Q}(\hat{\beta}_n) = \{\theta: |\theta - \hat{\beta}_n| \leq n^{-1/2} \log n\}$ и предварительная оценка $\hat{\beta}_n$, так же, как и $l_n^S(\hat{\beta}_n, \theta)$, основана на Y_n .

Условие 6.7.1. Для предварительной оценки $\hat{\beta}_n$ существуют предел $\hat{\beta}_n \xrightarrow{P_\beta} \theta_\gamma$ и функционал влияния $IF(\theta_\gamma, \mu_\xi)$.

Тогда в силу определения $\tilde{\beta}_{n,S}$ существуют предел

$$\tilde{\beta}_{n,S} \xrightarrow{P_\beta} \tilde{\theta}_\gamma^S = \theta_\gamma$$

и функционал

$$IF(\tilde{\theta}_\gamma^S, \mu_\xi) = IF(\theta_\gamma, \mu_\xi).$$

Последнее означает, что знаковая оценка $\tilde{\beta}_{n,S}$ наследует функционал влияния предварительной оценки $\hat{\beta}_n$. Если в качестве $\hat{\beta}_n$ берется одна из оценок $\hat{\beta}_{n,LS}$, $\hat{\beta}_{n,LD}$ или $\hat{\beta}_{n,M}$, то функционал влияния $\tilde{\beta}_{n,S}$ будет даваться соответственно соотношениями (5.6.9), (5.6.12), (5.6.15).

6.7.3. Функционал влияния знаковой оценки $\beta_{n,S}^*$. В заключение параграфа рассмотрим третью знаковую оценку $\beta_{n,S}^*$, определенную в теореме 6.6.3 для схемы без засорения. В случае засоренных данных эта оценка определяется соотношением

$$\beta_{n,S}^* = \hat{\beta}_n - (n\hat{e}_n)^{-1} l_n^S(\hat{\beta}_n), \quad (6.7.17)$$

в котором $\hat{\beta}_n$, \hat{e}_n и $l_n^S(\theta)$ основаны на векторе данных Y_n .

Пусть выполнены условия леммы 6.7.1, условие 6.7.1 и, кроме того, справедливо

Условие 6.7.2. $\hat{\varepsilon}_n \xrightarrow{P_\beta} \theta_\gamma^e$, $\theta_\gamma^e \rightarrow \theta_0^e = \lambda_S(\beta)$, $\gamma \rightarrow 0$.

Условию 6.7.2 удовлетворяет, например, оценка $\hat{\varepsilon}_n = -2\hat{g}_n \hat{m}_n (1 - \hat{\beta}_n^2)^{-1}$, построенная по Y_n аналогично (6.6.9)–(6.6.10), если только $\sqrt{n}(\hat{\beta}_n - \theta_\gamma) = O_p(1)$.

Покажем, что при малых γ

$$n^{-1}l_n^S(\hat{\beta}_n) \xrightarrow{P_\beta} \Lambda_S(\gamma, \theta_\gamma). \quad (6.7.18)$$

Действительно, пусть $\delta > 0$ и γ таковы, что $|\beta| < 1 - \delta$, и $|\theta_\gamma| < 1 - \delta$. Тогда для любого $\varepsilon > 0$

$$\begin{aligned} P_\beta \{ |n^{-1}l_n^S(\hat{\beta}_n) - \Lambda_S(\gamma, \theta_\gamma)| > 2\varepsilon \} &\leq P_\beta \{ |n^{-1}l_n^S(\hat{\beta}_n) - \Lambda_S(\gamma, \hat{\beta}_n)| > \varepsilon \} + \\ &+ P_\beta \{ |\Lambda_S(\gamma, \hat{\beta}_n) - \Lambda_S(\gamma, \theta_\gamma)| > \varepsilon \} \leq \\ &\leq P_\beta \{ |n^{-1}l_n^S(\hat{\beta}_n) - \Lambda_S(\gamma, \hat{\beta}_n)| > \varepsilon, |\hat{\beta}_n| \leq 1 - \delta \} + P_\beta \{ |\hat{\beta}_n| > 1 - \delta \} + \\ &+ P_\beta \{ |\Lambda_S(\gamma, \hat{\beta}_n) - \Lambda_S(\gamma, \theta_\gamma)| > \varepsilon \} \leq \\ &\leq P_\beta \left\{ \sup_{|\theta| \leq 1 - \delta} |n^{-1}l_n^S(\theta) - \Lambda_S(\gamma, \theta)| > \varepsilon \right\} + P_\beta \{ |\hat{\beta}_n| > 1 - \delta \} + \\ &+ P_\beta \{ |\Lambda_S(\gamma, \hat{\beta}_n) - \Lambda_S(\gamma, \theta_\gamma)| > \varepsilon \}. \end{aligned}$$

Все три последние вероятности сколь угодно малы при достаточно больших n : первая — в силу соотношения (6.7.4), вторая — в силу условия 6.7.1 и неравенства $|\theta_\gamma| < 1 - \delta$, третья — из-за непрерывности по θ функции $\Lambda_S(\gamma, \theta)$ и того же условия 6.7.1. Итак, справедливость (6.7.18) доказана.

Из (6.7.17)–(6.7.18) и условий 6.7.1, 6.7.2 следует, что

$$\beta_{n,S}^* \xrightarrow{P_\beta} \theta_\gamma^{*S} := \theta_\gamma - (\theta_\gamma^e)^{-1} \Lambda_S(\gamma, \theta_\gamma). \quad (6.7.19)$$

Последнее соотношение, непрерывная дифференцируемость $\Lambda_S(\gamma, \theta_\gamma)$, условия 6.7.1, 6.7.2 и тот факт, что

$$\Lambda_S(0, \beta) = 0, \quad IF(\theta_\gamma^S, \mu_\xi) = -\lambda_S^{-1}(\beta) \frac{\partial}{\partial \gamma} \Lambda_S(0, \beta),$$

влекут

$$\begin{aligned} \theta_\gamma^{*S} &= \theta_\gamma - [\lambda_S^{-1}(\beta) + o(1)] \times \\ &\times \left[\Lambda_S(0, \beta) + \frac{\partial \Lambda_S(0, \beta)}{\partial \gamma} \gamma + \frac{\partial \Lambda_S(0, \beta)}{\partial \theta} (\theta_\gamma - \beta) + o(\gamma) \right] = \\ &= \theta_\gamma + IF(\theta_\gamma^S, \mu_\xi) \gamma - IF(\theta_\gamma, \mu_\xi) \gamma + o(\gamma). \end{aligned}$$

Последнее соотношение, определение (5.6.2) и условие 6.7.1 влекут

$$IF(\theta_\gamma^{*S}, \mu_\xi) := \lim_{\gamma \rightarrow 0} \frac{\theta_\gamma^{*S} - \theta_0}{\gamma} = IF(\theta_\gamma^S, \mu_\xi),$$

где функционал $IF(\theta_\gamma^S, \mu_\xi)$ определен в теореме 6.7.1. Итак, оценки $\hat{\beta}_{n,S}$ и $\beta_{n,S}^*$ имеют одинаковые функционалы влияния, и, следовательно, $\beta_{n,S}^*$ слабо зависит от выбросов в данных.

§ 6.8. Результаты моделирования: оценивание квантилей, доверительное оценивание, засоренные выборки

В этом параграфе мы рассматриваем схему (6.1.1) и приводим результаты численных экспериментов по вычислению квантилей $c_n^\alpha(\beta_0)$ из § 6.2, доверительному оцениванию неизвестного параметра β схемы (6.1.1) при умеренных и больших n , а также численным экспериментом иллюстрируем устойчивость знаковых оценок к выбросам в наблюдениях. Начнем с квантилей.

6.8.1. Оценивание квантилей.

О п р е д е л е н и е. Будем называть α -квантилью, $0 < \alpha < 1$, дискретной случайной величины с функцией распределения $F(x)$ число ζ_α , полученное следующим образом:

если уравнение $F(x) = \alpha$ не имеет решения, то

$$\zeta_\alpha = \sup\{x: F(x) < \alpha\};$$

если же уравнение $F(x) = \alpha$ разрешимо, то ζ_α — середина интервала решений этого уравнения.

В силу результатов § 6.2 при гипотезе $H_0: \beta = \beta_0$ и условии 6.1.1 знаковая тестовая статистика $l_n^S(\beta_0)$, определенная соотношениями (6.2.10), (6.2.9), свободна от $G(x)$. То же верно, разумеется, и для нормированной статистики $T_{n,S}(\beta_0) = n^{-1/2} \sqrt{1 - \beta_0^2} l_n^S(\beta_0)$, которая распределена при H_0 и условии 6.1.1 асимптотически как стандартная гауссовская величина. Для конечных n распределение $T_{n,S}(\beta_0)$ при гипотезе и условии 6.1.1 совпадает с распределением величины

$$\nu_n(\beta_0) = \sqrt{\frac{1 - \beta_0^2}{n}} \sum_{t=1}^{n-1} \beta_0^{t-1} \sum_{k=t+1}^n \xi_{k-t} \xi_k,$$

где ξ_1, \dots, ξ_n — независимые, одинаково распределенные случайные величины, $P\{\xi_i = 1\} = P\{\xi_i = -1\} = 1/2$. Это распределение, а также квантили $c_n^\alpha(\beta_0)$ уровня α величины $\nu_n(\beta_0)$ можно при любом β_0 вычислить методом статистических испытаний.

(6.8.2). Представим статистику $l_n^S(\beta, \eta)$ из (6.8.1) в виде

$$l_n^S(\beta, \eta) = \sum_{t=0}^{n-2} a_t \beta^t,$$

где коэффициенты полинома $l_n^S(\beta, \eta)$ есть $a_t = \sum_{k=t+1}^{n-1} \prod_{r=k-t}^k \eta_r$. Объединяя слагаемые при четных и нечетных степенях β , запишем статистику $\nu_n(\beta, \eta)$ в виде суммы двух слагаемых

$$\nu_n(\beta, \eta) = \nu_{n1}(\beta, \eta) + \nu_{n2}(\beta, \eta),$$

где

$$\nu_{n1}(\beta, \eta) := \sqrt{\frac{1-\beta^2}{n}} \sum_{t=0}^{[(n-2)/2]} \beta^{2t} \sum_{k=2t+1}^{n-1} \prod_{r=k-2t}^k \eta_r, \quad (6.8.4)$$

$$\nu_{n2}(\beta, \eta) := \sqrt{\frac{1-\beta^2}{n}} \sum_{t=0}^{[(n-3)/2]} \beta^{2t+1} \sum_{k=2t+2}^{n-1} \prod_{r=k-2t-1}^k \eta_r. \quad (6.8.5)$$

Тогда

$$\begin{aligned} & F_n(x+0, \beta) + F_n(-x, -\beta) = \\ & = \mathbf{P}_\beta \{ \nu_{n1}(\beta, \eta) + \nu_{n2}(\beta, \eta) \leq x \} + \mathbf{P}_\beta \{ \nu_{n1}(-\beta, \eta) + \nu_{n2}(-\beta, \eta) < -x \} = \\ & = \mathbf{P}_\beta \{ \nu_{n1}(\beta, \eta) + \nu_{n2}(\beta, \eta) \leq x \} + \mathbf{P}_\beta \{ -\nu_{n1}(\beta, \eta) + \nu_{n2}(\beta, \eta) > x \} = \\ & = \frac{1}{2^{n-1}} \sum_{k=1}^{2^{n-1}} \{ I(\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) \leq x) + \\ & \quad + I(-\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) > x) \}, \quad (6.8.6) \end{aligned}$$

где $\nu_{n1}(\beta, \eta^{(k)})$ и $\nu_{n2}(\beta, \eta^{(k)})$ — значения статистик $\nu_{n1}(\beta, \eta)$ и $\nu_{n2}(\beta, \eta)$ на одной из 2^{n-1} возможных комбинаций значений последовательности $\eta_1, \eta_2, \dots, \eta_{n-1}$, $\eta^{(k)} := (\eta_1^{(k)}, \eta_2^{(k)}, \dots, \eta_{n-1}^{(k)})$, а суммирование производится по всем возможным комбинациям. Теперь (6.8.6) тождественно равно

$$\begin{aligned} & \frac{1}{2} \frac{1}{2^{n-1}} \sum_{k=1}^{2^{n-1}} \{ I(\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) \leq x) + \\ & \quad + I(-\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) > x) + \\ & \quad + I(\nu_{n1}(\beta, -\eta^{(k)}) + \nu_{n2}(\beta, -\eta^{(k)}) \leq x) + \\ & \quad + I(-\nu_{n1}(\beta, -\eta^{(k)}) + \nu_{n2}(\beta, -\eta^{(k)}) > x) \}. \quad (6.8.7) \end{aligned}$$

Коэффициенты при четных степенях β в определении $\nu_{n1}(\beta, \eta^{(k)})$ (6.8.4) содержат нечетное число сомножителей из последовательности $\eta^{(k)}$, а при нечетных степенях β в определении $\nu_{n2}(\beta, \eta^{(k)})$ (6.8.5) — четное число сомножителей. Значит,

$$\nu_{n1}(\beta, -\eta^{(k)}) = -\nu_{n1}(\beta, \eta^{(k)}), \quad \nu_{n2}(\beta, -\eta^{(k)}) = \nu_{n2}(\beta, \eta^{(k)}).$$

Поэтому (6.8.7) равно

$$\begin{aligned} \frac{1}{2} \frac{1}{2^{n-1}} \sum_{k=1}^{2^{n-1}} \{ & I(\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) \leq x) + \\ & + I(-\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) > x) + \\ & + I(-\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) \leq x) + \\ & + I(\nu_{n1}(\beta, \eta^{(k)}) + \nu_{n2}(\beta, \eta^{(k)}) > x) \} = 1. \end{aligned}$$

Таким образом, (6.8.3), а следовательно, и (6.8.2), доказано.

Т а б л и ц а 6.8.1

Квантили $c_{\alpha}^*(\beta_0)$ ($\alpha = 0.05$)

β_0	n			
	50	100	150	200
-0.9	-1.85	-1.85	-1.84	-1.82
-0.8	-1.83	-1.81	-1.79	-1.77
-0.7	-1.81	-1.78	-1.75	-1.75
-0.6	-1.78	-1.75	-1.74	-1.73
-0.5	-1.75	-1.73	-1.71	-1.71
-0.4	-1.73	-1.71	-1.70	-1.69
-0.3	-1.70	-1.69	-1.68	-1.68
-0.2	-1.68	-1.67	-1.67	-1.66
-0.1	-1.65	-1.65	-1.65	-1.65
0.0	-1.56	-1.70	-1.71	-1.63
0.1	-1.60	-1.62	-1.63	-1.63
0.2	-1.57	-1.60	-1.61	-1.62
0.3	-1.55	-1.58	-1.60	-1.60
0.4	-1.52	-1.56	-1.57	-1.58
0.5	-1.47	-1.53	-1.55	-1.57
0.6	-1.42	-1.50	-1.53	-1.54
0.7	-1.35	-1.45	-1.49	-1.51
0.8	-1.26	-1.37	-1.43	-1.46
0.9	-1.07	-1.23	-1.31	-1.35

Используя (6.8.2), можно моделировать квантили $c_n^\alpha(\beta)$ лишь при $\beta \geq 0$, а их значения при $\beta < 0$ получать центральной симметрией относительно нуля. Аналогично, зная $c_n^\alpha(\beta)$ при $\beta \in (-1, 1)$, можно находить $c_n^{1-\alpha}(\beta)$. Для $n = 50(50), 200$, $\beta_0 = 0, \pm 0.1, \pm 0.2, \dots, \pm 0.9$ и $\alpha = 0.05$ мы приводим результаты вычислений $c_n^\alpha(\beta_0)$ в табл. 6.8.1. Для этого было выполнено по 60000 моделирований величины $\nu_n(\beta_0)$ в каждой точке β_0 заданной сетки.

Анализируя данные табл. 6.8.1, видим, что скорость сходимости квантили случайной величины $\nu_n(\beta_0)$ к квантили гауссовской $N(0, 1)$ случайной величины (равной 1.65) существенно зависит от β_0 : при β_0 , близких к ± 1 , она хуже, чем при малых β_0 . Можно показать, что эта сходимость равномерная при $|\beta_0| \leq 1 - \delta$, $0 < \delta < 1$.

Представление о распределении случайной величины $\nu_n(\beta_0)$ при разных β_0 и о скорости сходимости к нормальному закону дают также рис. 6.8.1 и 6.8.2. На рис. 6.8.1 приведены графики функции распределения $F_n(x, \beta_0)$ величины $\nu_n(\beta_0)$ для $\beta_0 = \pm 0.5$ и $n = 100$, а на рис. 6.8.2 — те же графики в нормальном масштабе. (Функции распределения получены также моделированием 60000 значений случайной величины $\nu_n(\beta_0)$). Отличие от гауссовской кривой явное.

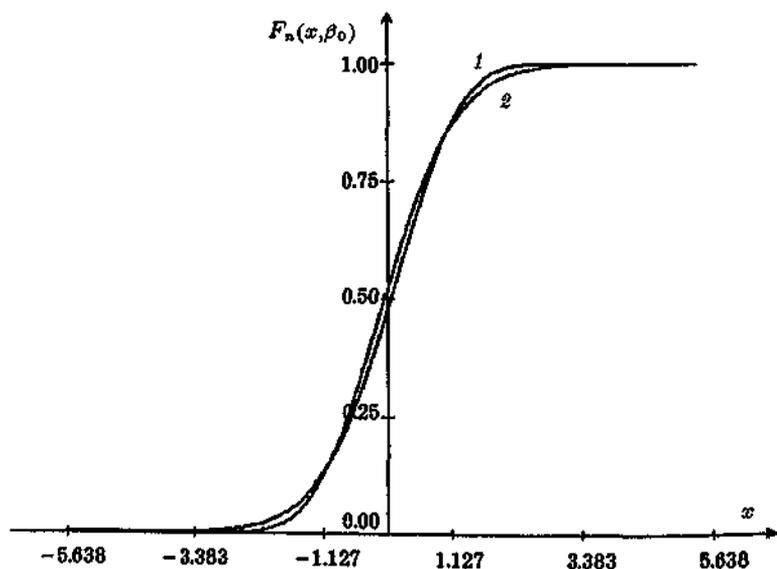


Рис. 6.8.1. Функции распределения $\nu_n(\beta_0)$, $n = 100$; 1 — $\beta_0 = -0.5$; 2 — $\beta_0 = 0.5$

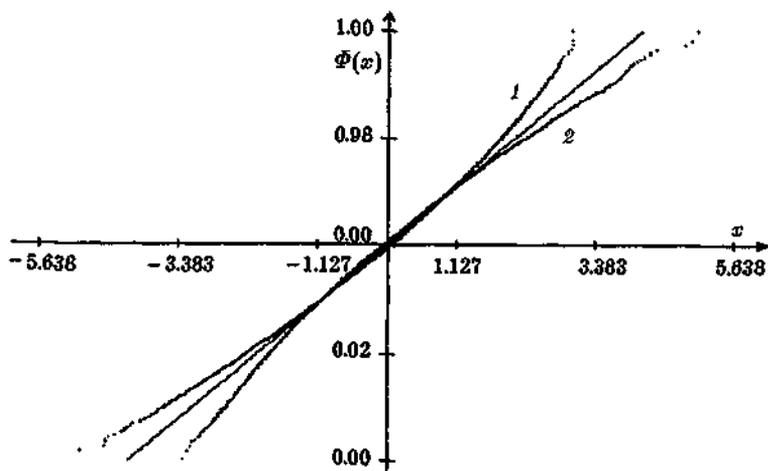


Рис. 6.8.2. Функции распределения $\nu_n(\beta_0)$ в нормальном масштабе, $n = 100$;
1 — $\beta_0 = -0.5$; 2 — $\beta_0 = 0.5$

Хотя представленные результаты отчетливо демонстрируют справедливость нормальной аппроксимации для тестовых знаковых статистик и их квантилей, разумная точность этой аппроксимации достижима лишь при объемах выборок n порядка нескольких сотен. Это лишний раз свидетельствует в пользу непараметрических точных методов проверки гипотез и оценивания (разумеется, там, где это возможно). К точному доверительному оцениванию параметра β мы и переходим.

6.8.2. Доверительное оценивание параметра β . Пусть

$$T_{n,S}(\theta) = n^{-1/2} \sqrt{1 - \theta^2} I_n^S(\theta),$$

где

$$I_n^S(\theta) = \sum_{i=1}^{n-1} \theta^{i-1} \sum_{i=i+1}^n \text{sign}((u_i - \theta u_{i-1})(u_{i-1} - \theta u_{i-2})),$$

а наблюдения u_i подчиняются модели (6.1.1). При единственном условии 6.1.1 множество

$$\mathbb{D}_{n\alpha} = \{\theta: c_n^\alpha(\theta) \leq T_{n,S}(\theta) \leq c_n^{1-\alpha}(\theta)\}$$

есть доверительное множество для параметра β надежности не меньше $1 - 2\alpha$.

Квантили $c_n^\alpha(\theta)$ и $c_n^{1-\alpha}(\theta)$ могут быть оценены на сетке значений θ , $|\theta| < 1$, как это описано в п. 6.8.1. Строя одновременно кривые $c_n^\alpha(\theta)$, $c_n^{1-\alpha}(\theta)$ и $T_{n,S}(\theta)$, можно построить явно множество $\mathbb{D}_{n\alpha}$.

Результаты моделирования для $n = 100, 200, 400$ и двух типов распределения ошибок — гауссовского и Коши — представлены на рис. 6.8.3–6.8.8. На каждом рисунке изображены три независимые реализации (помечены метками *, + и x) случайного процесса $T_{n,S}(\theta)$, построенного по наблюдениям u_0, u_1, \dots, u_n при $\beta = 0.2$, и кривые квантилей $c_n^\alpha(\theta)$, $c_n^{1-\alpha}(\theta)$ при $\alpha = 0.05$. Соответствующие доверительные множества для β — это множества тех значений θ , при которых процесс $T_{n,S}(\theta)$ находится в полосе между кривыми $c_n^\alpha(\theta)$ и $c_n^{1-\alpha}(\theta)$. Для того чтобы было удобно сравнивать результаты, полученные для различных объемов выборок и различных типов распределений, мы построили все графики в одном масштабе. За каждым графиком непосредственно следует таблица точечных и доверительных оценок параметра β . Точечная оценка получена решением уравнения $l_n^S(\theta) \div 0$, $|\theta| < 1$.

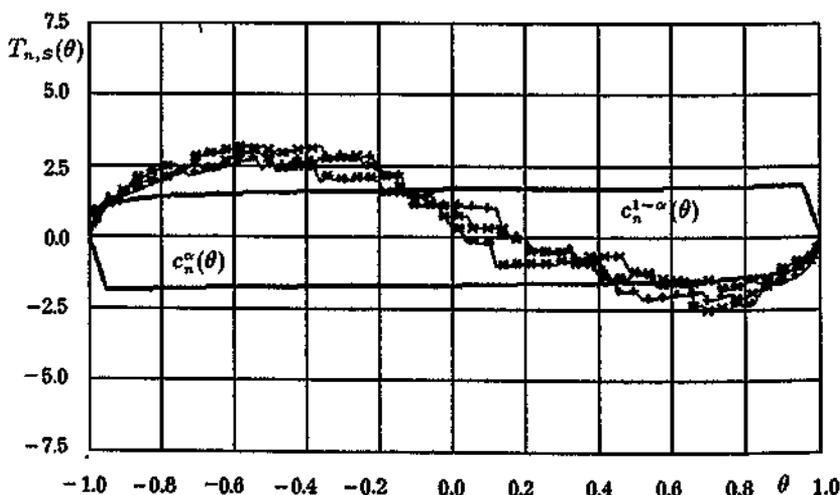


Рис. 6.8.3. Гауссовское $N(0,1)$ -распределение ошибок; $n = 100$, $\beta = 0.2$

Т а б л и ц а 6.8.2

Номер реализации	Точечная оценка	90-процентный доверительный интервал
1 (*)	0.04	(-0.18, 0.54)
2 (+)	0.19	(-0.13, 0.45)
3 (x)	0.15	(-0.13, 0.56)

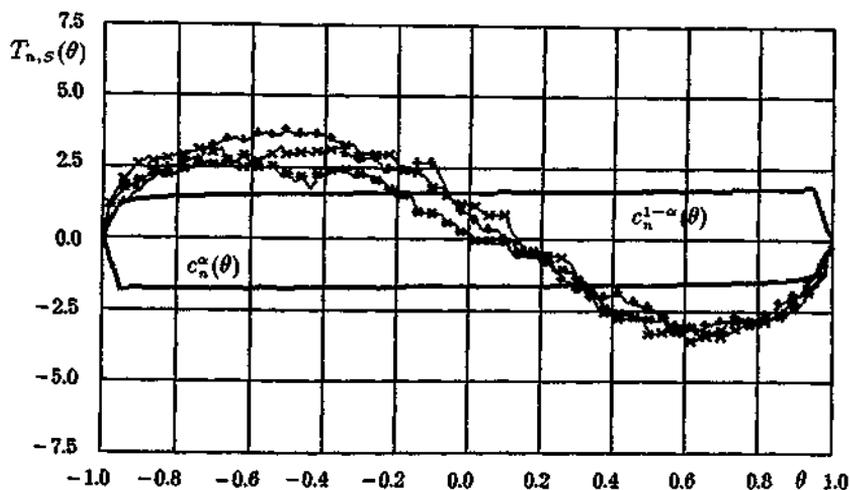
Рис. 6.8.4. Гауссовское $N(0,1)$ -распределение ошибок; $n = 200$, $\beta = 0.2$

Таблица 6.8.3

Номер реализации	Точечная оценка	90-процентный доверительный интервал
1 (*)	0.02	(-0.21, 0.32)
2 (+)	0.11	(-0.06, 0.30)
3 (x)	0.14	(-0.06, 0.33)

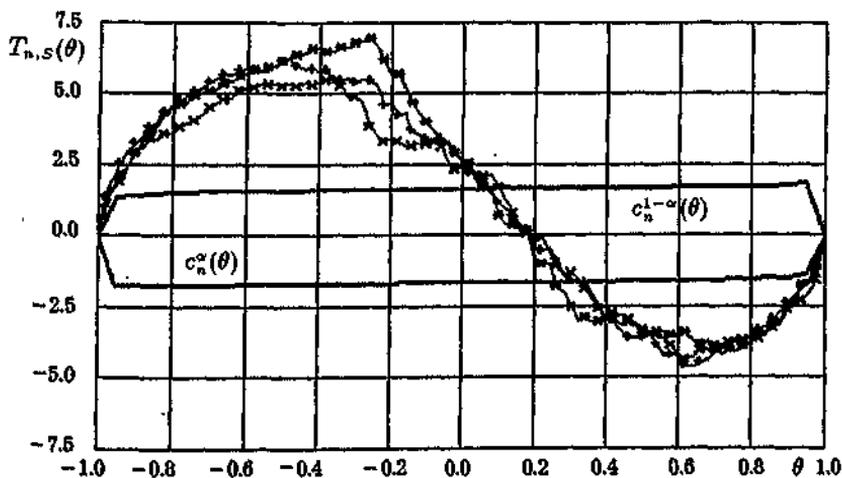
Рис. 6.8.5. Гауссовское $N(0,1)$ -распределение ошибок; $n = 400$, $\beta = 0.2$

Таблица 6.8.4

Номер реализации	Точечная оценка	90-процентный доверительный интервал
1 (*)	0.18	(0.09, 0.25)
2 (+)	0.19	(0.08, 0.30)
3 (x)	0.19	(0.10, 0.32)

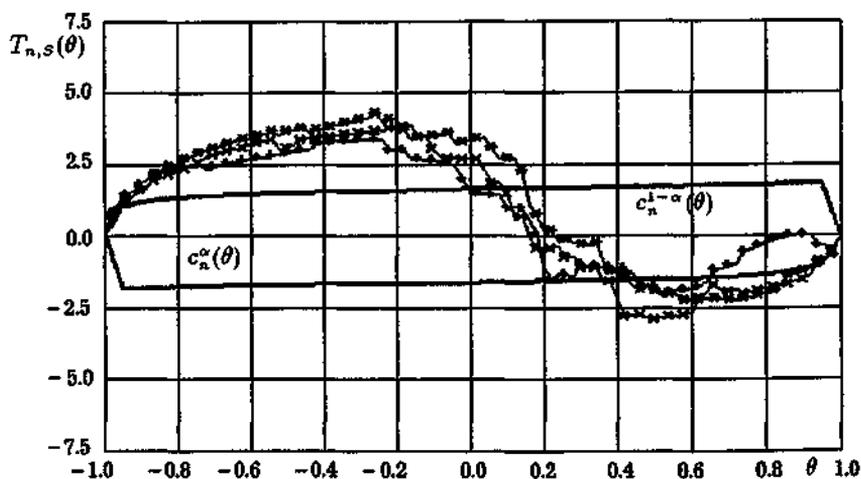
Рис. 6.8.6. Распределение ошибок Коши (0,1); $n = 100$, $\beta = 0.2$

Таблица 6.8.5

Номер реализации	Точечная оценка	90-процентный доверительный интервал
1 (*)	0.21	(0.15, 0.44)
2 (+)	0.19	(0.00, 0.46) ∪ (0.64, 1.0)
3 (x)	0.15	(0.09, 0.35)

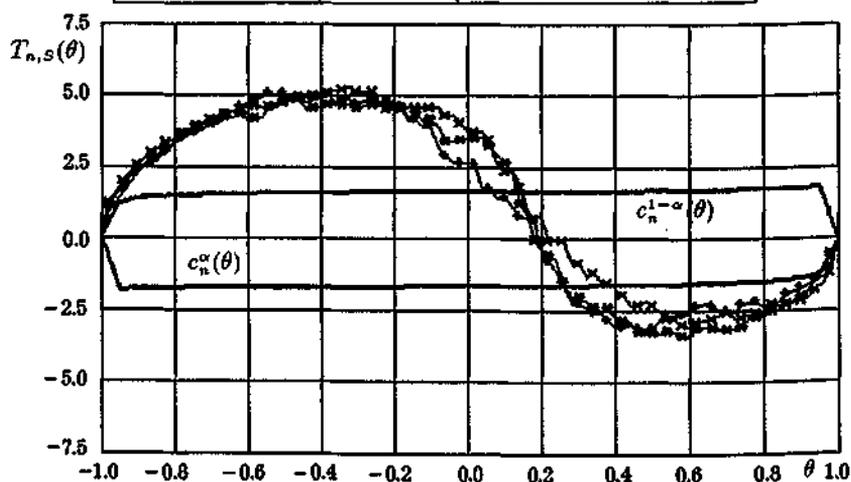
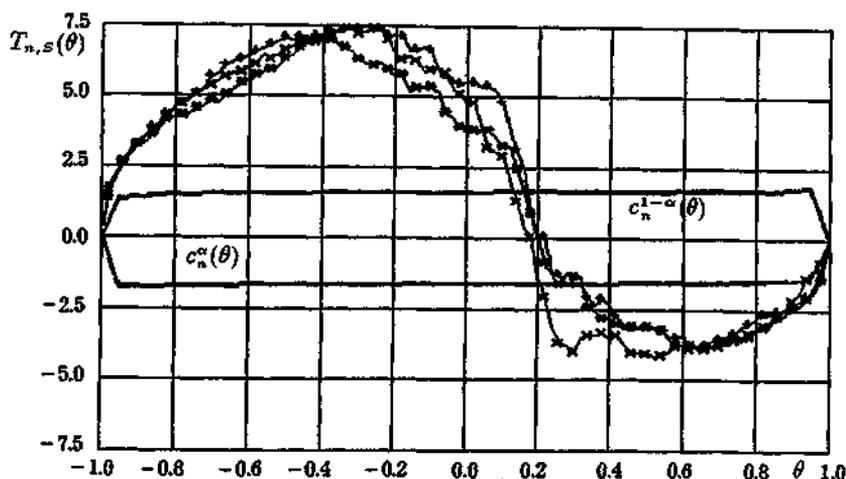
Рис. 6.8.7. Распределение ошибок Коши (0,1); $n = 200$, $\beta = 0.2$

Таблица 6.8.6

Номер реализации	Точечная оценка	90-процентный доверительный интервал
1 (*)	0.18	(0.12, 0.28)
2 (+)	0.20	(0.07, 0.26)
3 (x)	0.22	(0.15, 0.39)

Рис. 6.8.8. Распределение ошибок Коши (0, 1); $n = 400$, $\beta = 0.2$

Т а б л и ц а 6.8.7

Номер реализации	Точечная оценка	90-процентный доверительный интервал
1 (*)	0.20	(0.16, 0.32)
2 (+)	0.20	(0.18, 0.28)
3 (x)	0.18	(0.14, 0.20)

По результатам моделирования можно сделать следующие выводы.

Во-первых, для уверенного оценивания параметра авторегрессии необходимо не несколько десятков (как в случае независимых наблюдений), а несколько сотен наблюдений. Так, из рис. 6.8.6 видно, что при $n = 100$ среди смоделированных есть траектории, хвосты которых попадают внутрь полосы $(c_n^\alpha(\theta), c_n^{1-\alpha}(\theta))$, что соответствует появлению дополнительного ложного доверительного интервала параметра β . При $n = 400$ все три траектории находятся в полосе $(c_n^\alpha(\theta), c_n^{1-\alpha}(\theta))$, когда параметр θ лежит в окрестности истинного параметра β .

Во-вторых, утяжеление «хвостов» в распределении ошибок приводит к увеличению точности в оценивании параметра β . Так, кривая $T_{n,s}(\theta)$, построенная по наблюдениям с ошибками, распределенными по Коши, спадает «круче» в окрестности истинного параметра β , чем кривая, построенная по наблюдениям с гауссовскими ошибками (ср. рис. 6.8.3 и 6.8.6; рис. 6.8.4 и 6.8.7; рис. 6.8.5 и 6.8.8). А поскольку от «крутизны» траектории $T_{n,s}(\theta)$ зависит длина доверительного интервала, то можно сказать, что при одном и том же числе наблюдений доверительные интервалы при ошибках, подчиняющихся распределению Коши, получаются, как правило, уже, чем при гауссовских ошибках. Обратим внимание: в случае ошибок $\{\epsilon_i\}$, имеющих распределение Коши, асимптотические доверительные интервалы, основанные на общеупотребительной оценке наименьших квадратов, построить

нельзя. В то же время наши знаковые правила не только применимы, но дают (качественно) вполне приемлемые результаты.

6.8.3. Знаковое оценивание по засоренным выборкам. Сравним теперь знаковые оценки и оценки наименьших квадратов (ОНК) в случае засоренных наблюдений. Предположим, что вместо $\{u_i\}$ в схеме (6.1.1) мы наблюдаем процесс $\{y_i\}$, $i \in \mathbb{Z}$, из (6.7.1) для $\beta = 0.2$, $\gamma = 0.1$, $\varepsilon_i \sim N(0, 1)$, $\xi_i \sim N(0, 25)$. Насколько сильно ухудшает качество оценки присутствие грубых ошибок? Чтобы составить представление об этом, мы провели моделирование процесса $\{y_i\}$.

На рис. 6.8.9 для $n = 400$ приведены графики трех независимых реализаций процесса $T_{n,s}(\theta, Y_n) = n^{-1/2} \sqrt{1 - \theta^2} l_n^S(\theta, Y_n)$. Процесс $l_n^S(\theta, Y_n)$ строится по $Y_n = (y_0, \dots, y_n)$ так же, как $l_n^S(\theta)$ строится по u_0, \dots, u_n .

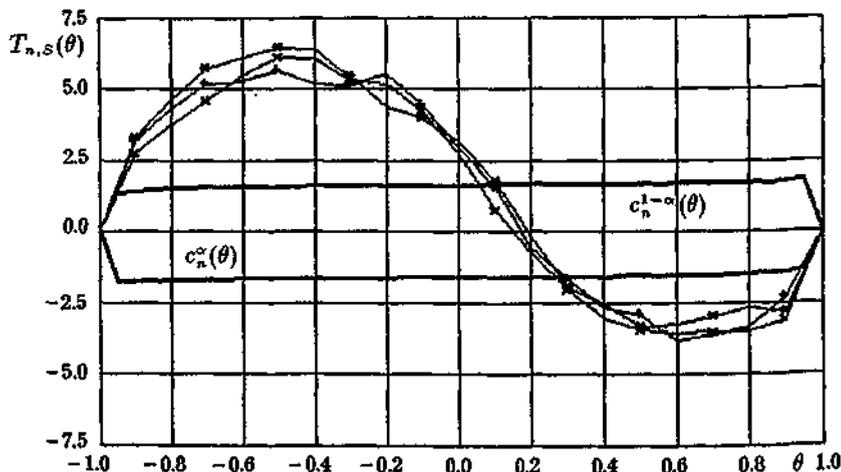


Рис. 6.8.9. Засоренные данные; $\gamma = 0.1$, $\varepsilon_i \sim N(0, 1)$, $\xi_i \sim N(0, 25)$, $n = 400$, $\beta = 0.2$

В табл. 6.8.8 мы приводим одновременно значения знаковых оценок $\hat{\beta}_{n,s}$ и оценок наименьших квадратов, построенные по этим засоренным и соответствующим незасоренным (т.е. при $\gamma = 0$) данным. Мы приводим также 90-процентные доверительные интервалы, причем в случае оценок наименьших квадратов это асимптотические доверительные интервалы из (5.3.9). Разумеется, при наличии засорения доверительные интервалы нельзя строить по стандартной методике, так как квантили статистики $T_{n,s}(\theta, Y_n)$ вовсе не равны $c_n^\alpha(\theta)$, а оценка наименьших квадратов не является \sqrt{n} -асимптотически нормальной. Тем не менее мы привели именно результаты стандартных расчетов (формально некорректных) с тем, чтобы качественно продемонстрировать характер изменений при наличии засорения.

Все численные результаты убеждают в высокой устойчивости знаковых оценок по отношению к грубым выбросам. Они же подтверждают тот факт, что оценки наименьших квадратов весьма чувствительны к выбросам.

Т а б л и ц а 6.8.8

Метод оценивания	Точечная оценка	90-процентный доверительный интервал
МНК 1 (*)	0.06	(-0.02, 0.14)
МНК 1 ($\gamma=0$) (*)	0.21	(0.13, 0.29)
Sign 1 (*)	0.20	(0.11, 0.29)
Sign 1 ($\gamma=0$) (*)	0.26	(0.11, 0.41)
МНК 2 (+)	-0.03	(-0.11, 0.05)
МНК 2 ($\gamma=0$) (+)	0.20	(0.12, 0.28)
Sign 2 (+)	0.16	(0.08, 0.28)
Sign 2 ($\gamma=0$) (+)	0.16	(0.13, 0.25)
МНК 3 (x)	0.10	(0.01, 0.18)
МНК 3 ($\gamma=0$) (x)	0.23	(0.14, 0.31)
Sign 3 (x)	0.15	(0.04, 0.28)
Sign 3 ($\gamma=0$) (x)	0.23	(0.10, 0.33)

Мы подробно обсуждали в § 5.6 и 6.7 как численно охарактеризовать устойчивость оценок с помощью функционалов влияния и чувствительностей к большим ошибкам. Для знаковой оценки $\hat{\beta}_n, s$ эти характеристики вычислены в п. 6.7.1 (см. соотношения (6.7.13)–(6.7.15)).

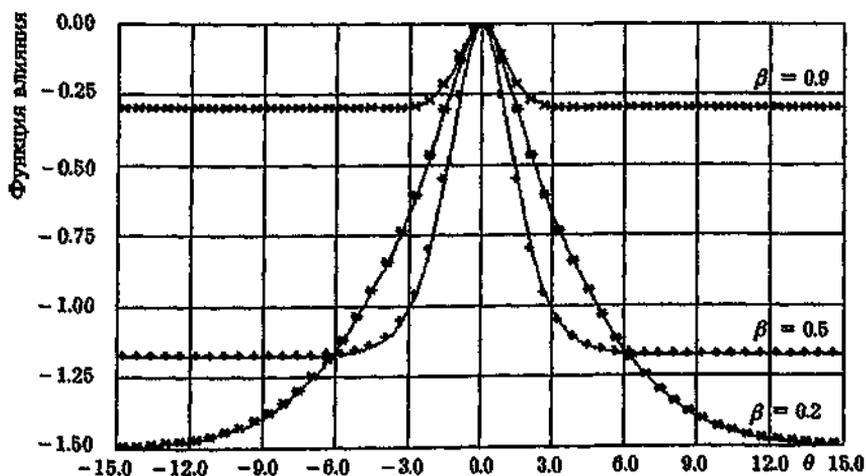


Рис. 6.8.10. Функции влияния знаковой оценки; * — $\beta=0.2$; + — $\beta=0.5$; x — $\beta=0.9$

На рис. 6.8.10 изображены графики функций $IF(\theta_\gamma^S, \xi)$ при трех различных значениях параметра β в зависимости от величины случайного засорения ξ для стандартной гауссовской $G(x)$. Как видно, эти функции ограничены в согласии с конечной чувствительностью знаковых оценок.

§ 6.9. Приложение. Доказательство теоремы 6.4.1

Из определения (6.2.10) следует

$$n^{-1/2} l_n^S(\beta_n + n^{-1/2}\theta) - n^{-1/2} l_n^S(\beta_n) = \\ = n^{-1/2} \sum_{t=1}^{n-1} (\beta_n + n^{-1/2}\theta)^{t-1} [\Gamma_{tn}(\beta_n + n^{-1/2}\theta) - \Gamma_{tn}(\beta_n)] + \quad (6.9.1)$$

$$+ n^{-1/2} \sum_{t=1}^{n-1} [(\beta_n + n^{-1/2}\theta)^{t-1} - \beta_n^{t-1}] \Gamma_{tn}(\beta_n). \quad (6.9.2)$$

Рассмотрим сначала (6.9.2). Далее индекс β_n у символов вероятности и математического ожидания будет опущен.

Лемма 6.9.1. Пусть верна альтернатива $H_{1n}(a)$, выполнено условие 6.1.1, $0 < \Theta < \infty$.

Тогда супремум (6.9.2) по $|\theta| \leq \Theta n^r$, $r < 1/2$, есть $o_p(1)$.

Доказательство. Обозначим для краткости $\Theta_n = \Theta n^r$. Для b такого, что $|\beta_0| < b < 1$ и $\forall \varepsilon > 0$ при $n > n_0$

$$\mathbf{P} \left\{ \sup_{|\theta| \leq \Theta_n} \left| n^{-1/2} \sum_{t=1}^{n-1} [(\beta_n + n^{-1/2}\theta)^{t-1} - \beta_n^{t-1}] \Gamma_{tn}(\beta_n) \right| \geq \varepsilon \right\} \leq \\ \leq \mathbf{P} \left\{ \Theta_n n^{-1} \sup_{|\theta| \leq \Theta_n} \sum_{t=2}^{n-1} \sum_{s=0}^{t-2} |(\beta_n + n^{-1/2}\theta)|^{t-2-s} |\beta_n|^s |\Gamma_{tn}(\beta_n)| \geq \varepsilon \right\} \leq \\ \leq \mathbf{P} \left\{ \Theta_n n^{-1} \sum_{t=2}^{n-1} (t-1) b^{t-2} |\Gamma_{tn}(\beta_n)| \geq \varepsilon \right\}.$$

Последняя вероятность в силу неравенства Чебышева не больше

$$\varepsilon^{-1} \Theta_n n^{-1} \sum_{t=2}^{n-1} (t-1) b^{t-2} \mathbf{E} |\Gamma_{tn}(\beta_n)| = O(n^{-1/2+r}) = o(1),$$

так как $n^{-1/2} \mathbf{E} |\Gamma_{tn}(\beta_n)| \leq n^{-1/2} (\mathbf{E} \Gamma_{tn}^2(\beta_n))^{1/2} = ((n-t)/n)^{1/2} \leq 1$ в силу (6.2.14). Лемма 6.9.1 доказана. \square

Рассмотрим (6.9.1). Пусть

$$\Delta_k(x) = \begin{cases} 1, & \varepsilon_k < x, \\ 1/2, & \varepsilon_k = x, \\ 0, & \varepsilon_k > x. \end{cases} \quad (6.9.3)$$

Обозначим для краткости $\eta_n = n^{-1/2}\theta$. Тогда

$$\begin{aligned} S_k(\beta_n + n^{-1/2}\theta) &= \text{sign}(u_k - (\beta_n + n^{-1/2}\theta)u_{k-1}) = \\ &= \text{sign}(\varepsilon_k - \eta_n u_{k-1}) = 1 - 2\Delta_k(\eta_n u_{k-1}). \end{aligned}$$

Отсюда и из определения $\Gamma_{tn}(\theta)$ в (6.2.9) следует представление

$$\begin{aligned} n^{-1/2}\Gamma_{tn}(\beta_n + n^{-1/2}\theta) - n^{-1/2}\Gamma_{tn}(\beta_n) &= \\ &= 4x_{tn}(\eta_n) + 4y_{tn}(\eta_n) - 2u_{tn}(\eta_n), \quad (6.9.4) \end{aligned}$$

где

$$\begin{aligned} x_{tn}(\eta_n) &= n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1})\Delta_k(\eta_n u_{k-1}) - \\ &\quad - \Delta_{k-t}(\eta_n u_{k-t-1})G(\eta_n u_{k-1}) - \Delta_{k-t}(0)\Delta_k(0) + \Delta_{k-t}(0)G(0)], \\ y_{tn}(\eta_n) &= n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1})G(\eta_n u_{k-1}) - \Delta_{k-t}(0)G(0)], \\ u_{tn}(\eta_n) &= n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)] + \\ &\quad + n^{-1/2} \sum_{k=t+1}^n [\Delta_k(\eta_n u_{k-1}) - \Delta_k(0)]. \end{aligned}$$

С помощью (6.9.4) перепишем (6.9.1) в виде

$$\sum_{i=1}^{n-1} (\beta_n + \eta_n)^{t-1} [4x_{tn}(\eta_n) + 4y_{tn}(\eta_n) - 2u_{tn}(\eta_n)]. \quad (6.9.5)$$

Л е м м а 6.9.2. Пусть выполнены условия 6.1.2, 6.1.3, верна альтернатива $H_{1n}(a)$, $r > 1/4$ и $\Theta_n = \Theta_n^{-r}$, $N_n \sim \log_b n^{-1}$, где $|\beta_0| < b < 1$.

Тогда:

- 1) $\sup_{t \leq N_n} \sup_{|\eta_n| \leq \Theta_n} |x_{tn}(\eta_n)| = o_p(1)$;
- 2) $\sup_{|\eta_n| \leq \Theta_n} \left| n^{-1/2} \sum_{k=1}^n (\Delta_k(\eta_n u_{k-1}) - \Delta_k(0)) \right| = o_p(1)$.

Пункты 1) и 2) леммы доказываются единым методом, поэтому ограничимся доказательством п. 1).

Разделим отрезок $[-\Theta_n, \Theta_n]$ на 3^{m_n} частей точками

$$\eta_{sn} = -\Theta_n + 2\Theta_n 3^{-m_n} s, \quad s = 0, 1, \dots, 3^{m_n}.$$

Выберем r_1 так, что $0 < r_1 < r$ и $r + r_1 > 1/2$, пусть $3^{m_n} \sim n^{r_1}$. Введем величины

$$\hat{u}_{ks} = u_k(1 - 2\Theta_n 3^{-m_n} \eta_{sn}^{-1} I(u_k \leq 0)), \quad (6.9.6)$$

$$\tilde{u}_{ks} = u_k(1 - 2\Theta_n 3^{-m_n} \eta_{sn}^{-1} I(u_k \geq 0)). \quad (6.9.7)$$

Для любого $|\eta_n| \leq \Theta_n$ выберем среди точек $\{\eta_{sn}\}$ ту точку η_{ln} , для которой

$$0 \leq \eta_{ln} - \eta_n \leq \Theta_n 3^{-m_n}. \quad (6.9.8)$$

Из определений (6.9.6)–(6.9.7) и неравенства (6.9.8) следует

$$\eta_{ln} \tilde{u}_{kl} \leq \eta_n u_k \leq \eta_{ln} \hat{u}_{kl}, \quad k = 0, 1, \dots, n-1. \quad (6.9.9)$$

Введем векторы

$$\hat{U}_{sn} = (\hat{u}_{0s}, \dots, \hat{u}_{n-1,s}), \quad \tilde{U}_{sn} = (\tilde{u}_{0s}, \dots, \tilde{u}_{n-1,s}), \quad U_n = (u_0, \dots, u_{n-1}).$$

Чтобы подчеркнуть зависимость величины $x_{tn}(\eta_n)$ от компонент вектора U_n , переобозначим

$$x_{tn}(\eta_n) = x_{tn}(\eta_n, U_n).$$

Далее нам понадобятся величины $x_{tn}(\eta_{sn}, \hat{U}_{sn})$ и $x_{tn}(\eta_{sn}, \tilde{U}_{sn})$, получаемые формальной заменой в $x_{tn}(\eta_n, U_n)$ вектора U_n на \hat{U}_{sn} и \tilde{U}_{sn} , η_n на η_{sn} . Из монотонности функций $\Delta_k(y)$, $G(y)$ и (6.9.9) следуют неравенства

$$\begin{aligned} x_{tn}(\eta_n, U_n) &\leq x_{tn}(\eta_{ln}, \hat{U}_{ln}) + \\ &+ n^{-1/2} \sum_{k=i+1}^n [\Delta_{k-i}(\eta_{ln} \hat{u}_{k-i-1,i}) G(\eta_{ln} \hat{u}_{k-1,i}) - \\ &\quad - \Delta_{k-i}(\eta_{ln} \tilde{u}_{k-i-1,i}) G(\eta_{ln} \tilde{u}_{k-1,i})], \\ x_{tn}(\eta_n, U_n) &\geq x_{tn}(\eta_{ln}, \tilde{U}_{ln}) - \\ &- n^{-1/2} \sum_{k=i+1}^n [\Delta_{k-i}(\eta_{ln} \hat{u}_{k-i-1,i}) G(\eta_{ln} \hat{u}_{k-1,i}) - \\ &\quad - \Delta_{k-i}(\eta_{ln} \tilde{u}_{k-i-1,i}) G(\eta_{ln} \tilde{u}_{k-1,i})]. \end{aligned}$$

Из двух последних неравенств получаем

$$\sup_{t \leq N_n} \sup_{|\eta_n| \leq \Theta_n} |x_{tn}(\eta_n, U_n)| \leq \sup_{t \leq N_n} \sup_{s \leq 3^{mn}} \{ |x_{tn}(\eta_{sn}, \hat{U}_{sn})| + |x_{tn}(\eta_{sn}, \tilde{U}_{sn})| \} + \quad (6.9.10)$$

$$+ \sup_{t \leq N_n} \sup_{s \leq 3^{mn}} n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_{sn} \hat{u}_{k-t-1, s}) \times \\ \times G(\eta_{sn} \hat{u}_{k-1, s}) - \Delta_{k-t}(\eta_{sn} \tilde{u}_{k-t-1, s}) G(\eta_{sn} \tilde{u}_{k-1, s})]. \quad (6.9.11)$$

Рассмотрим в (6.9.10) первое слагаемое $x_{tn}(\eta_{sn} \hat{U}_{sn})$. Обозначим для краткости

$$\xi_k = \xi_k(n, s, t) = \Delta_{k-t}(\eta_{sn} \hat{u}_{k-t-1, s}) \Delta_k(\eta_{sn} \hat{u}_{k-1, s}) - \\ - \Delta_{k-t}(\eta_{sn} \tilde{u}_{k-t-1, s}) G(\eta_{sn} \hat{u}_{k-1, s}) - \\ - \Delta_{k-t}(0) \Delta_k(0) + \Delta_{k-t}(0) G(0). \quad (6.9.12)$$

Тогда

$$x_{tn}(\eta_{sn}, \hat{U}_{sn}) = n^{-1/2} \sum_{k=t+1}^n \xi_k.$$

Пусть сигма-алгебра Ω_j порождается случайными величинами $\{\epsilon_i, i \leq j\}$. Тогда $\mathbf{E} \xi_k = \mathbf{E} \{ \mathbf{E}(\xi_k | \Omega_{k-1}) \} = 0$, так как очевидно, $\mathbf{E} \{ \xi_k | \Omega_{k-1} \} = 0$ п.н. (т.е. последовательность $\{\xi_k, \Omega_k\}, k = 1, 2, \dots$, — мартингал-разность). Аналогично при $k < j$

$$\mathbf{E} \xi_k \xi_j = \mathbf{E} \{ \mathbf{E}(\xi_k \xi_j | \Omega_{j-1}) \} = \mathbf{E} \{ \xi_k \mathbf{E}(\xi_j | \Omega_{j-1}) \} = 0.$$

Столь же простые вычисления дают

$$\mathbf{E} \xi_k^2 \leq 2 \sup_x g(x) \{ \mathbf{E} |\eta_{sn} \hat{u}_{k-1, s}| + \mathbf{E} |\eta_{sn} \tilde{u}_{k-t-1, s}| \} = O(n^{-r})$$

равномерно по k, s, t , так как $|\eta_{sn}| \leq \Theta_n = O(n^{-r})$, $|\hat{u}_{k-1, s}| \leq 3|u_{k-1}|$. Следовательно,

$$\sup_{t \leq N_n} \sup_{s \leq 3^{mn}} \mathbf{E} x_{tn}^2(\eta_{sn}, \hat{U}_{sn}) = O(n^{-r}). \quad (6.9.13)$$

Из (6.9.13) в силу неравенства Чебышева следует, что $\forall \epsilon > 0$

$$\mathbf{P} \left\{ \sup_{t \leq N_n} \sup_{s \leq 3^{mn}} |x_{tn}(\eta_{sn}, \hat{U}_{sn})| > \epsilon \right\} \leq \sum_{t=1}^{N_n} \sum_{s=0}^{3^{mn}} \mathbf{P} \{ |x_{tn}(\eta_{sn}, \hat{U}_{sn})| > \epsilon \} \leq$$

$$\leq \varepsilon^{-2} \sum_{t=1}^{N_n} \sum_{s=0}^{3^{m_n}} \mathbf{E} x_{tn}^2(\eta_{sn}, \hat{U}_{sn}) = O(N_n 3^{m_n} n^{-r}) = O(n^{r_1-r} \log n) = o(1).$$

Аналогично показывается, что

$$\sup_{t \leq N_n} \sup_{s \leq 3^{m_n}} |x_{tn}(\eta_{sn}, \tilde{U}_{sn})| = o_p(1).$$

Следовательно, (6.9.10) есть $o_p(1)$.

Рассмотрим теперь (6.9.11). Очевидно, (6.9.11) не больше

$$\sup_{t \leq N_n} \sup_{s \leq 3^{m_n}} n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_{sn} \hat{u}_{k-t-1,s}) - \Delta_{k-t}(\eta_{sn} \tilde{u}_{k-t-1,s})] G(\eta_{sn} \hat{u}_{k-1,s}) + \quad (6.9.14)$$

$$+ \sup_{t \leq N_n} \sup_{s \leq 3^{m_n}} n^{-1/2} \sum_{k=t+1}^n [G(\eta_{sn} \hat{u}_{k-1,s}) - G(\eta_{sn} \tilde{u}_{k-1,s})] \Delta_{k-t}(\eta_{sn} \tilde{u}_{k-t-1,s}). \quad (6.9.15)$$

В свою очередь (6.9.15) не больше

$$\begin{aligned} \sup_{s \leq 3^{m_n}} n^{-1/2} \sum_{k=1}^n [G(\eta_{sn} \hat{u}_{k-1,s}) - G(\eta_{sn} \tilde{u}_{k-1,s})] &\leq \\ &\leq \sup_{s \leq 3^{m_n}} \sup_x g(x) n^{-1/2} \sum_{k=1}^n (\eta_{sn} \hat{u}_{k-1,s} - \eta_{sn} \tilde{u}_{k-1,s}). \end{aligned} \quad (6.9.16)$$

Из определений $\hat{u}_{k-1,s}$ и $\tilde{u}_{k-1,s}$ в (6.9.6) и (6.9.7) следует, что последнее выражение не больше

$$2\Theta_n \sup_x g(x) 3^{-m_n} n^{-1/2} \sum_{k=1}^n |u_{k-1}| = O_p(n^{1/2-r_1-r}) = o_p(1).$$

Осталось рассмотреть (6.9.14). Очевидно, (6.9.14) не больше

$$\begin{aligned} \sup_{s \leq 3^{m_n}} n^{-1/2} \sum_{k=1}^n [\Delta_k(\eta_{sn} \hat{u}_{k-1,s}) - \Delta_k(\eta_{sn} \tilde{u}_{k-1,s})] &\leq \\ \leq \sup_{s \leq 3^{m_n}} \left| n^{-1/2} \sum_{k=1}^n [\Delta_k(\eta_{sn} \hat{u}_{k-1,s}) - G(\eta_{sn} \hat{u}_{k-1,s}) - \right. & \\ \left. - \Delta_k(\eta_{sn} \tilde{u}_{k-1,s}) + G(\eta_{sn} \tilde{u}_{k-1,s})] \right| + \end{aligned} \quad (6.9.17)$$

$$+ \sup_{s \leq 3^{m_n}} n^{-1/2} \sum_{k=1}^n [G(\eta_{sn} \hat{u}_{k-1,s}) - G(\eta_{sn} \tilde{u}_{k-1,s})]. \quad (6.9.18)$$

Выражение (6.9.18) совпадает с (6.9.16) и есть $o_p(1)$.

Рассмотрим (6.9.17). Обозначим для краткости

$$\begin{aligned} \nu_k &= \nu_k(n, s) = \\ &= \Delta_k(\eta_{sn} \hat{u}_{k-1,s}) - G(\eta_{sn} \hat{u}_{k-1,s}) - \Delta_k(\eta_{sn} \tilde{u}_{k-1,s}) + G(\eta_{sn} \tilde{u}_{k-1,s}). \end{aligned}$$

Тогда выражение (6.9.17) есть

$$\sup_{s \leq 3^{m_n}} \left| n^{-1/2} \sum_{k=1}^n \nu_k \right|.$$

Аналогично рассмотренным ранее в (6.9.12) величинам ξ_k величины ν_k центрированы, не коррелированы и $\sup_{s \leq 3^{m_n}} \mathbf{E} \nu_k^2 = O(n^{-r} 3^{-m_n})$ равномерно по k . Отсюда $\forall \varepsilon > 0$ в силу неравенства Чебышева

$$\begin{aligned} \mathbf{P} \left\{ \sup_{t \leq N_n} \sup_{s \leq 3^{m_n}} \left| n^{-1/2} \sum_{k=1}^n \nu_k \right| \geq \varepsilon \right\} &\leq \\ &\leq \varepsilon^{-2} \sum_{t=1}^{N_n} \sum_{s=0}^{3^{m_n}} \mathbf{E} \left(n^{-1/2} \sum_{k=1}^n \nu_k \right)^2 = O(N_n n^{-r}) = o(1). \end{aligned}$$

Итак, (6.9.17) есть $o_p(1)$. Лемма 6.9.2 доказана. \square

Лемма 6.9.3. Пусть выполнены условия 6.1.2, 6.1.3, верна альтернатива $H_{1n}(a)$, $N_n \sim \log_b n^{-1}$, где $|\beta_0| < b < 1$; $\delta = \min(\Delta, 1)$, $r > 0$, $\Theta_n = \Theta n^{-r}$, $0 < \Theta < \infty$.

Тогда

$$\begin{aligned} \mathbf{E} \sup_{t \leq N_n} \sup_{|\eta_n| \leq \Theta_n} \left| n^{-1} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)] u_{k-1} \right| &= \\ &= O(n^{-r\delta/(1+\delta)} \log^2 n). \end{aligned}$$

Доказательство. Разделим отрезок $[-\Theta_n, \Theta_n]$ на 3^{m_n} равных частей точками $\eta_{sn} = -\Theta_n + 2\Theta_n 3^{-m_n} s$, $s = 0, 1, \dots, 3^{m_n}$, и пусть $3^{m_n} \sim N_n$. Пусть случайные величины \hat{u}_k и \tilde{u}_k определены соотношениями (6.9.6) и (6.9.7) соответственно. Для любого $|\eta_n| \leq \Theta_n$ выберем среди $\{\eta_{sn}\}$ ту точку η_n , для которой

$$0 \leq \eta_n - \eta_n \leq 2\Theta_n 3^{-m_n}.$$

Тогда справедливо неравенство (6.9.9). В силу этого неравенства и

монотонности $\Delta_k(y)$ имеем

$$\begin{aligned} \Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0) &\leq \Delta_{k-t}(\eta_n \hat{u}_{k-t-1,l}) - \Delta_{k-t}(0) \leq \\ &\leq |\Delta_{k-t}(\eta_l \hat{u}_{k-t-1,l}) - \Delta_{k-t}(0)| + |\Delta_{k-t}(\eta_l \tilde{u}_{k-t-1,l}) - \Delta_{k-t}(0)|. \end{aligned}$$

Аналогично,

$$\begin{aligned} \Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0) &\geq \Delta_{k-t}(\eta_l \tilde{u}_{k-t-1,l}) - \Delta_{k-t}(0) \geq \\ &\geq -|\Delta_{k-t}(\eta_l \hat{u}_{k-t-1,l}) - \Delta_{k-t}(0)| - |\Delta_{k-t}(\eta_l \tilde{u}_{k-t-1,l}) - \Delta_{k-t}(0)|. \end{aligned}$$

Следовательно,

$$\begin{aligned} \sup_{|\eta_n| \leq \Theta_n} |\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)| &\leq \\ &\leq \sup_{s \leq 3^{m_n}} |\Delta_{k-t}(\eta_s \hat{u}_{k-t-1,s}) - \Delta_{k-t}(0)| + \end{aligned} \quad (6.9.19)$$

$$+ \sup_{s \leq 3^{m_n}} |\Delta_{k-t}(\eta_s \tilde{u}_{k-t-1,s}) - \Delta_{k-t}(0)|. \quad (6.9.20)$$

Рассмотрим (6.9.19). Очевидно, для среднего от выражения (6.9.19) справедливо неравенство

$$\begin{aligned} \mathbf{E} \sup_{s \leq 3^{m_n}} |\Delta_{k-t}(\eta_s \hat{u}_{k-t-1,s}) - \Delta_{k-t}(0)| &\leq \\ &\leq \sum_{s=0}^{3^{m_n}} \mathbf{E} |\Delta_{k-t}(\eta_s \hat{u}_{k-t-1,s}) - \Delta_{k-t}(0)| \leq \\ &\leq \sum_{s=0}^{3^{m_n}} \left[\mathbf{P}\{0 < \varepsilon_{k-t} < |\eta_s \hat{u}_{k-t-1,s}|\} + \right. \\ &\quad \left. + \mathbf{P}\{-|\eta_s \hat{u}_{k-t-1,s}| < \varepsilon_{k-t} < 0\} \right]. \end{aligned} \quad (6.9.21)$$

Первая из вероятностей под знаком суммы в (6.9.21) не больше

$$\mathbf{E}[G(|\eta_s \hat{u}_{k-t-1,s}|) - G(0)] \leq \sup_x g(x) \mathbf{E} |\eta_s \hat{u}_{k-t-1,s}| = O(n^{-r})$$

равномерно по s, k, t . То же самое верно и для второй вероятности в (6.9.21), так что (6.9.19) есть $O(n^{-r} \log n)$. Аналогично оценивая среднее (6.9.20), получаем

$$\mathbf{E} \sup_{|\eta_n| \leq \Theta_n} |\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)| = O(n^{-r} \log n) \quad (6.9.22)$$

равномерно по k и t . Отсюда, используя неравенство Гёльдера, полу-

чаем

$$\begin{aligned} & \mathbf{E} \sup_{t \leq N_n} \sup_{|\eta_n| \leq \Theta_n} \left| n^{-1} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)] u_{k-1} \right| \leq \\ & \leq \sum_{t=1}^{N_n} \left\{ n^{-1} \sum_{k=t+1}^n \mathbf{E} \sup_{|\eta_n| \leq \Theta_n} |\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)| |u_{k-1}| \right\} \leq \\ & \leq N_n n^{-1} \sum_{k=t+1}^n \left\{ \mathbf{E} \left\{ \sup_{|\eta_n| \leq \Theta_n} |\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)| \right\}^{(1+\delta)/\delta} \right\}^{\delta/(1+\delta)} \times \\ & \quad \times \left\{ \mathbf{E} |u_{k-1}|^{1+\delta} \right\}^{1/(1+\delta)} = \\ & = O(N_n (n^{-r} \log n)^{\delta/(1+\delta)}) = O(n^{-r\delta/(1+\delta)} \log^2 n) \end{aligned}$$

в силу (6.9.22). Лемма 6.9.3 полностью доказана. \square

Л е м м а 6.9.4. Пусть выполнены условия 6.1.2, 6.1.3, верна альтернатива $H_{1n}(a)$, $\Theta_n = \Theta n^{-r}$, $r > 1/4$, $0 < \Theta < \infty$.

Тогда:

- 1) $\sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{n-1} (\beta_n + \eta_n)^{t-1} x_{tn}(\eta_n) \right| = o_p(1)$;
- 2) $\sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{n-1} (\beta_n + \eta_n)^{t-1} u_{tn}(\eta_n) \right| = o_p(1)$.

Доказательство. Пункты 1) и 2) леммы 6.9.4 доказываются с помощью п. 1) и 2) леммы 6.9.2 соответственно. Докажем п. 1).

Очевидно, что $|x_{tn}(\eta_n)| \leq 4n^{1/2}$ при любом $t = 1, \dots, n-1$. Отсюда $\forall \varepsilon > 0$ для b такого, что $|\beta_0| < b < 1$, при $n > n_0$ и $N_n \sim \log_3 n^{-1}$ имеем

$$\begin{aligned} \mathbf{P} \left\{ \sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=N_n+1}^{n-1} (\beta_n + \eta_n)^{t-1} x_{tn}(\eta_n) \right| \geq \varepsilon \right\} & \leq \\ & \leq \mathbf{P} \left\{ 4n^{1/2} \sum_{t \geq N_n+1} b^{t-1} > \varepsilon \right\} = 0, \end{aligned}$$

так как $n^{1/2} \sum_{t \geq N_n+1} b^{t-1} = O(n^{1/2} b^{N_n}) = o(1)$. Далее, при $n > n_0$

$$\begin{aligned} \sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{N_n} (\eta_n + \beta_n)^{t-1} x_{tn}(\eta_n) \right| & \leq \\ & \leq \sup_{t \leq N_n} \sup_{|\eta_n| \leq \Theta_n} |x_{tn}(\eta_n)| \sum_{t \geq 1} b^{t-1} = o_p(1) \end{aligned}$$

в силу п. 1) леммы 6.9.2. Пункт 1) леммы 6.9.4 доказан. \square

Л е м м а 6.9.5. Пусть выполнено условие 6.1.2 и верна альтернатива $H_{1n}(a)$.
Тогда

$$n^{-1} \sum_{k=t+1}^n \Delta_{k-t}(0) u_{k-1} = -\frac{1}{2} \mathbf{E} |\varepsilon_1| \beta_0^{t-1} + \delta_{tn},$$

где при $N_n \sim \log_b n^{-1}$ и $|\beta_0| < b < 1$ имеем

$$\sup_{t \leq N_n} \mathbf{E} |\delta_{tn}| = O(n^{-\delta/(1+\delta)} \log n), \quad \delta = \min(\Delta, 1).$$

Д о к а з а т е л ь с т в о. Рассмотрим более трудный случай $0 < \Delta < 1$.

Заметим прежде всего, что если $\xi_1, \xi_2, \dots, \xi_n$ — н.о.р.сл.в., $\mathbf{E} \xi_1 = 0$, $\mathbf{E} |\xi_1|^{1+\Delta} < \infty$, то

$$n^{-1/(1+\Delta)} \sum_{i=1}^n \xi_i \xrightarrow{п.н.} 0 \quad (6.9.23)$$

[44, с. 383]. Далее, поскольку при $0 < \Delta < 1$

$$\mathbf{E} |\xi_1 + \dots + \xi_n|^{1+\Delta} \leq 2^{1+\Delta} n \mathbf{E} |\xi_1|^{1+\Delta}$$

(см. [46]), то

$$\mathbf{E} \left| n^{-1/(1+\Delta)} \sum_{i=1}^n \xi_i \right|^{1+\Delta} \leq 2^{1+\Delta} \mathbf{E} |\xi_1|^{1+\Delta},$$

и последовательность $n^{-1/(1+\Delta)} \sum_{i=1}^n \xi_i$ равномерно интегрируема.

Следовательно,

$$\mathbf{E} \left| n^{-1/(1+\Delta)} \sum_{i=1}^n \xi_i \right| \rightarrow 0.$$

Последнее соотношение влечет

$$\mathbf{E} \left| n^{-1} \sum_{i=1}^n \xi_i \right| = o(n^{-\Delta/(1+\Delta)}). \quad (6.9.24)$$

Теперь

$$u_{k-1} = \sum_{s=0}^{t-2} \beta_n^s \varepsilon_{k-1-s} + \beta_n^{t-1} \varepsilon_{k-t} + \beta_n^t u_{k-t-1},$$

и отсюда

$$n^{-1} \sum_{k=t+1}^n \Delta_{k-t}(0) u_{k-1} =$$

$$= n^{-1} \sum_{k=t+1}^n \sum_{s=0}^{t-2} \beta_n^s \Delta_{k-t}(0) \varepsilon_{k-1-s} + \quad (6.9.25)$$

$$+ \beta_n^t n^{-1} \sum_{k=t+1}^n \Delta_{k-t}(0) \varepsilon_{k-t} + \quad (6.9.26)$$

$$+ \beta_n^t n^{-1} \sum_{k=t+1}^n \Delta_{k-t}(0) u_{k-t-1}. \quad (6.9.27)$$

Рассмотрим (6.9.25). Среднее модуля этого выражения не больше

$$\sum_{s=0}^{t-2} \beta_n^s \mathbb{E} \left| n^{-1} \sum_{k=1}^{n-t} \Delta_k(0) \varepsilon_{k+t-s-1} \right|. \quad (6.9.28)$$

При $N_n \sim \log_b n^{-1}$

$$\sup_{t \leq N_n} \sup_{1 \leq m \leq N_n} \mathbb{E} \left| n^{-1} \sum_{k=1}^{n-t} \Delta_k(0) \varepsilon_{k+m} \right| =$$

$$= \sup_{1 \leq m \leq N_n} \mathbb{E} \left| n^{-1} \sum_{k=1}^n \Delta_k(0) \varepsilon_{k+m} \right| + \delta_{1n} =$$

$$= \sup_{1 \leq m \leq N_n} \mathbb{E} \left| n^{-1} \sum_{j=1}^m \sum_{i=0}^{n/m} \Delta_{j+(m+1)i} \varepsilon_{j+(m+1)i+m} \right| + \delta_{1n} + \delta_{2n}, \quad (6.9.29)$$

$$\mathbb{E} |\delta_{jn}| \leq N_n^{-1} n^{-1} = O(n^{-1} \log n), \quad j = 1, 2.$$

Теперь среднее в (6.9.29) не больше

$$m^{-1} \sum_{j=1}^m \mathbb{E} \left| (mn^{-1}) \sum_{i=0}^{n/m} \Delta_{j+(m+1)i} \varepsilon_{j+(m+1)i+m} \right|. \quad (6.9.30)$$

Слагаемые под знаком внутренней суммы в (6.9.30) представляют собой н.о.р.сл.в. со средним нуль и конечным моментом порядка $1 + \Delta$. В силу (6.9.24) выражение в (6.9.30) есть $o((n/m)^{-\Delta/(1+\Delta)})$. Из последнего соотношения легко следует, что (6.9.29) есть $O(n^{-\Delta/(1+\Delta)} \log n)$, а значит, и (6.9.28) есть $O(n^{-\Delta/(1+\Delta)} \log n)$. Итак,

(6.9.25) есть δ_{3tn} , где

$$\sup_{t \leq N_n} \mathbf{E}|\delta_{3tn}| = O(n^{-\Delta/(1+\Delta)} \log n).$$

Далее, (6.9.26) есть $-1/2 \beta_0^{t-1} \mathbf{E}|\varepsilon_1| + \delta_{4tn}$, где

$$\sup_{t \leq N_n} \mathbf{E}|\delta_{4tn}| = O(n^{-1} \log n + n^{-\Delta/(1+\Delta)}) = O(n^{-\Delta/(1+\Delta)}).$$

Наконец, пользуясь представлением

$$u_{k-t-1} = \sum_{s=0}^{\infty} \beta_n^s \varepsilon_{k-t-1-s},$$

и рассуждая, как при анализе (6.9.25), получаем, что (6.9.27) есть δ_{5tn} и

$$\sup_{t \leq N_n} \mathbf{E}|\delta_{5tn}| = O(n^{-\Delta/(1+\Delta)} \log n).$$

Лемма 6.9.5 доказана. \square

Л е м м а 6.9.6. Пусть выполнены условия 6.1.2, 6.1.3, верна альтернатива $H_{1n}(a)$, $\delta = \min(\Delta, 1)$, $\Theta_n = \Theta n^{-r}$, $r > \frac{1+\delta}{2(1+2\delta)}$, $0 < \Theta < \infty$.

Тогда

$$\sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{n-1} (\beta_n + \eta_n)^{t-1} y_{tn}(\eta_n) + \frac{1}{2} g(0) \mathbf{E}|\varepsilon_1| (1 - \beta_0^2)^{-1} \theta \right| = o_p(1).$$

Доказательство. Пусть $N_n \sim \log_b n^{-1}$, где $|\beta_0| < b < 1$. Как при доказательстве леммы 6.9.4, проверяется, что

$$\sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=N_n+1}^{n-1} (\beta_n + \eta_n)^{t-1} y_{tn}(\eta_n) \right| = o_p(1).$$

В силу формулы Тейлора

$$\sum_{t=1}^{N_n} (\beta_n + \eta_n)^{t-1} y_{tn}(\eta_n) =$$

$$= G(0) \sum_{t=1}^{N_n} (\beta_n + \eta_n)^{t-1} \times \\ \times \left[n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)] \right] + \quad (6.9.31)$$

$$+ \sum_{t=1}^{N_n} (\beta_n + \eta_n)^{t-1} \times \\ \times \left[n^{-1/2} \sum_{k=t+1}^n \Delta_{k-t}(\eta_n u_{k-t-1}) g(\theta_k \eta_n u_{k-1}) \eta_n u_{k-1} \right], \quad (6.9.32)$$

где $\theta_k \in (0, 1)$. Супремум модуля (6.9.31) по $|\eta_n| \leq \Theta_n$ есть $o_p(1)$ в силу п. 2) леммы 6.9.2, так как $r > \frac{1+\delta}{2(1+2\delta)} > \frac{1}{4}$. Выражение (6.9.32) в силу условия

$$|g(x) - g(0)| \leq h|x|^\delta, \quad \delta = \min(\Delta, 1)$$

(см. доказательство теоремы 6.2.1), перепишем в виде

$$g(0)\eta_n \sum_{t=1}^{N_n} (\beta_n + \eta_n)^{t-1} \left[n^{-1/2} \sum_{k=t+1}^n \Delta_{k-t}(\eta_n u_{k-t+1}) u_{k-1} \right] + \varepsilon_n, \quad (6.9.33)$$

причем для $|\eta_n| \leq \Theta_n$ и $|\beta| < b < 1$ при $n > n_0$

$$|\varepsilon_n| \leq h n^{1/2} |\eta_n|^{1+\delta} \sum_{t=1}^{N_n} b^{t-1} \left[n^{-1} \sum_{k=1}^n |u_{k-1}|^{1+\delta} \right] = o_p(1)$$

при $r > \frac{1+\delta}{2(1+2\delta)}$. Главный член в (6.9.33) представим в виде

$$g(0)\eta_n n^{1/2} \sum_{t=1}^{N_n} (\beta_n + \eta_n)^{t-1} \times \\ \times \left[n^{-1} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n u_{k-t-1}) - \Delta_{k-t}(0)] u_{k-1} \right] + \quad (6.9.34)$$

$$+ g(0)\eta_n n^{1/2} \sum_{t=1}^{N_n} (\beta_n + \eta_n)^{t-1} \left[n^{-1} \sum_{k=t+1}^n \Delta_{k-t}(0) u_{k-1} \right]. \quad (6.9.35)$$

Среднее супремума по $|\eta_n| \leq \Theta_n$ модуля (6.9.34) есть $o(1)$ в силу

леммы 6.9.3 для $r > \frac{1 + \delta}{2(1 + 2\delta)}$. Наконец, (6.9.35) есть

$$-\frac{1}{2}g(0)E|\varepsilon_1|(1 - \beta_0^2)^{-1}\theta + \varepsilon_{1n},$$

где $\sup_{|\eta_n| \leq \Theta_n} |\varepsilon_{1n}| = o_p(1)$ в силу леммы 6.9.5. Лемма 6.9.6 полностью доказана. \square

Теперь из (6.9.5), лемм 6.9.4 и 6.9.6 и определения $\eta_n = n^{-1/2}\theta$ непосредственно следует, что (6.9.1) равно

$$\lambda_S(\beta_0)\theta + \varepsilon_{2n}(\theta), \quad \sup_{|\theta| \leq \Theta_n} |\varepsilon_{2n}(\theta)| = o_p(1) \quad (6.9.36)$$

для $r < \frac{\delta}{2(1 + 2\delta)}$. Из леммы 6.9.1 и (6.9.36) следует утверждение теоремы 6.4.1. \square

Глава 7

ЗНАКОВЫЙ АНАЛИЗ МНОГОПАРАМЕТРИЧЕСКОЙ АВТОРЕГРЕССИИ

§ 7.1. Введение

В этой главе будем рассматривать схему авторегрессии вида

$$u_i = \beta_1 u_{i-1} + \dots + \beta_q u_{i-q} + \varepsilon_i, \quad i \in \mathbb{Z}. \quad (7.1.1)$$

В (7.1.1) порядок q известен; β_1, \dots, β_q — такие неизвестные неслучайные коэффициенты, что характеристическое уравнение

$$x^q = \beta_1 x^{q-1} + \dots + \beta_q \quad (7.1.2)$$

имеет корни по модулю меньше единицы; $\{\varepsilon_i\}$ — н.о.р.сл.в. с неизвестной невырожденной функцией распределения $G(x)$. Эти условия дальше особо оговариваться не будут.

Обозначим $\beta = (\beta_1, \dots, \beta_q)^T$ и определим последовательность $\{\delta_t = \delta_t(\beta), t = 1 - q, 2 - q, \dots\}$ рекуррентным соотношением

$$\delta_t = \beta_1 \delta_{t-1} + \dots + \beta_q \delta_{t-q}, \quad t = 1, 2, \dots, \quad (7.1.3)$$

с начальными условиями $\delta_{1-q} = \dots = \delta_{-1} = 0$, $\delta_0 = 1$. Если корни уравнения (7.1.2) лежат внутри единичного круга, то величины δ_t убывают к нулю с экспоненциальной скоростью,

$$|\delta_t(\beta)| \leq c b^t, \quad t = 1, 2, \dots, \quad (7.1.4)$$

для некоторых $c > 0$ и $0 < b < 1$, не зависящих от n (см., например, [2, 45, § 5.2.1]). Отсюда, рассуждая как при доказательстве леммы 5.2.2, легко получить: при $E \log^+ |\varepsilon_1| < \infty$ существует п.н. единственное строго стационарное решение уравнения (7.1.1), представимое в виде

$$u_i = \sum_{t=0}^{\infty} \delta_t(\beta) \varepsilon_{i-t}, \quad (7.1.5)$$

ряд в (7.1.5) сходится п.н. абсолютно. Если $E|\varepsilon_1|^{1+\Delta} < \infty$, то ряд

в (7.1.5) сходится и в $\mathbb{L}^{1+\Delta}$, и это устанавливается рассуждениями, аналогичными доказательству леммы 5.2.1.

Пусть u_{1-q}, \dots, u_n — наблюдения, удовлетворяющие (7.1.5). По этим наблюдениям мы построим непараметрические знаковые тесты для проверки гипотез относительно вектора β , а также соответствующие непараметрические знаковые оценки, обобщающие тесты и оценки из гл. 6. Наши процедуры основываются на векторе знаков

$$S^n(\theta) = (S_1(\theta), \dots, S_n(\theta)), \quad \theta = (\theta_1, \dots, \theta_q)^T \in \mathbb{R}^q, \quad (7.1.6)$$

$$S_k(\theta) = \text{sign}(u_k - \theta_1 u_{k-1} - \dots - \theta_q u_{k-q}), \quad k = 1, \dots, n. \quad (7.1.7)$$

Мы начинаем в § 7.2 с проверки гипотезы

$$H_0: \beta = \beta_0, \quad (7.1.8)$$

где вектор $\beta_0 = (\beta_{10}, \dots, \beta_{q0})^T$ полностью известен.

Пусть $\beta_0^j = (\beta_{10}, \dots, \beta_{j-1,0}, \beta_j, \beta_{j+1,0}, \dots, \beta_{q0})^T$ — вектор с единственной свободной j -й компонентой β_j . Альтернативами к (7.1.8) берутся гипотезы

$$H_j^+: \beta = \beta_0^j, \quad \beta_j > \beta_{j0}, \quad j = 1, \dots, q, \quad (7.1.9)$$

и аналогичные левосторонние альтернативы H_j^- .

Тестовые статистики $l_{j,n}^S(\beta_0)$ основаны на векторе $S^n(\beta_0)$, соответствующие тесты являются ЛНМ среди всех, основанных на векторе знаков $S^n(\beta_0)$. Затем мы строим статистику $L_{n,S}(\beta_0)$ для проверки H_0 против альтернативы

$$H_1: \beta \neq \beta_0. \quad (7.1.10)$$

Статистика $L_{n,S}(\beta_0)$ является квадратичной формой от $l_{j,n}^S(\beta_0)$, $j = 1, \dots, q$. Матрица квадратичной формы пропорциональна матрице $I_n^{-1}(\beta_0, \beta_0)$, где $I_n(\beta, \theta)$ — информация Фишера о векторе β , содержащаяся в векторе знаков $S^n(\theta)$. Оказывается,

$$I_n(\beta, \beta) \sim n(2g(0) \mathbb{E}|\varepsilon_1|)^2 K(\beta),$$

где

$$K(\beta) = (k_{ij}(\beta))_{i,j=1,\dots,q}, \quad k_{ij}(\beta) = \sum_{t=0}^{\infty} \delta_t(\beta) \delta_{t+|i-j|}(\beta). \quad (7.1.11)$$

Матрица $K(\beta)$ играет центральную роль в нашем рассмотрении. Объяснение просто: при стандартных предположениях относительно $G(x)$, в частности, если $\mathbb{E}\varepsilon_1^2 < \infty$, $G(x)$ имеет плотность $g(x)$ и информация Фишера $g(x)$ относительно сдвига конечна (обозначим ее, как

в п. 5.3.2, $I(g)$), то информационная матрица наблюдений u_{1-g}, \dots, u_n относительно вектора β есть

$$J_n(\beta) \sim n I(g) E \varepsilon_1^2 K(\beta)$$

(см., например, [18, 59, п. 9.2.3]. Уместно здесь же отметить, что $\det(K(\beta)) \neq 0$ [2, 45, лемма 5.5.5]. Кроме того, при $E \varepsilon_1 = 0$, $E \varepsilon_1^2 < \infty$ для оценки наименьших квадратов $\hat{\beta}_{n,LS}$ справедливо

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta) \xrightarrow{d} N(0, \Sigma_{LS}(\beta)),$$

где $\Sigma_{LS}(\beta) = K^{-1}(\beta)$ [2, 45, теорема 5.5.7].

В § 7.3 мы исследуем поведение тестовых статистик из § 7.2 при близких альтернативах

$$H_{1n}(a): \beta = \beta_n := \beta_0 + a n^{-1/2} + o(n^{-1/2}), \quad (7.1.12)$$

a — постоянный вектор. Асимптотическая относительная эффективность (АОЭ) нашего знакового теста относительно теста, основанного на статистике $\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0)$, оказывается равной

$$e_{S,LS} = (2g(0) E|\varepsilon_1|)^2,$$

т.е. такой же, как при $g = 1$.

Соответствующий результат получается с помощью линейного равномерного при $|\theta| \leq \Theta_n$ стохастического разложения векторного поля $I_n^S(\beta_n + n^{-1/2} \theta)$, где $I_n^S(\theta) = (I_{1n}^S(\theta), \dots, I_{qn}^S(\theta))^T$, $\Theta_n \rightarrow \infty$ со степенной скоростью. Здесь и далее $|\cdot|$ означает евклидову норму вектора или матрицы.

В § 7.4 мы строим тесты для проверки линейной гипотезы относительно β . А именно, пусть вектор β разбит на подвекторы $\beta^T = (\beta^{(1)T}, \beta^{(2)T})$, $\beta^{(i)}$, $i = 1, 2$, размерностей соответственно m и $q - m$, $1 \leq m < q$. Мы проверяем гипотезу

$$H'_0: \beta^{(2)} = \beta_0^{(2)}, \quad (7.1.13)$$

$\beta_0^{(2)}$ — известный вектор, вектор $\beta^{(1)}$ является мешающим параметром. При $\beta_0^{(2)} = 0$ гипотеза (7.1.13) означает, что истинный порядок уравнения (7.1.1) равен m и $m < q$.

Мы исследуем поведение тестовой статистики при близких альтернативах. АОЭ нашего теста относительно теста, основанного на $\hat{\beta}_{n,LS}$, оказывается равной $e_{S,LS}$.

Наконец, в § 7.5 мы строим непараметрические знаковые оценки, а в § 7.6 находим их функционалы влияния для схемы засорения независимыми выбросами. Как и в одномерной авторегрессионной модели,

мы строим несколько знаковых оценок, все они распределены асимптотически одинаково. Например, для знаковой оценки $\tilde{\beta}_{n,S}$ в случае схемы (7.1.1) (т.е. без засорения) мы показываем, что

$$\sqrt{n}(\tilde{\beta}_{n,S} - \beta) \xrightarrow{d} N(0, \Sigma_S(\beta)),$$

где $\Sigma_S(\beta) = (2g(0) \mathbb{E}|\varepsilon_1|)^{-2} K^{-1}(\beta)$.

Матрица ковариаций $\Sigma_S(\beta)$ лишь множителем $e_{S,LS}^{-1}$ отличается от матрицы ковариаций $\Sigma_{LS}(\beta)$ оценки наименьших квадратов.

Мы завершаем главу § 7.7, в котором исследуем свойства эмпирической функции распределения $\hat{G}_n(x)$, построенной по оценкам $\varepsilon_1(\hat{\beta}_n), \dots, \varepsilon_n(\hat{\beta}_n)$ ненаблюдаемых величин $\varepsilon_1, \dots, \varepsilon_n$. Эти оценки получаются с помощью \sqrt{n} -состоятельной оценки $\hat{\beta}_n = (\hat{\beta}_{1n}, \dots, \hat{\beta}_{qn})^T$ вектора β

$$\varepsilon_k(\hat{\beta}_n) := u_k - \hat{\beta}_{1n}u_{k-1} - \dots - \hat{\beta}_{qn}u_{k-q}, \quad k = 1, \dots, n,$$

$$\hat{G}_n(x) := n^{-1} \sum_{k=1}^n I(\varepsilon_k(\hat{\beta}_n) < x),$$

$I(\cdot)$ — индикатор события. Мы показываем: если $G_n(x)$ есть истинная (но недоступная для построения) эмпирическая функция распределения величин $\varepsilon_1, \dots, \varepsilon_n$, то

$$\sup_x \sqrt{n} |\hat{G}_n(x) - G_n(x)| = o_p(1). \quad (7.1.14)$$

Условия, при которых верно (7.1.14), незначительно отличаются от условий 6.1.2, 6.1.3.

Пользуясь (7.1.14), мы переносим хорошо известные свойства $G_n(x)$ на $\hat{G}_n(x)$. В частности, процессы типа эмпирического и параметрического эмпирического, основанные на $\hat{G}_n(x)$, имеют те же слабые пределы в пространстве Скорохода $\mathbb{D}[0, 1]$, что и аналогичные процессы, основанные на $G_n(x)$. Это позволяет основывать на $\hat{G}_n(x)$ тесты типа Колмогорова и омега-квадрат для проверки простых и сложных гипотез относительно $G(x)$.

§ 7.2. Тестовые статистики и их распределения при гипотезе

Рассмотрим схему (7.1.1), и пусть наблюдения u_{-q+1}, \dots, u_n удовлетворяют (7.1.5), а вектор $S^n(\theta)$ определен соотношениями (7.1.6)–(7.1.7).

Поставим задачей этого параграфа построить по вектору $S^n(\beta_0)$ ЛНМ тесты для проверки гипотезы (7.1.8) $H_0: \beta = \beta_0$ против правосторонних альтернатив H_j^+ из (7.1.9), $j = 1, 2, \dots, q$, и аналогичных

левосторонних альтернатив H_j^- . Построим также тест для проверки H_0 против альтернативы (7.1.10) $H_1: \beta \neq \beta_0$.

Будем обозначать $s^n = (s_1, \dots, s_n)$, $s_k = \pm 1$ реализации вектора $S^n(\beta_0)$. Обозначим также

$$\gamma_{tn}(s^n) = \sum_{k=t+1}^n s_{k-t}s_k, \quad t = 1, \dots, n-1. \quad (7.2.1)$$

Пусть $\{\delta_t(\beta), t = 1-q, 2-q, \dots\}$ — последовательность, определенная (7.1.3). Основой предлагаемых процедур служит следующая теорема.

Теорема 7.2.1. Пусть выполнены условия 6.1.1–6.1.3, $\beta = \beta_0^j$, $n \geq j+1$, $j = 1, \dots, q$.

Тогда при $\beta_j \rightarrow \beta_{j0}$

$$\begin{aligned} P_{\beta_0^j} \{S^n(\beta_0) = s^n\} &= \left(\frac{1}{2}\right)^n \left\{ 1 + 2g(0)E|\varepsilon_1| \sum_{t=0}^{n-j-1} \delta_t(\beta_0) \times \right. \\ &\quad \left. \times \gamma_{t+j,n}(s^n)(\beta_j - \beta_{j0}) \right\} + o(\beta_j - \beta_{j0}). \end{aligned}$$

При $q = 1$ утверждение теоремы 7.2.1 совпадает с утверждением теоремы 6.2.1.

Доказательство. При $\beta = \beta_0^j$, $j = 1, \dots, q$, имеем

$$\begin{aligned} u_k - \beta_{10}u_{k-1} - \dots - \beta_{j0}u_{k-j} - \dots - \beta_{q0}u_{k-q} &= \\ &= \varepsilon_k + (\beta_j - \beta_{j0})u_{k-j}, \quad k = 1, \dots, n. \quad (7.2.2) \end{aligned}$$

Установим рекуррентное соотношение между правдоподобиями $S^n(\beta_0)$ и $S^{n-1}(\beta_0)$. Имеем для $n \geq j+1$

$$\begin{aligned} P_{\beta_0^j} \{S^n(\beta_0) = s^n\} &= E_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \times \right. \\ &\quad \left. \times \left[I(S_n(\beta_0) = 1) \frac{1+s_n}{2} + I(S_n(\beta_0) = -1) \frac{1-s_n}{2} \right] \right\} = \\ &= E_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \times \right. \\ &\quad \left. \times \left[\frac{1+s_n}{2} - s_n I(u_n - \beta_{10}u_{n-1} - \dots - \beta_{q0}u_{n-q} < 0) \right] \right\} = \\ &= E_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \times \right. \\ &\quad \left. \times \left[\frac{1+s_n}{2} - s_n I(\varepsilon_n < (\beta_{j0} - \beta_j)u_{n-j}) \right] \right\} \quad (7.2.3) \end{aligned}$$

в силу (7.2.2).

Пусть $\Omega_{\leq i}$ — сигма-алгебра, порожденная $\{\varepsilon_r, r \leq i\}$. В силу представления (7.1.5) ε_n и $\Omega_{\leq n-1}$ независимы. При условии 6.1.3 $g(x)$ удовлетворяет условию Гельдера произвольного порядка $\delta \in (0, 1]$. Положим $\delta = \min(1, \Delta)$. Тогда (7.2.3) в силу формулы полного математического ожидания равно

$$\begin{aligned} & \mathbf{E}_{\beta_0^j} \left\{ \mathbf{E}_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \times \right. \right. \\ & \quad \times \left. \left[\frac{1+s_n}{2} - s_n I(\varepsilon_n < (\beta_{j0} - \beta_j) u_{n-j}) \right] \middle| \Omega_{n-1} \right\} \right\} = \\ & = \mathbf{E}_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \left[\frac{1+s_n}{2} - s_n G((\beta_{j0} - \beta_j) u_{n-j}) \right] \right\} = \\ & = \mathbf{E}_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \left[\frac{1+s_n}{2} - s_n \left(\frac{1}{2} + g(0)(\beta_{j0} - \beta_j) u_{n-j} \right) \right] \right\} + \\ & \quad + O(|\beta_j - \beta_{j0}|^{1+\delta}) = \\ & = \mathbf{E}_{\beta_0^j} \left\{ I(S^{n-1}(\beta_0) = s^{n-1}) \left[\frac{1}{2} + s_n g(0) u_{n-j} (\beta_j - \beta_{j0}) \right] \right\} + \\ & \quad + o(\beta_j - \beta_{j0}) = \frac{1}{2} \mathbf{P}_{\beta_0^j} \{ S^{n-1}(\beta_0) = s^{n-1} \} + \\ & + s_n g(0) \mathbf{E}_{\beta_0^j} \{ I(S^{n-1}(\beta_0) = s^{n-1}) u_{n-j} \} (\beta_j - \beta_{j0}) + o(\beta_j - \beta_{j0}). \quad (7.2.4) \end{aligned}$$

В силу (7.1.5)

$$u_{n-j} = \sum_{t=0}^{\infty} \delta_t(\beta_0^j) \varepsilon_{n-j-t}.$$

Подставим последнее выражение в (7.2.4), и для вычисления среднего

$$\mathbf{E}_{\beta_0^j} \{ I(S^{n-1}(\beta_0) = s^{n-1}) \varepsilon_{n-j-t} \}$$

воспользуемся опять описанным выше методом с использованием формулы полного математического ожидания. Подсчеты позволяют переписать (7.2.4) в виде

$$\left(\frac{1}{2} \right)^{n-1} g(0) \mathbf{E}|\varepsilon_1| \sum_{t=0}^{n-j-1} s_n s_{n-j-t} \delta_t(\beta_0^j) (\beta_j - \beta_{j0}) + o(\beta_j - \beta_{j0}). \quad (7.2.5)$$

В силу определения (7.1.3) каждая из величин $\delta_t(\beta)$ является непрерывной функцией β . Поэтому (7.2.5) можно переписать в виде

$$\left(\frac{1}{2} \right)^{n-1} g(0) \mathbf{E}|\varepsilon_1| \sum_{t=0}^{n-j-1} s_n s_{n-j-t} \delta_t(\beta_0) (\beta_j - \beta_{j0}) + o(\beta_j - \beta_{j0}). \quad (7.2.6)$$

Теперь соотношение (7.2.6) влечет

$$\begin{aligned} P_{\beta_0^j} \{S^n(\beta_0) = s^n\} &= \frac{1}{2} P_{\beta_0^j} \{S^{n-1}(\beta_0) = s^{n-1}\} + \\ &+ \left(\frac{1}{2}\right)^{n-1} g(0) E|\varepsilon_1| \sum_{t=0}^{n-j-1} s_n s_{n-j-t} \delta_t(\beta_0) (\beta_j - \beta_{j0}) + \\ &+ o(\beta_j - \beta_{j0}). \end{aligned} \quad (7.2.7)$$

Аналогично (7.2.7) получаем при $n = j$

$$P_{\beta_0^j} \{S^j(\beta_0) = s^j\} = \left(\frac{1}{2}\right)^j + o(\beta_j - \beta_{j0}). \quad (7.2.8)$$

Рекуррентное соотношение (7.2.7) при начальном условии (7.2.8) влечет утверждение теоремы 7.2.1. \square

Непосредственным следствием теоремы 7.2.1 является

Т е о р е м а 7.2.2. Пусть выполнены условия 6.1.1–6.1.3, $n \geq j+1$, $j = 1, \dots, q$.

Тогда ЛНМ тест для проверки $H_0: \beta = \beta_0$ против H_j^+ из (7.1.9), основанный на векторе знаков $S^n(\beta_0)$, имеет критическую область

$$Q_{jn}^+ = \left\{ s^n: \sum_{t=0}^{n-j-1} \delta_t(\beta_0) \gamma_{t+j,n}(s^n) > \text{const} \right\}. \quad (7.2.9)$$

Критическая область для проверки H_0 против левосторонней альтернативы $H_j^-: \beta = \beta_0^j$, $\beta_j < \beta_{j0}$ отличается от (7.2.9) знаком неравенства.

Для проверки H_0 против двусторонней альтернативы $H_{1j}: \beta = \beta_0^j$, $\beta_j \neq \beta_{j0}$ можно брать двустороннюю версию (7.2.9)

$$Q_{jn} = \left\{ s^n: \left| \sum_{t=0}^{n-j-1} \delta_t(\beta_0) \gamma_{t+j,n}(s^n) \right| > \text{const} \right\}. \quad (7.2.10)$$

Для $S_k(\theta)$, $k = 1, \dots, n$, определенных (7.1.7), положим

$$\Gamma_{tn}(\theta) = \sum_{k=t+1}^n S_{k-t}(\theta) S_k(\theta), \quad t = 1, \dots, n-1, \quad (7.2.11)$$

$$l_{jn}^S(\theta) = \sum_{t=0}^{n-j-1} \delta_t(\theta) \Gamma_{t+j,n}(\theta), \quad j = 1, \dots, q. \quad (7.2.12)$$

Тогда $l_{jn}^S(\beta_0)$ есть статистика тестов (7.2.9) и (7.2.10).

При гипотезе $H_0: \beta = \beta_0$

$$\Gamma_{tn}(\beta_0) = \sum_{k=t+1}^n \text{sign}(\varepsilon_{k-t}\varepsilon_k). \quad (7.2.13)$$

Очевидно, при условии 6.1.1 статистика (7.2.13) свободна от $G(x)$. Кроме того, как отмечалось в § 6.2 (см. соотношение (6.2.14)),

$$E_{\beta_0} \Gamma_{tn}(\beta_0) = 0, \quad E_{\beta_0} \Gamma_{tn}(\beta_0) \Gamma_{rn}(\beta_0) = (n-t)\delta_{tr}, \quad (7.2.14)$$

δ_{tr} — символ Кронекера, $t, r = 1, \dots, n-1$,

$$\frac{1}{2}(n-t + \Gamma_{tn}(\beta_0)) \sim \text{Bi}\left(n-t, \frac{1}{2}\right).$$

Поскольку статистики $\Gamma_{tn}(\beta_0)$ при H_0 и условии 6.1.1 свободны от $G(x)$, то этим же свойством обладают и статистики $l_{jn}^S(\beta_0)$, $j = 1, \dots, q$. Распределение $l_{jn}^S(\beta_0)$ зависит от β_0 , однако его можно оценить методом статистических испытаний аналогично тому, как это делалось в § 6.8. Далее в теореме 7.2.3 мы укажем нормальную аппроксимацию для распределения $l_{jn}^S(\beta_0)$ при H_0 и больших n .

Построим теперь статистику для проверки гипотезы $H_0: \beta = \beta_0$ против альтернативы $H_1: \beta \neq \beta_0$. Пусть $n \geq q+1$. Для $l_{jn}^S(\theta)$ из (7.2.12) положим

$$l_n^S(\theta) = (l_{1n}^S(\theta), \dots, l_{qn}^S(\theta))^T. \quad (7.2.15)$$

В силу теоремы 7.2.1

$$\begin{aligned} \frac{\partial \log P_{\beta_0^j}(S^n(\beta_0) = s^n)}{\partial \beta_j} \Big|_{\beta_j = \beta_{j0}} &= \\ &= 2g(0) E|\varepsilon_1| \sum_{t=0}^{n-j-1} \delta_t(\beta_0) \gamma_{tn}(s^n), \quad j = 1, \dots, q. \end{aligned} \quad (7.2.16)$$

Из (7.2.16) и определений (7.2.12), (7.2.15) следует, что естественной статистикой для проверки H_0 против H_1 является квадратичная форма

$$L_{n,S}(\beta_0) = (2g(0) E|\varepsilon_1|)^2 (l_n^S(\beta_0))^T I_n^{-1}(\beta_0, \beta_0) l_n^S(\beta_0), \quad (7.2.17)$$

где $I_n(\beta, \theta)$ — информация Фишера о векторе β , содержащаяся в векторе $S^n(\theta)$.

Найти $I_n(\beta, \beta)$ несложно. Действительно, в силу (7.2.16), (7.2.12)

и определения информации Фишера

$$I_n(\beta, \beta) = (2g(0) \mathbf{E}|\varepsilon_1|)^2 \mathbf{E}_\beta \mathbf{l}_n^S(\beta) (\mathbf{l}_n^S(\beta))^T.$$

Из (7.2.14) следует, что

$$\mathbf{E}_\beta \mathbf{l}_n^S(\beta) = \mathbf{0}, \quad \mathbf{E}_\beta \mathbf{l}_n^S(\beta) (\mathbf{l}_n^S(\beta))^T = \mathbf{K}_n(\beta), \quad (7.2.18)$$

где

$$\begin{aligned} \mathbf{K}_n(\beta) &= (k_{ij}^n(\beta)), \\ k_{ij}^n(\beta) &= \sum_{t=0}^{n-\max(i,j)-1} \delta_t(\beta) \delta_{t+j-i}(\beta) (n-t-\max(i,j)). \end{aligned} \quad (7.2.19)$$

Следовательно,

$$I_n(\beta, \beta) = (2g(0) \mathbf{E}|\varepsilon_1|)^2 \mathbf{K}_n(\beta). \quad (7.2.20)$$

Поскольку в силу (7.1.4) $\delta_t(\beta)$ экспоненциально убывают, то

$$n^{-1} k_{ij}^n(\beta) \rightarrow k_{ij}(\beta),$$

$k_{ij}(\beta)$ определены в (7.1.11). Следовательно,

$$n^{-1} \mathbf{K}_n(\beta) \rightarrow \mathbf{K}(\beta), \quad (7.2.21)$$

матрица $\mathbf{K}(\beta)$ определена в (7.1.11), и

$$I_n(\beta, \beta) \sim n (2g(0) \mathbf{E}|\varepsilon_1|)^2 \mathbf{K}(\beta). \quad (7.2.22)$$

При обычных условиях регулярности на $G(x)$ информационная матрица Фишера наблюдений u_{-g+1}, \dots, u_n равна

$$\mathbf{J}_n(\beta) \sim n I(g) \mathbf{E} \varepsilon_1^2 \mathbf{K}(\beta), \quad (7.2.23)$$

$I(g)$ — информация Фишера $g(x)$.

Легко проверить, что $I(g) \geq (2g(0))^2$, и равенство достигается лишь на распределении Лапласа. Поэтому из (7.2.22), (7.2.23) и неравенства $\mathbf{E} \varepsilon_1^2 \geq (\mathbf{E}|\varepsilon_1|)^2$ следует, что имеет место строгое неравенство $I_n(\beta, \beta) < \mathbf{J}_n(\beta)$ при любом распределении $G(x)$ (по крайней мере для больших n).

С помощью (7.2.20) перепишем (7.2.17) в виде

$$L_{n,S}(\beta_0) = (\mathbf{l}_n^S(\beta_0))^T \mathbf{K}_n^{-1}(\beta_0) \mathbf{l}_n^S(\beta_0). \quad (7.2.24)$$

При H_0 и условии 6.1.1 распределение $L_{n,s}(\beta_0)$ свободно от $G(x)$ при конечных n . Его аппроксимацию при больших n мы укажем в теореме 7.2.4. В важном частном случае $\beta_0 = \mathbf{0}$, соответствующем независимости наблюдений, (7.2.24) имеет вид

$$L_{n,s}(\mathbf{0}) = \sum_{t=1}^q (n-t)^{-1} \Gamma_{tn}^2(\mathbf{0}), \quad \Gamma_{tn}(\mathbf{0}) = \sum_{k=t+1}^n \text{sign}(u_{k-t} u_k).$$

Распределение статистики $L_{n,s}(\mathbf{0})$ при гипотезе может быть табулировано при любом q .

Следующая теорема дает предельное распределение $l_n^S(\beta_0)$ при гипотезе H_0 .

Теорема 7.2.3. При гипотезе $H_0: \beta = \beta_0$ и условии 6.1.1

$$n^{-1/2} l_n^S(\beta_0) \xrightarrow{d_{\beta_0}} N(\mathbf{0}, K(\beta_0)).$$

Доказательство. Ограничимся схемой рассуждения. Нам достаточно показать, что для любого вектора постоянных $c = (c_1, \dots, c_q)^T$ имеет место слабая сходимость скалярных величин

$$n^{-1/2} c^T l_n^S(\beta_0) \xrightarrow{d_{\beta_0}} N(0, c^T K(\beta_0) c).$$

Из (7.2.18) и (7.2.21) следует, что

$$E_{\beta_0} n^{-1/2} c^T l_n^S(\beta_0) = 0,$$

$$\begin{aligned} D_{\beta_0} \{n^{-1/2} c^T l_n^S(\beta_0)\} &= n^{-1} c^T E_{\beta_0} l_n^S(\beta_0) (l_n^S(\beta_0))^T c = \\ &= n^{-1} c^T K_n(\beta_0) c \rightarrow c^T K(\beta_0) c. \end{aligned}$$

Таким образом, сходимость моментов величины $n^{-1/2} c^T l_n^S(\beta_0)$ к моментам предельного гауссовского закона имеет место. Собственно асимптотическая гауссовость скалярной величины $n^{-1} c^T l_n^S(\beta_0)$ устанавливается по той же схеме, что при доказательстве теоремы 6.2.3. \square

Из теоремы 7.2.3 следует: при H_0 и условии 6.1.1

$$\begin{aligned} n^{-1/2} l_{jn}^S(\beta_0) &\xrightarrow{d_{\beta_0}} N(0, k_{jj}(\beta_0)), \\ k_{jj}(\beta_0) &= \sum_{i=0}^{\infty} \delta_i^2(\beta_0), \quad j = 1, \dots, q. \end{aligned} \tag{7.2.25}$$

Соотношение (7.2.25) позволяет использовать для отыскания постоянной в (7.2.9) и (7.2.10) нормальную аппроксимацию.

Обозначим $\chi^2(q)$ распределение хи-квадрат с q степенями свободы. Из теоремы 7.2.3 и (7.2.21) следует

Теорема 7.2.4. Пусть выполнено условие 6.1.1.

Тогда при гипотезе $H_0: \beta = \beta_0$ $L_{n,S}(\beta_0) \xrightarrow{d_{\beta_0}} \chi^2(q)$.

Если $\chi_{q,\alpha}$ — квантиль уровня α распределения $\chi^2(q)$, а $C(\theta)$ — множество $\theta = (\theta_1, \dots, \theta_q)^T$, для которых уравнение $x^q = \theta_1 x^{q-1} + \dots + \theta_q$ имеет корни, по модулю меньшие единицы, то множество

$$A_{n\alpha} = \{\theta: L_{n,S}(\theta) < \chi_{q,1-\alpha}\} \cap C(\theta) \quad (7.2.26)$$

является асимптотическим доверительным множеством для вектора β надежности $1 - \alpha$.

В заключение параграфа отметим полезный факт, аналогичный отмечавшемуся в § 6.2. А именно, пусть $\{\varepsilon_i\}$ будет последовательностью разнораспределенных случайных величин, удовлетворяющих условию

$$P\{\varepsilon_i < 0\} = P\{\varepsilon_i > 0\} = 1/2, \quad i \in \mathbb{Z}.$$

Тогда при H_0 :

— распределения $I_{j_n}^S(\beta_0)$, $I_n^S(\beta_0)$, $L_{n,S}(\beta_0)$ останутся свободными от $G(x)$ и такими же, как в случае одинаково распределенных $\{\varepsilon_i\}$;

— утверждения теорем 7.2.3 и 7.2.4 останутся верными;

— множество $A_{n\alpha}$ из (7.2.26) будет асимптотическим доверительным множеством для β надежности $1 - \alpha$.

Основные результаты этого параграфа опубликованы в [53].

§ 7.3. Теорема о равномерном стохастическом разложении: мощность знаковых тестов при близких альтернативах

В этом параграфе мы найдем мощность знаковых тестов из § 7.2 при близких альтернативах (7.1.12)

$$H_{1n}(a): \beta = \beta_n := \beta_0 + a n^{-1/2} + o(n^{-1/2}).$$

Аналогично случаю однопараметрической авторегрессии (см. § 6.4) соответствующее утверждение будет получено с помощью теоремы о равномерном линейном стохастическом разложении векторного случайного поля $n^{-1/2} I_n^S(\beta_n + n^{-1/2}\theta)$, справедливом для $|\theta| \leq \Theta_n$, где $\Theta_n \rightarrow \infty$ со степенной скоростью. Этот результат будет использован также в следующем параграфе при исследовании тестов для проверки линейных гипотез и в § 7.5 при построении и исследовании знаковых оценок.

Пусть вектор $I_n^S(\theta)$ определен соотношениями (7.2.15) и (7.2.12)–(7.2.11), матрица $K(\beta)$ — соотношением (7.1.11),

$$\lambda_S(\beta) := -2g(0) E|\varepsilon_1| K(\beta). \quad (7.3.1)$$

Теорема 7.3.1. Пусть выполнены условия 6.1.1–6.1.3.

Тогда если справедлива альтернатива $H_{1n}(\alpha)$ из (7.1.12) и $\Theta_n = \Theta n^\alpha$, где $\alpha < \frac{\delta}{2(1+2\delta)}$, $\delta = \min(\Delta, 1)$, $0 < \Theta < \infty$, то

$$\sup_{|\theta| \leq \Theta_n} \left| n^{-1/2} l_n^S(\beta_n + n^{-1/2}\theta) - n^{-1/2} l_n^S(\beta_n) - \lambda_S(\beta_0)\theta \right| = o_p(1). \quad (7.3.2)$$

Для случая, когда $\alpha = 0$ и корни характеристического уравнения (7.1.2) различны, теорема 7.3.1 была доказана в [53].

Доказательство. Идейно оно похоже на доказательство теоремы 6.3.1. Это позволяет сократить утомительные подробности.

Легко показать (см. [93, раздел 2]), что для любого β в некоторой окрестности β_0

$$|\delta_t(\beta) - \delta_t(\beta_0)| \leq c |\beta - \beta_0| b^t \quad (7.3.3)$$

при некоторых $c > 0$ и $0 < b < 1$, не зависящих от t и β . Теперь соотношения (7.3.3) и (7.1.4) влекут неравенства

$$|\delta_t(\beta_n)| \leq c_1 b^t, \quad |\delta_t(\beta_n + n^{-1/2}\theta)| \leq c_1 b^t, \quad (7.3.4)$$

$$|\delta_t(\beta_n + n^{-1/2}\theta) - \delta_t(\beta_n)| \leq c_2 \Theta_n n^{-1/2} b^t \quad (7.3.5)$$

при $n > n_0$, $|\theta| \leq \Theta_n$ и $t = 0, 1, 2, \dots$

Напомним, что в однопараметрической авторегрессии при $q = 1$

$$\delta_t(\beta) = \beta^t, \quad t = 0, 1, 2, \dots,$$

и, кроме того, для некоторого b_1 , $0 < b < b_1 < 1$ и $|\theta| \leq \Theta_n$

$$|(\beta_n + n^{-1/2}\theta)^t - \beta_n^t| \leq c_3 \Theta_n n^{-1/2} t b^{t-1} \leq c_4 n^{-1/2} b^t$$

(см. доказательство леммы 6.9.1). Таким образом, неравенства (7.3.4) и (7.3.5) устанавливают аналогичные свойства последовательности $\{\delta_t, t = 0, 1, 2, \dots\}$ при произвольном q .

Пусть $k_j^T(\beta) = (k_{j1}(\beta), \dots, k_{jq}(\beta))$ — j -я строка матрицы $K(\beta)$. Для доказательства теоремы достаточно показать

$$\begin{aligned} \sup_{|\theta| \leq \Theta_n} \left| n^{-1/2} l_{jn}^S(\beta_n + n^{-1/2}\theta) - n^{-1/2} l_{jn}^S(\beta_n) + 2g(0) E|\varepsilon_1| k_j^T(\beta) \theta \right| = \\ = o_p(1), \quad j = 1, \dots, q. \end{aligned} \quad (7.3.6)$$

Будем доказывать (7.3.6) для $j = 1$, это упростит обозначения.

Из определения (7.2.12) следует

$$n^{-1/2} l_{1n}^S(\beta_n + n^{-1/2}\theta) - n^{-1/2} l_{1n}^S(\beta_n) = \\ = n^{-1/2} \sum_{t=1}^{n-1} \delta_{t-1}(\beta_n + n^{-1/2}\theta) (\Gamma_{tn}(\beta_n + n^{-1/2}\theta) - \Gamma_{tn}(\beta_n)) + \quad (7.3.7)$$

$$+ n^{-1/2} \sum_{t=1}^{n-1} (\delta_{t-1}(\beta_n + n^{-1/2}\theta) - \delta_{t-1}(\beta_n)) \Gamma_{tn}(\beta_n). \quad (7.3.8)$$

Рассмотрим сначала (7.3.8).

Лемма 7.3.1. Пусть выполнено условие 6.1.1, верна $H_{1n}(a)$, $0 < \Theta < \infty$.

Тогда супремум (7.3.8) по $|\theta| \leq \Theta n^r$, $r < 1/2$, есть $o_p(1)$.

Лемма 7.3.1 доказывается аналогично лемме 6.9.1 с использованием неравенства (7.3.5) и соотношения (7.2.14).

Рассмотрим теперь (7.3.7). Рассуждая, как при доказательстве теоремы 6.4.1 (см. соотношение (6.9.4) и следующие за ним), перепишем это выражение в виде

$$\sum_{t=1}^{n-1} \delta_{t-1}(\beta_n + \eta_n) (4x_{tn}(\theta) + 4y_{tn}(\theta) - 2u_{tn}(\theta)),$$

где $\tilde{u}_t = (u_t, u_{t-1}, \dots, u_{t-g+1})^T$, $\Delta_k(x)$ определены в (6.9.3), $\eta_n = n^{-1/2}\theta$ и

$$x_{tn}(\theta) = n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n^T \tilde{u}_{k-t-1}) \Delta_k(\eta_n^T \tilde{u}_{k-1}) - \\ - \Delta_{k-t}(\eta_n^T \tilde{u}_{k-t-1}) G(\eta_n^T \tilde{u}_{k-1}) - \Delta_{k-t}(0) \Delta_k(0) + \Delta_{k-t}(0) G(0)],$$

$$y_{tn}(\theta) = n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n^T \tilde{u}_{k-t-1}) G(\eta_n^T \tilde{u}_{k-1}) - \Delta_{k-t}(0) G(0)],$$

$$u_{tn}(\theta) = n^{-1/2} \sum_{k=t+1}^n [\Delta_{k-t}(\eta_n^T \tilde{u}_{k-t-1}) - \Delta_{k-t}(0)] + \\ + n^{-1/2} \sum_{k=t+1}^n [\Delta_k(\eta_n^T \tilde{u}_{k-1}) - \Delta_k(0)].$$

Лемма 7.3.2. Пусть выполнены условия 6.1.2, 6.1.3, верна альтернатива $H_{1n}(a)$, $\Theta_n = \Theta n^{-r}$, $r > 1/4$, $0 < \Theta < \infty$.

Тогда:

$$1) \sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{n-1} \delta_{t-1}(\beta_n + \eta_n) x_{tn}(\theta) \right| = o_p(1);$$

$$2) \sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{n-1} \delta_{t-1}(\beta_n + \eta_n) u_{tn}(\theta) \right| = o_p(1).$$

Лемма 7.3.2 доказывается аналогично лемме 6.9.4 с использованием неравенства (7.3.4).

Л е м м а 7.3.3. Пусть выполнены условия 6.1.2, 6.1.3, верна альтернатива $H_{1n}(\alpha)$, $\delta = \min(\Delta, 1)$, $\Theta_n = \Theta n^{-r}$, $r > \frac{1+\delta}{2(1+2\delta)}$, $0 < \Theta < \infty$.

Тогда

$$\sup_{|\eta_n| \leq \Theta_n} \left| \sum_{t=1}^{n-1} \delta_{t-1}(\beta_n + \eta_n) y_{tn}(\theta) + \frac{1}{2} g(0) \mathbf{E}|\varepsilon_1| k_1^T(\beta_0) \theta \right| = o_p(1).$$

Лемма 7.3.3 доказывается аналогично лемме 6.9.6. При этом аналогом соотношения (6.9.35) является соотношение

$$g(0) \eta^T n^{-1/2} \sum_{t=1}^{N_n} \delta_{t-1}(\beta_n + \eta_n) \left(n^{-1} \sum_{k=t+1}^n \Delta_{k-t}(0) \tilde{u}_{k-1} \right).$$

Последнее выражение в силу представления (7.1.5) и закона больших чисел равно

$$-\frac{1}{2} g(0) \mathbf{E}|\varepsilon_1| \theta^T \sum_{t=1}^{\infty} \delta_{t-1}(\beta_0) (\delta_{t-1}(\beta_0), \delta_{t-2}(\beta_0), \dots, \delta_{t-q}(\beta_0))^T +$$

$$+ o_p(1) = -\frac{1}{2} g(0) \mathbf{E}|\varepsilon_1| k_1^T(\beta_0) \theta + o_p(1),$$

что и объясняет утверждение леммы.

Теперь из лемм 7.3.1–7.3.3, соотношений (7.3.7)–(7.3.8) и неравенства $\frac{\delta}{2(1+2\delta)} < \frac{1}{4}$ следует справедливость (7.3.5) для $j = 1$. Теорема доказана. \square

Аналогом следствия 6.4.1 является полезное в дальнейшем

С л е д с т в и е 7.3.1. Если выполнены условия 6.1.1–6.1.3, верна альтернатива $H_{1n}(\alpha)$, $\hat{\beta}_n$ — любая последовательность, для

которой $n^{1/2-\alpha}(\hat{\beta}_n - \beta_n) = O_p(1)$, $\alpha < \frac{\delta}{2(1+2\delta)}$, $\delta = \min(\Delta, 1)$,
то

$$n^{-1/2}I_n^S(\hat{\beta}_n) = n^{-1/2}I_n^S(\beta_n) + \lambda_S(\beta_0)\sqrt{n}(\hat{\beta}_n - \beta_n) + o_p(1). \quad (7.3.9)$$

Легко убедиться, что асимптотическое распределение статистики $n^{-1/2}I_n^S(\beta_n)$ при альтернативе $H_{1n}(\alpha)$ и условии 6.1.1 то же самое, что и статистики $n^{-1/2}I_n^S(\beta_0)$ при гипотезе H_0 . С учетом этого замечания из теоремы 7.2.3 и следствия 7.3.1 следует

Теорема 7.3.2. Пусть выполнены условия 6.1.1–6.1.3. Тогда при альтернативе $H_{1n}(\alpha)$

$$n^{-1/2}I_n^S(\beta_0) = n^{-1/2}I_n^S(\beta_n) - \lambda_S(\beta_0)\alpha + o_p(1),$$

и, следовательно,

$$n^{-1/2}I_n^S(\beta_0) \xrightarrow{d_{P_n}} N(-\lambda_S(\beta_0)\alpha, K(\beta_0)).$$

Итак, асимптотический сдвиг статистики $n^{-1/2}I_n^S(\beta_0)$ при альтернативе $H_{1n}(\alpha)$ равен

$$-\lambda_S(\beta_0)\alpha = 2g(0)E|\varepsilon_1|K(\beta_0)\alpha.$$

Следовательно, асимптотический сдвиг $n^{-1/2}I_{jn}^S(\beta_0)$, $j = 1, \dots, q$ (напомним, что $I_{jn}^S(\beta_0)$ является тестовой статистикой для проверки гипотезы $H_0: \beta = \beta_0$ против альтернатив H_j^+ , H_j^- , H_{1j} ; см. (7.2.9), (7.2.10)), при $H_{1n}(\alpha)$ с $\alpha = (0, \dots, 0, \alpha_j, 0, \dots, 0)^T$ равен

$$2g(0)E|\varepsilon_1|k_{jj}(\beta_0)\alpha_j = 2g(0)E|\varepsilon_1|\alpha_j \sum_{i=0}^{\infty} \delta_i^2(\beta_0).$$

Обозначим $\chi^2(q, \lambda^2)$ нецентральное хи-квадрат распределение с q степенями свободы и параметром нецентральности λ^2 . Теорема 7.3.2 влечет следующее утверждение для $L_{n,S}$ из (7.2.24).

Теорема 7.3.3. Пусть выполнены условия 6.1.1–6.1.3 и верна альтернатива $H_{1n}(\alpha)$.

Тогда

$$L_{n,S}(\beta_0) \xrightarrow{d_{P_n}} \chi^2(q, \lambda_S^2),$$

где параметр нецентральности $\lambda_S^2 = (2g(0)E|\varepsilon_1|)^2 \alpha^T K(\beta_0)\alpha$.

Отметим, что в силу (7.2.22) $\lambda_S^2 \sim n^{-1}\alpha^T I_n(\beta_0, \beta_0)\alpha$, где $I_n(\beta, \theta)$ — информация Фишера о векторе β , содержащаяся в $S^n(\theta)$.

Теорема 7.3.3 позволяет находить АОЭ знакового теста со статистикой $L_{n,S}(\beta_0)$ относительно известных тестов. В частности, при $E\varepsilon_1 = 0$, $E\varepsilon_1^2 < \infty$ и $H_{1n}(\mathbf{a})$

$$\sqrt{n}(\hat{\beta}_{n,LS} - \beta_0) \xrightarrow{d_{\beta_n}} N(\mathbf{a}, \mathbf{K}^{-1}(\beta_0)), \quad (7.3.10)$$

откуда

$$L_{n,LS}(\beta) := n(\hat{\beta}_{n,LS} - \beta_0)^T \mathbf{K}(\beta_0)(\hat{\beta}_{n,LS} - \beta_0) \xrightarrow{d_{\beta_n}} \chi^2(q, \lambda_{LS}^2),$$

где $\lambda_{LS}^2 = \mathbf{a}^T \mathbf{K}(\beta_0) \mathbf{a}$. Доказательство (7.3.10) аналогично доказательству теоремы 5.3.2.

Поэтому АОЭ знакового теста со статистикой $L_{n,S}(\beta_0)$ относительно общеупотребительного теста со статистикой $L_{n,LS}(\beta_0)$ есть

$$e_{S,LS} = (2g(0)E|\varepsilon_1|)^2$$

и такая же, как в однопараметрической схеме при $q = 1$ (см. § 6.5).

§ 7.4. Проверка линейных гипотез

В этом параграфе нашей целью будет построение знаковых тестов для проверки линейных гипотез относительно вектора β и исследование асимптотической мощности этих тестов при близких альтернативах.

Пусть вектор β разбит на подвекторы $\beta^T = (\beta^{(1)T}, \beta^{(2)T})$, где векторы $\beta^{(i)}$, $i = 1, 2$, имеют размерности m и $q - m$ соответственно, $1 \leq m < q$. Мы рассматриваем линейную гипотезу вида

$$H'_0: \beta^{(2)} = \beta_0^{(2)}, \quad (7.4.1)$$

где $\beta_0^{(2)}$ — известный вектор, а $\beta^{(1)}$ — мешающий параметр.

При $\beta_0^{(2)} = \mathbf{0}$ гипотеза H'_0 из (7.4.1) означает, что истинный порядок уравнения (7.1.1) равен m и $m < q$.

Асимптотическую мощность строящихся далее тестов будем исследовать при близких альтернативах

$$H_{1n}(\mathbf{a}): \beta = \beta_n := \beta_0 + \mathbf{a}n^{-1/2} + o(n^{-1/2}), \quad (7.4.2)$$

где $\beta_0^T = (\beta^{(1)T}, \beta_0^{(2)T})$, $\mathbf{a}^T = (\mathbf{a}^{(1)T}, \mathbf{a}^{(2)T})$ — постоянный вектор с подвекторами размерностей m и $q - m$ соответственно. Мы допускаем, таким образом, что при альтернативе $H_{1n}(\mathbf{a})$ мешающий параметр $\beta^{(1)}$ меняется на величину порядка $O(n^{-1/2})$.

Перейдем к построению тестовых статистик. Определим матрицы $\mathbf{J}(\beta)$ порядка $m \times m$, $\mathbf{M}(\beta)$ порядка $(q - m) \times (q - m)$ и $\mathbf{B}^T(\beta)$

порядка $m \times (q - m)$ равенством

$$K(\beta) = \begin{pmatrix} J(\beta) & B^T(\beta) \\ B(\beta) & M(\beta) \end{pmatrix},$$

где матрица $K(\beta)$ определена в (7.1.11). Пусть

$$V(\beta) = \begin{pmatrix} J^{-1/2}(\beta) & -J^{-1}(\beta)B^T(\beta)C^{-1/2}(\beta) \\ 0 & C^{-1/2}(\beta) \end{pmatrix},$$

где $C(\beta) = M(\beta) - B(\beta)J^{-1}(\beta)B^T(\beta)$.

Легко убедиться, что

$$V(\beta)V^T(\beta) = K^{-1}(\beta). \quad (7.4.3)$$

Пусть $\hat{\beta}_n$ — последовательность, зависящая от наблюдений u_{1-q}, \dots, u_n и гипотетического вектора $\beta_0^{(2)}$ из (7.4.1), такая, что $\sqrt{n}(\hat{\beta}_n - \beta_n) = O_p(1)$ при $H_{1n}(\alpha)$. В частности, можно взять $\hat{\beta}_n$ равным вектору $\hat{\beta}_{n0}$, где

$$\hat{\beta}_{n0}^T := (\hat{\beta}_n^{(1)T}, \beta_0^{(2)T}), \quad (7.4.4)$$

а $\hat{\beta}_n^{(1)}$ — \sqrt{n} -состоятельная оценка $\beta^{(1)}$. При $E\epsilon_1 = 0$ и $E\epsilon_1^2 < \infty$ в силу (7.3.10) в качестве $\hat{\beta}_n$ можно взять $\hat{\beta}_{n,LS}$.

Обозначим \hat{V}_n состоятельную при $H_{1n}(\alpha)$ оценку матрицы $V(\beta_0)$, например, $\hat{V}_n = V(\hat{\beta}_n)$.

Аналогично вектору β разобьем вектор $\hat{\beta}_n$ на подвекторы $\hat{\beta}_n^T = (\hat{\beta}_n^{(1)T}, \hat{\beta}_n^{(2)T})$ размерностей m и $q - m$ соответственно.

Пусть вектор $l_n^S(\theta)$ определен соотношениями (7.2.15), (7.2.12). Статистикой теста для проверки гипотезы H'_0 из (7.4.1) возьмем

$$L_{n,S}^\pi(\hat{\beta}_{n0}) := n^{-1} |\pi \circ \hat{V}_n^T l_n^S(\hat{\beta}_{n0})|^2, \quad (7.4.5)$$

где π — ортогональная проекция на последние $q - m$ компонент.

Предельное распределение статистики $L_{n,S}^\pi(\hat{\beta}_{n0})$ при альтернативе $H_{1n}(\alpha)$ из (7.4.2) описывается следующей теоремой.

Теорема 7.4.1. Пусть выполнены условия 6.1.1–6.1.3.

Тогда при альтернативе $H_{1n}(\alpha)$ из (7.4.2)

$$L_{n,S}^\pi(\hat{\beta}_{n0}) \xrightarrow{d_{\mathcal{D}_n}} \chi^2(q - m, \lambda_{S,\pi}^2),$$

где параметр нецентральности

$$\lambda_{S,\pi}^2 = (2g(0) \mathbf{E}|\varepsilon_1|)^2 \mathbf{a}^{(2)T} \mathbf{C}(\beta_0) \mathbf{a}^{(2)}.$$

В частности, при гипотезе H'_0 из (7.4.1)

$$L_{n,S}^{\pi}(\hat{\beta}_{n0}) \xrightarrow{d_{P_0}} \chi^2(q-m).$$

Доказательство. При $H_{1n}(\mathbf{a})$ по предположению $\sqrt{n}(\hat{\beta}_n - \beta_n) = O_p(1)$. Следовательно, и для вектора $\hat{\beta}_{n0}$ из (7.4.4) $\sqrt{n}(\hat{\beta}_{n0} - \beta_n) = O_p(1)$. Поэтому в силу следствия 7.3.1, состоятельности \hat{V}_n и соотношения (7.4.3) имеем

$$\begin{aligned} n^{-1/2} \hat{V}_n^T L_n^S(\hat{\beta}_{n0}) &= n^{-1/2} \hat{V}_n^T L_n^S(\beta_n) - 2g(0) \mathbf{E}|\varepsilon_1| \times \\ &\quad \times \hat{V}_n^T \mathbf{K}(\beta_0) \sqrt{n}(\hat{\beta}_{n0} - \beta_n) + o_p(1) = \\ &= n^{-1/2} \mathbf{V}^T(\beta_0) L_n^S(\beta_n) - 2g(0) \mathbf{E}|\varepsilon_1| \mathbf{V}^T(\beta_0) \mathbf{K}(\beta_0) \times \\ &\quad \times \sqrt{n}(\hat{\beta}_{n0} - \beta_n) + o_p(1) = \\ &= n^{-1/2} \mathbf{V}^T(\beta_0) L_n^S(\beta_n) - 2g(0) \mathbf{E}|\varepsilon_1| \mathbf{V}^{-1}(\beta_0) \times \\ &\quad \times (\sqrt{n}(\hat{\beta}_n^{(1)} - \beta_n^{(1)})^T, -\mathbf{a}^{(2)T})^T + o_p(1). \end{aligned} \quad (7.4.6)$$

Поскольку

$$\mathbf{V}^{-1}(\beta_0) = \begin{pmatrix} \mathbf{J}^{1/2}(\beta_0) & \mathbf{J}^{-1/2}(\beta_0) \mathbf{B}^T(\beta_0) \\ \mathbf{0} & \mathbf{C}^{1/2}(\beta_0) \end{pmatrix},$$

то из (7.4.6) следует

$$\begin{aligned} \pi \circ \hat{V}_n^T(\hat{\beta}_n) L_n^S(\beta_n) &= \pi \circ \mathbf{V}^T(\beta_0) L_n^S(\beta_n) + \\ &\quad + 2g(0) \mathbf{E}|\varepsilon_1| \mathbf{C}^{1/2}(\beta_0) \mathbf{a}^{(2)} + o_p(1). \end{aligned} \quad (7.4.7)$$

Утверждение теоремы следует теперь из сходимости при альтернативе $H_{1n}(\mathbf{a})$

$$n^{-1/2} \pi \circ \mathbf{V}(\beta_0) L_n^S(\beta_n) \xrightarrow{d_{P_0}} N(0, \mathbf{E}_{q-m})$$

(\mathbf{E}_{q-m} — единичная матрица порядка $q-m$) и соотношения (7.4.7). Теорема 7.4.1 доказана. \square

В [92] для схемы (7.1.1), удовлетворяющей условиям ЛАН, были построены, в частности, следующие тесты для проверки гипотезы H'_0 . Пусть

$$\psi_n(\theta) = n^{-1/2} \sum_{k=1}^n \psi(\varepsilon_k(\theta)) \tilde{\mathbf{u}}_{k-1},$$

где

$$\tilde{u}_{k-1} = (u_{k-1}, \dots, u_{k-q})^T, \quad \varepsilon_k(\theta) = u_k - \theta^T \tilde{u}_{k-1},$$

$$\int \psi^2 dG < \infty, \quad \int \psi dG = 0.$$

Пусть $\sigma^2 = \mathbf{E}\varepsilon_1^2$, $\mathbf{W} = \sigma^{-1}\mathbf{V}$ и $\widehat{\mathbf{W}}_n$ — состоятельная оценка \mathbf{W} . Тестовой статистикой бралась

$$\left(\int \psi^2 dG \right)^{-1} \left| \pi \circ \widehat{\mathbf{W}}_n^T \psi_n(\widehat{\beta}_{n0}) \right|^2. \quad (7.4.8)$$

Тогда [92, теорема 4.1]: если ψ и G дважды непрерывно дифференцируемы с ограниченными производными и $\mathbf{E}\varepsilon_1^4 < \infty$, то при $H_{1n}(\mathbf{a})$ статистика (7.4.8) слабо сходится к $\chi^2(q - m, \delta^2(\psi))$, где

$$\delta^2(\psi) = \mathbf{a}^{(2)T} \mathbf{C}(\beta) \mathbf{a}^{(2)} \left(\int \psi g' / g dG \right)^2 \left(\int \psi^2 dG \right)^{-1} \sigma^2.$$

При $\psi(x) = x/\sigma^2$ тест со статистикой (7.4.8) асимптотически эквивалентен тесту, основанному на оценке наименьших квадратов. В этом случае $\delta^2(\psi) = \mathbf{a}^{(2)T} \mathbf{C}(\beta_0) \mathbf{a}^{(2)}$, и АОЭ теста со статистикой (7.4.5) относительно теста со статистикой (7.4.8) есть опять, как в § 7.3,

$$\frac{\lambda_{5\pi}^2}{\delta^2(\psi)} = (2g(0) \mathbf{E}|\varepsilon_1|)^2 = e_{S,LS}.$$

При $\psi = g'/g$ тест со статистикой (7.4.8) является асимптотически максиминным в локальном смысле ([92, теорема 3.3], определения см. там же). В этом случае

$$\sigma^2(\psi) = \mathbf{a}^{(2)T} \mathbf{C}(\beta_0) \mathbf{a}^{(2)} I(g) \sigma^2,$$

и АОЭ знакового теста (7.4.4) относительно такого асимптотически оптимального параметрического теста равна

$$(2g(0) \mathbf{E}|\varepsilon_1|)^2 / (I(g) \sigma^2).$$

Последняя величина всегда меньше единицы (см. замечание в § 7.2 после соотношения (7.2.23)). Подчеркнем однако, что тесты из [92] и другие известные тесты для проверки линейных гипотез (см. например, [2, 45; 12, 71] и имеющуюся там библиографию) построены в предположении $\mathbf{E}\varepsilon_1^2 < \infty$ или при еще более жестких моментных ограничениях. В отличие от известных тестов, тест со статистикой (7.4.5) применим и в тех случаях, когда условие $\mathbf{E}\varepsilon_1^2 < \infty$ не выполнено.

§ 7.5. Знаковые оценки параметров

В этом параграфе поставим задачей оценить неизвестный вектор β в схеме (7.1.1) с помощью статистик знаковых тестов из § 7.2.

Возможны разные подходы к определению непараметрических знаковых оценок, в случае однопараметрической авторегрессии мы рассмотрели их в § 6.6. К сожалению, не все из них легко распространяются на многопараметрическую авторегрессию. Например, трудности возникают с аналогом оценки $\hat{\beta}_{n,s}$, получаемой решением уравнения (6.6.6). С подобной оценки и начнем.

7.5.1. Знаковая оценка $\hat{\beta}_{n,s}$. Будем предполагать выполненными условия 6.1.1–6.1.3. Рассмотрим вектор $l_n^S(\theta)$ из (7.2.15), компоненты которого определены соотношением (7.2.12). Следуя однопараметрическому случаю, естественно брать оценкой вектора β решение уравнения

$$l_n^S(\theta) \div 0. \quad (7.5.1)$$

В (7.5.1) символ \div означает, как и в одномерном случае, переход векторной функции $l_n^S(\theta)$ через нуль. Однако, в отличие от одномерного случая, в \mathbb{R}^q через точку можно «перейти» по многим направлениям, а потому соответствующее определение может быть дано несколькими способами. Один из них — простейший — состоит в следующем.

По определению векторная функция $l_n^S(\theta)$ переходит в точке $\theta_0 = (\theta_{10}, \dots, \theta_{q0})^T$ через нуль, если при любом $j = 1, \dots, q$ одномерная функция $l_{jn}^S(\theta_{10}, \dots, \theta_{j-1,0}, \theta_j, \theta_{j+1,0}, \dots, \theta_{q0})$ одной переменной θ_j переходит в точке θ_{j0} через нуль.

Аналогично одномерному случаю (см. доказательство теоремы 6.6.2) показывается, что для любого решения $\hat{\beta}_n$ уравнения (7.5.1)

$$n^{-1/2} l_n^S(\hat{\beta}_n) = o_p(1). \quad (7.5.2)$$

Поэтому, если существует \sqrt{n} -состоятельное решение $\hat{\beta}_{n,s}$ уравнения (7.5.1) (а только такие решения нас и интересуют), то в силу (7.5.2) и следствия 7.3.1 (с $\beta_n = \beta_0 = \beta$, $\hat{\beta}_n = \hat{\beta}_{n,s}$ и $\alpha = 0$)

$$\sqrt{n}(\hat{\beta}_{n,s} - \beta) = -\lambda_S^{-1}(\beta) n^{-1/2} l_n^S(\beta) + o_p(1), \quad (7.5.3)$$

где матрица $\lambda_S(\beta)$ определена в (7.3.1).

Соотношение (7.5.3) и теорема 7.2.3 непосредственно влекут

$$\sqrt{n}(\hat{\beta}_{n,s} - \beta) \xrightarrow{d_p} N(0, \Sigma_S(\beta)), \quad (7.5.4)$$

где

$$\Sigma_S(\beta) = (2g(0) \mathbf{E}|\varepsilon_1|)^{-2} K^{-1}(\beta),$$

матрица $K(\beta)$ определена в (7.1.11). Соотношения (7.5.3)–(7.5.4) обобщают утверждения теоремы 6.6.2 на случай произвольного q .

Отметим, что в силу (7.2.22) предельная ковариация $\Sigma_S(\beta)$ в (7.5.4) совпадает с предельным значением матрицы $nI_n^{-1}(\beta, \beta)$, где $I_n(\beta, \theta)$ — информация Фишера, содержащаяся в векторе $S^n(\theta)$, относительно векторного параметра β .

Доказательство существования \sqrt{n} -состоятельного решения уравнения (7.5.1) не очень простая задача даже при наличии теоремы 7.3.1, и в этой книге мы ею заниматься не будем. Более того, есть и другая неприятность, связанная с определением (7.5.1), — необходимость выбора нужного (т.е. \sqrt{n} -состоятельного) решения среди всех решений уравнения (7.5.1).

Обратимся к другой возможности. Определим оценку как решение экстремальной задачи

$$|l_n^S(\theta)| \Rightarrow \inf_{\theta \in \mathbb{R}^q}. \quad (7.5.5)$$

Легко показать, что \sqrt{n} -состоятельные решения задач (7.5.1) и (7.5.5) асимптотически эквивалентны в смысле предельного распределения. Действительно, пусть вектор $\tilde{\beta}_n$ определяется соотношением

$$\sqrt{n}(\tilde{\beta}_n - \beta) = -\lambda_S^{-1}(\beta)n^{-1/2}l_n^S(\beta). \quad (7.5.6)$$

В силу теоремы 7.2.3 вектор $\sqrt{n}(\tilde{\beta}_n - \beta)$ слабо сходится к гауссовскому, а потому $\sqrt{n}(\tilde{\beta}_n - \beta) = O_p(1)$. Подставляя $\tilde{\beta}_n$ вместо $\hat{\beta}_n$ в соотношение (7.3.9) (при $\beta_n = \beta_0 = \beta$ и $\alpha = 0$), убеждаемся, что

$$n^{-1/2}l_n^S(\tilde{\beta}_n) = o_p(1). \quad (7.5.7)$$

Из (7.5.7) следует, что для любого решения $\hat{\beta}_n$ задачи (7.5.5) и давно $n^{-1/2}l_n^S(\hat{\beta}_n) = o_p(1)$. Если же, вдобавок, решение задачи (7.5.5) является \sqrt{n} -состоятельным (обозначим такое решение $\hat{\beta}_{n,S}$), то для него в силу следствия 7.3.1 (опять с $\beta_n = \beta_0 = \beta$ и $\alpha = 0$) выполнены соотношения (7.5.3)–(7.5.4).

Итак, задачи (7.5.1) и (7.5.5) определяют асимптотически эквивалентные \sqrt{n} -состоятельные знаковые оценки (разумеется, если такие существуют).

Понятно, что доказать существование \sqrt{n} -состоятельного решения задачи (7.5.5) и выделить среди всех решений нужное — задача той же трудности, что и при определении (7.5.1). Подобные трудности возникают всегда при использовании M -оценок. Однако то, с чем относительно легко справиться в случае гладких или негладких, но выпуклых целевых функций, представляет гораздо более серьезную задачу применительно к $l_n^S(\theta)$.

К счастью, есть несколько возможностей обойти указанные трудности.

7.5.2. Знаковая оценка $\tilde{\beta}_{n,S}$. Будем строить оценку не с помощью функции $l_n^S(\theta)$, а с помощью более простой кусочно постоянной функции

$$l_n^S(\hat{\beta}_n, \theta) = (l_{1n}^S(\hat{\beta}_n, \theta), \dots, l_{qn}^S(\hat{\beta}_n, \theta))^T,$$

где

$$l_{jn}^S(\hat{\beta}_n, \theta) = \sum_{t=0}^{n-j-1} \delta_t(\hat{\beta}_n) \Gamma_{t+j,n}(\theta), \quad j = 1, \dots, q,$$

а $\hat{\beta}_n$ — какая-нибудь предварительная \sqrt{n} -состоятельная оценка β . Напомним, что величины $\Gamma_{tn}(\theta)$ определены в (7.2.11), а последовательность $\{\delta_t(\beta), t = 1 - q, 2 - q, \dots\}$ — в (7.1.3).

Назовем непараметрической знаковой оценкой вектора β (и обозначим ее $\tilde{\beta}_{n,S}$) любую случайную величину, являющуюся решением задачи

$$|l_n^S(\hat{\beta}_n, \theta)| \Rightarrow \inf_{\theta \in \mathbb{Q}(\hat{\beta}_n)}, \quad (7.5.8)$$

где $\mathbb{Q}(\hat{\beta}_n) = \{\theta: |\sqrt{n}(\theta - \hat{\beta}_n)| \leq \log n\}$. Свойства $\tilde{\beta}_{n,S}$ описывает следующая теорема.

Т е о р е м а 7.5.1. Если выполнены условия 6.1.1–6.1.3, $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$, то справедливы соотношения

$$\sqrt{n}(\tilde{\beta}_{n,S} - \beta) = -\lambda_S^{-1}(\beta) n^{-1/2} l_n^S(\beta) + o_p(1), \quad (7.5.9)$$

$$\begin{aligned} \sqrt{n}(\tilde{\beta}_{n,S} - \beta) &\xrightarrow{d} N(0, \Sigma_S(\beta)), \\ \Sigma_S(\beta) &= (2g(0) \mathbb{E}|\varepsilon_1|)^{-2} K^{-1}(\beta). \end{aligned} \quad (7.5.10)$$

Д о к а з а т е л ь с т в о. В условиях теоремы

$$\sup_{|\theta| \leq \Theta \log n} |n^{-1/2} l_n^S(\hat{\beta}_n, \beta + n^{-1/2} \theta) - n^{-1/2} l_n^S(\beta + n^{-1/2} \theta)| = o_p(1)$$

(см. [54]), что совместно с теоремой 7.3.1 (при $\beta_n = \beta_0 = \beta$) влечет стохастическое разложение

$$\sup_{|\theta| \leq \Theta \log n} |n^{-1/2} l_n^S(\hat{\beta}_n, \beta + n^{-1/2} \theta) - l_n^S(\beta) - \lambda_S(\beta) \theta| = o_p(1). \quad (7.5.11)$$

Из (7.5.11) аналогично следствию 7.3.1 получаем: если β_n^* — любая статистика, для которой

$$\sqrt{n}(\beta_n^* - \beta) / \log n = O_p(1),$$

то

$$n^{-1/2} l_n^S(\hat{\beta}_n, \beta_n^*) = n^{-1/2} l_n^S(\beta) + \lambda_S(\beta) \sqrt{n}(\beta_n^* - \beta) + o_p(1). \quad (7.5.12)$$

С вероятностью, стремящейся к единице, $\tilde{\beta}_n$ из (7.5.6) попадает в множество $\mathbb{Q}(\hat{\beta}_n)$, и в силу (7.5.12) $n^{-1/2} I_n^S(\tilde{\beta}_n, \tilde{\beta}_n) = o_p(1)$. Следовательно, и по-прежнему

$$n^{-1/2} I_n^S(\hat{\beta}_n, \tilde{\beta}_{n,S}) = o_p(1). \quad (7.5.13)$$

В силу определения $\tilde{\beta}_{n,S}$

$$\sqrt{n}(\tilde{\beta}_{n,S} - \beta) / \log n = O_p(1),$$

что совместно с (7.5.12) и (7.5.13) влечет (7.5.9). Соотношение (7.5.10) следует из (7.5.9) и теоремы 7.2.3. Теорема 7.5.1 полностью доказана. \square

Численное решение задачи (7.5.8) не представляется чрезмерно сложным, при малых q его можно искать простым перебором, так как $I_n^S(\hat{\beta}_n, \theta)$ принимает конечное множество значений внутри $\mathbb{Q}(\hat{\beta}_n)$.

7.5.3. Знаковая оценка $\beta_{n,S}^*$. Рассмотрим другую оценку, $\beta_{n,S}^*$, асимптотически эквивалентную в смысле предельного распределения решению задачи (7.5.8). А именно, пусть \hat{e}_n — состоятельная оценка матрицы $\lambda_S(\beta)$,

$$\beta_{n,S}^* := \hat{\beta}_n - (n\hat{e}_n)^{-1} I_n^S(\hat{\beta}_n). \quad (7.5.14)$$

Следствие 7.3.1 (с $\beta_n = \beta_0 = \beta$ и $\alpha = 0$) влечет следующую теорему.

Теорема 7.5.2. Пусть $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$, \hat{e}_n — состоятельная оценка матрицы $\lambda_S(\beta)$ и выполнены условия 6.1.1–6.1.3.

Тогда

$$\sqrt{n}(\beta_{n,S}^* - \beta) = -\lambda_S(\beta) n^{-1/2} I_n^S(\beta) + o_p(1) \xrightarrow{d} N(0, \Sigma_S(\beta)),$$

$$\Sigma_S(\beta) = (2g(0)) E|\varepsilon_1|^{-2} K^{-1}(\beta).$$

Отметим, что $\beta_{n,S}^*$ не является решением задачи (7.5.8), но удовлетворяет условию

$$n^{-1/2} I_n^S(\hat{\beta}_n, \beta_{n,S}^*) = o_p(1).$$

Для построения оценки \hat{e}_n достаточно состоятельно оценить постоянную $-2g(0) E|\varepsilon_1|$ и матрицу $K(\beta)$. Если $\hat{\beta}_n = (\hat{\beta}_{1n}, \dots, \hat{\beta}_{qn})^T$ и вектор $\hat{\beta}_n^1 = (\hat{\beta}_{1n}^1, \hat{\beta}_{2n}^1, \dots, \hat{\beta}_{qn}^1)^T$ отличается от $\hat{\beta}_n$ лишь первой компонентой $\hat{\beta}_{1n}^1 = \hat{\beta}_{1n} + hn^{-1/2}$, $h \neq 0$ — постоянная, то состоятельной оценкой постоянной $-2g(0) E|\varepsilon_1|$ будет

$$(hn^{1/2})^{-1} (\Gamma_{1n}(\hat{\beta}_n^1) - \Gamma_{1n}(\hat{\beta}_n)).$$

Состоятельной оценкой матрицы $K(\beta)$ будет $K(\hat{\beta}_n)$.

Другая состоятельная оценка $\lambda_S(\beta)$ имеет вид

$$\hat{\epsilon}_n = -2\hat{g}_n m_n K(\hat{\beta}_n),$$

где

$$\hat{g}_n = (nh_n)^{-1} \sum_{i=1}^n \varphi(h_n^{-1}(u_i - \hat{\beta}_n^T \tilde{u}_{i-1}))$$

— оценка Парзена—Розенблатта с гауссовским ядром $\varphi(x)$, $h_n = n^{-\alpha}$, $0 < \alpha < 1/4$, $\tilde{u}_{k-1} = (u_{k-1}, \dots, u_{k-q})^T$,

$$\hat{m}_n = n^{-1} \sum_{i=1}^n |u_i - \hat{\beta}_n^T \tilde{u}_{i-1}|.$$

Отметим, что оценки $\tilde{\beta}_{n,S}$ (и, разумеется, $\beta_{n,S}^*$, $\hat{\beta}_{n,S}$) и $\hat{\beta}_{n,LS}$, $\hat{\beta}_{n,LD}$ имеют асимптотические ковариационные матрицы, лишь множителем различающиеся между собой. А именно (см., [2, 45, теорема 5.5.7]), если $E\epsilon_1 = 0$, $E\epsilon_1^2 < \infty$, то

$$\sqrt{n}(\tilde{\beta}_{n,LS} - \beta) \xrightarrow{d} N(0, K^{-1}(\beta)).$$

Если же $E\epsilon_1 = 0$, $E\epsilon_1^2 < \infty$, $G(0) = 1/2$, $g(x)$ непрерывна в некоторой окрестности нуля и $g(0) > 0$, то (см. [102])

$$\sqrt{n}(\hat{\beta}_{n,LD} - \beta) \xrightarrow{d} N(0, [(2g(0))^2 E\epsilon_1^2]^{-1} K^{-1}(\beta)).$$

Если под АОЭ оценки $\tilde{\beta}_{n,S}$ относительно $\hat{\beta}_{n,LS}$ и $\hat{\beta}_{n,LD}$ понимать отношение коэффициентов перед $K^{-1}(\beta)$ в предельных ковариационных матрицах оценок, то такая АОЭ оказывается равной соответственно $e_{S,LS} = (2g(0) E|\epsilon_1|)^2$ и $e_{S,LD} = (E|\epsilon_1|)^2 / E\epsilon_1^2$ и такой же, как в случае однопараметрической авторегрессии (см. § 6.6 и § 6.5).

Завершим параграф замечаниями о предварительной \sqrt{n} -состоятельной оценке, участвующей в определении знаковых оценок $\hat{\beta}_{n,S}$ и $\beta_{n,S}^*$. Простейший подходящий пример — оценка наименьших квадратов $\hat{\beta}_{n,LS}$, которая \sqrt{n} -состоятельна при симметричной $G(x)$ и $E \log^+ |\epsilon| < \infty$ (см. [111]). К сожалению, $\hat{\beta}_{n,LS}$ имеет неограниченный функционал влияния, а $\tilde{\beta}_{n,S}$ будет его наследовать: для однопараметрической авторегрессии мы показали это в п. 5.6.1 и 6.7.2, а в общем случае рассмотрим эти вопросы в следующем параграфе. В связи с этим представляет интерес класс взвешенных оценок наименьших модулей. В однопараметрической авторегрессии такие оценки рассматривались в п. 5.3.3. Итак, будем брать оценкой β измеримое

решение задачи

$$\sum_{k=1}^n |\varphi(\tilde{u}_{k-1})(u_k - \theta^T \tilde{u}_{k-1})| \Rightarrow \inf_{\theta \in \mathbb{R}^q}. \quad (7.5.15)$$

Целевая функция в (7.5.15) выпукла вниз, и решение (обозначим его $\hat{\beta}_{n,LDW}$) всегда существует. Предположим, что выполнены следующие условия.

У с л о в и е 7.5.1.

$$\begin{aligned} \mathbb{E}\varepsilon_1 &= 0, & \mathbb{E}_\beta |\varphi(\tilde{u}_1)| &< \infty, \\ -\infty &< \lambda_{LDW}(\beta) := -2g(0) \mathbb{E}_\beta |\varphi(\tilde{u}_1)| \tilde{u}_1 \tilde{u}_1^T < 0, \\ 0 &< K_{LDW}(\beta) := \mathbb{E}_\beta \varphi^2(\tilde{u}_1) \tilde{u}_1 \tilde{u}_1^T < \infty. \end{aligned}$$

У с л о в и е 7.5.2. $\max_{1 \leq k \leq n} n^{-1/2} |\varphi(\tilde{u}_k) \tilde{u}_k| = o_p(1).$

У с л о в и е 7.5.3. $G(0) = 1/2$, $g(x)$ непрерывна в окрестности нуля.

Следуя схеме доказательства [102, теорема 2], получаем: при условиях 7.5.1–7.5.3

$$\sqrt{n}(\hat{\beta}_{n,LDW} - \beta) \xrightarrow{d_p} N(0, \Sigma_{LDW}(\beta)),$$

где

$$\Sigma_{LDW}(\beta) = \lambda_{LDW}^{-1}(\beta) K_{LDW}(\beta) \lambda_{LDW}^{-1}(\beta).$$

Уместно отметить: при изучении взвешенных оценок наименьших модулей в случае однопараметрической авторегрессии в п. 5.3.3 мы переходили от экстремальной задачи (5.5.8) к эквивалентному уравнению (5.5.9), которое и изучали. В случае многопараметрической авторегрессии переходить от выпуклой задачи (7.5.15) к аналогичному (5.5.9) уравнению

$$\sum_{k=1}^n \tilde{u}_{k-1} |\varphi(\tilde{u}_{k-1})| (u_k - \theta^T \tilde{u}_{k-1}) \div 0$$

нецелесообразно: это приведет к необходимости преодолевать те же трудности, что и при определении знаковой оценки в виде решения уравнения (7.5.1).

Результаты этого параграфа были опубликованы в [54].

§ 7.6. Функционалы влияния оценок в многопараметрической авторегрессии

Пусть наблюдается вектор $Y_n = (y_{1-g}, \dots, y_n)$, где

$$y_i = u_i + z_i^\gamma \xi_i, \quad i \in \mathbb{Z}. \quad (7.6.1)$$

В (7.6.1) величины $\{u_i\}$ удовлетворяют (7.1.1); $\{\xi_i\}$ — н.о.р.сл.в. с распределением μ_ξ из некоторого класса допустимых распределений \mathcal{M}_ξ ; $z_i^\gamma \sim \text{Bi}(1, \gamma)$, $0 \leq \gamma \leq 1$, γ — уровень засорения, $\{z_i^\gamma\}$ — н.о.р.сл.в.; последовательности $\{u_i\}$, $\{z_i^\gamma\}$, $\{\xi_i\}$ независимы между собой. Мы продолжаем рассматривать, таким образом, простейшую схему засорения данных независимыми (одиночными) выбросами, ранее это делалось в § 5.6 и § 6.7.

Определения функционалов влияния и чувствительности к большим ошибкам для оценок $\hat{\beta}_n$ векторного параметра β принципиально ничем не отличаются от определений из § 5.6 для скалярного случая. А именно, предположим, что существует предел

$$\hat{\beta}_n \xrightarrow{P} \theta_\gamma$$

и $\theta_0 = \beta$.

Назовем *функционалом влияния оценки* $\hat{\beta}_n$ вектор

$$IF(\theta_\gamma, \mu_\xi) = \lim_{\gamma \rightarrow 0} \frac{\theta_\gamma - \theta_0}{\gamma} \quad (7.6.2)$$

(разумеется, если этот предел существует). Функционал влияния характеризует величину главного линейного члена в разложении асимптотического смещения

$$\theta_\gamma - \theta_0 = IF(\theta_\gamma, \mu_\xi)\gamma + o(\gamma).$$

Предпочтительнее оценки, для которых величина

$$GES(\mathcal{M}_\xi, \theta_\gamma) := \sup_{\mu_\xi \in \mathcal{M}_\xi} |IF(\theta_\gamma, \mu_\xi)|$$

— чувствительность к большим ошибкам — конечна. В этом случае главный член разложения для асимптотического смещения равномерно мал при всех возможных засорениях и малых γ . Качественно последнее означает, что оценка при малых γ будет слабо зависеть даже от очень больших выбросов в данных.

Перейдем к примерам.

7.6.1. Функционал влияния оценки наименьших квадратов $\hat{\beta}_{n,LS}$.

Предположения. $E\varepsilon_1 = 0$, $E\varepsilon_1^2 < \infty$, $E\xi_1^2 < \infty$.

Оценка наименьших квадратов $\hat{\beta}_{n,LS}$ в случае засоренной выборки определяется как решение экстремальной задачи

$$\sum_{i=1}^n (y_i - \theta^T \tilde{y}_{i-1})^2 \Rightarrow \inf_{\theta \in \mathbb{R}^q},$$

где $\tilde{y}_{i-1} = (y_{i-1}, \dots, y_{i-q})^T$. Последняя задача эквивалентна уравнению

$$\sum_{i=1}^n \psi_i^{LS}(Y_n, \theta) = 0, \quad (7.6.3)$$

где

$$\psi_i^{LS}(Y_n, \theta) = (y_{i-1}(y_i - \theta^T \tilde{y}_{i-1}), \dots, y_{i-q}(y_i - \theta^T \tilde{y}_{i-q}))^T.$$

В силу эргодической теоремы для $0 \leq \gamma \leq 1$ и $\theta \in \mathbb{R}^q$

$$n^{-1} \sum_{i=1}^n \psi_i^{LS}(Y_n, \theta) \xrightarrow{P_\theta} \Lambda_{LS}(\gamma, \theta),$$

где

$$\begin{aligned} \Lambda_{LS}(\gamma, \theta) &= (\Lambda_1^{LS}(\gamma, \theta), \dots, \Lambda_q^{LS}(\gamma, \theta))^T, \\ \Lambda_j^{LS}(\gamma, \theta) &= E_{\beta} y_0 (y_j - \theta^T \tilde{y}_{j-1}), \quad j = 1, \dots, q. \end{aligned} \quad (7.6.4)$$

Легко видеть, $\Lambda_{LS}(0, \beta) = 0$. Заменяя в (7.6.4) y_j его выражением из (7.6.1) и пользуясь независимостью последовательностей $\{u_i\}$, $\{z_i^j\}$, $\{\xi_i\}$, а также величин $\{z_i^j\}$ и $\{\xi_i\}$, легко находим, что $\frac{\partial}{\partial \gamma} \Lambda_{LS}(\gamma, \theta)$ и $\frac{\partial}{\partial \theta} \Lambda_{LS}(\gamma, \theta)$ существуют и непрерывны по (γ, θ) при $0 \leq \gamma \leq 1$, $\theta \in \mathbb{R}^q$;

$$\frac{\partial}{\partial \gamma} \Lambda_{LS}(0, \beta) = -E \xi_1^2 \beta, \quad (7.6.5)$$

$$\frac{\partial}{\partial \theta} \Lambda_{LS}(0, \beta) = -(E_{\beta} u_0 u_{|i-j|})_{i,j=1,\dots,q} = -E \varepsilon_1^2 K(\beta). \quad (7.6.6)$$

Поскольку последняя матрица невырождена, в силу теоремы о неявной функции уравнение

$$\Lambda_{LS}(\gamma, \theta) = 0$$

в некоторой окрестности точки $(0, \beta)$ имеет единственное и дифференцируемое по γ решение θ_{γ}^{LS} , $\theta_0^{LS} = \beta$, и в силу (7.6.5)–(7.6.6)

$$\left. \frac{d\theta_{\gamma}^{LS}}{d\gamma} \right|_{\gamma=0} = -\frac{E \xi_1^2}{E \varepsilon_1^2} K^{-1}(\beta) \beta. \quad (7.6.7)$$

При малых γ корень θ_γ^{LS} единственный во всем пространстве \mathbb{R}^q , он выписывается явно. Решение уравнения (7.6.3) (т.е. оценка наименьших квадратов $\hat{\beta}_{n,LS}$) при малых γ с вероятностью, стремящейся к единице, существует и единственно, причем также выписывается явно. Из этих явных соотношений легко следует, что

$$\hat{\beta}_{n,LS} \xrightarrow{P_s} \theta_\gamma^{LS}.$$

Последнее равенство совместно с (7.6.7) и определением (7.6.2) влечет

$$IF(\theta_\gamma^{LS}, \mu_\xi) = -(\mathbf{E}\xi_1^2/\mathbf{E}\varepsilon_1^2)\mathbf{K}^{-1}(\beta)\beta.$$

При $q = 1$ последнее соотношение совпадает с (5.6.9). Если \mathcal{M}_i , $i = 1, 2$, — класс распределений μ_ξ с конечным i -м моментом, то

$$GES(\mathcal{M}_2, \theta_\gamma) = \infty, \quad \beta \neq 0.$$

7.6.2. Функционал влияния оценки наименьших модулей $\hat{\beta}_{n,LD}$.

Предположения. $\mathbf{E}\varepsilon_1 = 0$, $\mathbf{E}\varepsilon_1^2 < \infty$, $\mathbf{E}\xi_1^2 < \infty$, $g(x)$ непрерывна, $\sup_x g(x) < \infty$, $g(0) > 0$, $G(0) = 1/2$.

Оценка наименьших модулей $\hat{\beta}_{n,LD}$ определяется как решение задачи

$$\sum_{k=1}^n |y_k - \theta^T \tilde{y}_{k-1}| \Rightarrow \inf_{\theta \in \mathbb{R}^q}. \quad (7.6.8)$$

Целевая функция в (7.6.8) выпукла вниз, а потому это решение существует, оно п.н. единственно в силу непрерывности распределения наблюдений. В силу эргодической теоремы

$$n^{-1} \sum_{k=1}^n |y_k - \theta^T \tilde{y}_{k-1}| \xrightarrow{P_s} \mathbf{E}_\beta |y_1 - \theta^T \tilde{y}_0|.$$

Функция $\mathbf{E}_\beta |y_1 - \theta^T \tilde{y}_0|$ есть предел выпуклых функций, а потому выпукла. Это гладкая функция, и ее минимум достигается в точках, удовлетворяющих уравнению

$$\frac{\partial}{\partial \theta} \mathbf{E}_\beta |y_1 - \theta^T \tilde{y}_0| = -\mathbf{E}_\beta \tilde{y}_0 \operatorname{sign}(y_1 - \theta^T \tilde{y}_0) = 0. \quad (7.6.9)$$

Обозначим

$$\Lambda_{LD}(\gamma, \theta) = \mathbf{E}_\beta \tilde{y}_0 \operatorname{sign}(y_1 - \theta^T \tilde{y}_0),$$

тогда $\Lambda_{LD}(0, \beta) = 0$. Перепишем $\Lambda_{LD}(\gamma, \theta)$ по формуле полного математического ожидания в виде

$$\begin{aligned} \Lambda_{LD}(\gamma, \theta) &= E_{\beta} \tilde{y}_0 \left[1 - 2G \left((\theta - \beta)^T \tilde{u}_0 + \sum_{j=1}^q \theta_j z_{1-j}^{\gamma} \xi_{1-j} + z_1^{\gamma} \xi_1 \right) \right] = \\ &= \sum_{i=0}^q E_{\beta} \left[\tilde{y}_0 \left(1 - 2G \left((\theta - \beta)^T \tilde{u}_0 + \sum_{j=1}^q \theta_j z_{1-j}^{\gamma} \xi_{1-j} + z_1^{\gamma} \xi_1 \right) \right) | \mathbb{H}_i \right] P \{ \mathbb{H}_i \}, \end{aligned}$$

где случайные события \mathbb{H}_i состоят в том, что точно i величин из ξ_{-q+1}, \dots, ξ_1 отличны от нуля. В последнем выражении условные средние от γ не зависят, а вероятности представляют собой полиномы от γ . Отсюда легко усмотреть, что существует непрерывная по (γ, θ) при $\theta \in \mathbb{R}^q$ и $0 \leq \gamma \leq 1$ матрица $\frac{\partial \Lambda_{LD}(\gamma, \theta)}{\partial \theta}$ и, в частности,

$$\frac{\partial \Lambda_{LD}(0, \beta)}{\partial \theta} = -2g(0) E \varepsilon_1^2 K(\beta); \quad (7.6.10)$$

существует непрерывная производная $\frac{\partial \Lambda_{LD}(\gamma, \theta)}{\partial \gamma}$ и, в частности,

$$\Delta_{LD}(\beta) := \frac{\partial \Lambda_{LD}(0, \beta)}{\partial \gamma} = (\Delta_1^{LD}(\beta), \dots, \Delta_q^{LD}(\beta))^T, \quad (7.6.11)$$

где $\Delta_j^{LD}(\beta) = E_{\beta} \xi_1 (1 - 2G(\beta_j \xi_1))$, $j = 1, \dots, q$.

Следовательно, в некоторой окрестности точки $(0, \beta)$ уравнение (7.6.9) имеет единственное и дифференцируемое решение θ_{γ}^{LD} , $\theta_0^{LD} = \beta$, и в силу (7.6.10)–(7.6.11)

$$\left. \frac{d\theta_{\gamma}^{LD}}{d\gamma} \right|_{\gamma=0} = [2g(0) E \varepsilon_1^2 K(\beta)]^{-1} \Delta_{LD}(\beta). \quad (7.6.12)$$

Поскольку матрица $-\lambda_{LD}(\beta)$ положительно определена, в некоторой окрестности $(0, \beta)$ положительно определенной является матрица $-\frac{\partial \Lambda_{LD}(\gamma, \theta)}{\partial \theta}$. Следовательно, в точке θ_{γ}^{LD} функция $E_{\beta} |y_1 - \theta^T \tilde{y}_0|$ имеет строгий минимум и, более того, у выпуклой функции он только один. Этого достаточно, чтобы, повторяя доказательство теоремы 2.2.1 в [51], установить сходимость $\hat{\beta}_{n,LD}$ к θ_{γ}^{LD} по вероятности (при малых γ).

В силу (7.6.12) и (7.6.2) получаем

$$IF(\theta_{\gamma}^{LD}, \mu_{\varepsilon}) = [2g(0) E \varepsilon_1^2 K(\beta)]^{-1} \Delta_{LD}(\beta). \quad (7.6.13)$$

При $q = 1$ последнее соотношение совпадает с (5.6.12).

Если $\xi_i = \xi$ и ξ — постоянная, то при $\beta \neq 0$ $|IF(\theta_\gamma^{LD}, \mu_\xi)| \rightarrow \infty$, когда $\xi \rightarrow \infty$. Следовательно,

$$GES(\mathcal{M}_2, \theta_\gamma) = \infty, \quad \beta \neq 0.$$

7.6.3. Функционал влияния взвешенных оценок наименьших модулей $\hat{\beta}_{n,LDW}$.

Предположения.

$$\mathbb{E}\varepsilon_1 = 0, \quad \mathbb{E}|\varepsilon_1| < \infty, \quad \mathbb{E}|\xi_1| < \infty;$$

$$\mathbb{E}_\beta |\varphi(\tilde{y}_1)| < \infty, \quad \mathbb{E}_\beta |\varphi(\tilde{y}_1)\tilde{y}_1\tilde{y}_1^T| < \infty \text{ при малых } \gamma;$$

$$-\infty < \lambda_{LDW}(\beta) := -2g(0)\mathbb{E}|\varphi(\tilde{u}_1)|\tilde{u}_1\tilde{u}_1^T < 0;$$

$$g(x) \text{ непрерывна, } \sup_x g(x) < \infty, \quad G(0) = 1/2, \quad g(0) > 0.$$

Аналогично оценке наименьших модулей показывается, что: любая последовательность случайных величин $\hat{\beta}_{n,LDW}$, являющихся при каждом n решением задачи

$$\sum_{k=1}^n |\varphi(\tilde{y}_{k-1})(y_k - \theta^T \tilde{y}_{k-1})| \Rightarrow \inf_{\theta \in \mathbb{R}^q}, \quad (7.6.14)$$

при малых γ сходится по вероятности к θ_γ^{LDW} ; функционал влияния $\hat{\beta}_{n,LDW}$ равен

$$IF(\theta_\gamma^{LDW}, \mu_\xi) = -\lambda_{LDW}^{-1}(\beta)\Delta_{LDW}(\beta), \quad (7.6.15)$$

где

$$\Delta_{LDW}(\beta) = (\Delta_1^{LDW}(\beta), \dots, \Delta_q^{LDW}(\beta))^T,$$

$$\Delta_j^{LDW}(\beta) = \mathbb{E}_\beta \left\{ u_{1-j} |\varphi(\tilde{u}_0)| [1 - 2G(-\xi_1)] + \right. \\ \left. + \sum_{k=1}^q (u_{1-j} + \xi_1 \delta_{kj}) |\varphi(\tilde{u}_{k,\xi})| [1 - 2G(\beta_k \xi_1)] \right\},$$

$$\tilde{u}_{k,\xi} = (u_0, \dots, u_{2-k}, u_{1-k} + \xi_1, u_{-k}, \dots, u_{-q})^T,$$

δ_{ij} — символ Кронекера.

Соотношение (7.6.15) совместно с условием

$$\sup_{\alpha \in \mathbb{R}^q} |\alpha \varphi(\alpha)| < \infty \quad (7.6.16)$$

влечет неравенство

$$GES(\mathcal{M}_\xi, \theta_\gamma^{LDW}) < \infty.$$

Здесь \mathcal{M}_ξ — класс распределений μ_ξ , удовлетворяющих условиям $E|\xi_1| < \infty$ и $E_\beta|\varphi(\tilde{y}_1)| < \infty$, $E_\beta|\varphi(\tilde{y}_1)\tilde{y}_1\tilde{y}_1^T| < \infty$.

Стоит отметить, что при $q = 1$ мы вместо экстремальной задачи (7.6.14) рассматривали в п. 5.6.3 эквивалентное ей уравнение (5.6.14). Это позволило найти функционал влияния $\hat{\beta}_{n,LDW}$ при условиях 5.6.5, 5.6.6, которые слабее сформулированных здесь. Разумеется, при $q = 1$ из (7.6.15) следуют соотношения п. 5.6.3, а при $\varphi(\tilde{y}_1) = 1$ (7.6.15) совпадает с (7.6.13).

7.6.4. Функционал влияния знаковой оценки $\tilde{\beta}_{n,S}$. Для незасоренных данных оценка $\tilde{\beta}_{n,S}$ была определена в (7.5.8). В случае схемы (7.6.1) $\tilde{\beta}_{n,S}$ определяется как измеримое решение задачи

$$|l_n^S(\hat{\beta}_n, \theta)| \Rightarrow \inf_{\theta \in Q(\hat{\beta}_n)},$$

где

$$Q(\hat{\beta}_n) = \{\theta: |\theta - \hat{\beta}_n| \leq n^{-1/2} \log n\}$$

и предварительная оценка $\hat{\beta}_n$ так же, как и $l_n^S(\hat{\beta}_n, \theta)$, основаны на Y_n .

Предположения. Предварительная оценка сходится: $\hat{\beta}_n \xrightarrow{P_\beta} \theta_\gamma$, и существует функционал влияния $IF(\theta_\gamma, \mu_\xi)$.

Тогда из определения оценки $\tilde{\beta}_{n,S}$ непосредственно следует

$$\tilde{\beta}_{n,S} \xrightarrow{P_\beta} \tilde{\theta}_\gamma^S = \theta_\gamma,$$

и, значит, $IF(\tilde{\theta}_\gamma^S, \mu_\xi) = IF(\theta_\gamma, \mu_\xi)$. Таким образом, знаковая оценка $\tilde{\beta}_{n,S}$ наследует функционал влияния предварительной оценки $\hat{\beta}_n$. В качестве последней можно взять какую-нибудь взвешенную оценку наименьших модулей, удовлетворяющую условию (7.6.16); тогда чувствительность $\tilde{\beta}_{n,S}$ будет конечной.

7.6.5. Функционал влияния знаковой оценки $\beta_{n,S}^*$. В силу (7.5.14) эта оценка в случае засоренной выборки определяется соотношением

$$\beta_{n,S}^* = \hat{\beta}_n - (n\hat{e}_n)^{-1}l_n^S(\hat{\beta}_n),$$

где предварительная оценка $\hat{\beta}_n$, а также \hat{e}_n и l_n^S строятся по Y_n . Пусть справедливы следующие

Предположения.

$$\mathbf{E}\varepsilon_1 = 0, \quad \mathbf{E}|\xi_1| < \infty;$$

$$G(0) = 1/2, \quad \sup_x g(x) < \infty, \quad g(0) > 0, \quad g(x) \text{ непрерывна};$$

предварительная оценка $\hat{\beta}_n$ имеет функционал влияния,

$$\hat{\varepsilon}_n \xrightarrow{P_S} \theta_\gamma^e, \quad \theta_0^e = \lambda_S(\beta) \text{ и функция } \theta_\gamma^e \text{ непрерывна в точке } \gamma = 0.$$

Если $\sqrt{n}(\hat{\beta}_n - \theta_\gamma) = O_p(1)$, то последнему требованию удовлетворяет оценка

$$\hat{\varepsilon}_n = -2\hat{g}_n \hat{m}_n \mathbf{K}(\hat{\beta}_n),$$

где

$$\hat{g}_n = (nh_n)^{-1} \sum_{i=1}^n \varphi(h_n^{-1}(y_i - \hat{\beta}_n^T \tilde{y}_{i-1}))$$

— оценка плотности типа Парзена—Розенблатта с гауссовским ядром $\varphi(x)$, $h_n = n^{-\alpha}$, $0 < \alpha < 1/4$;

$$\hat{m}_n = n^{-1} \sum_{i=1}^n |y_i - \hat{\beta}_n^T \tilde{y}_{i-1}|.$$

При малых γ

$$\beta_{n,S}^* \xrightarrow{P_S} \theta_\gamma^{*S} := \theta_\gamma - (\theta_\gamma^e)^{-1} \Lambda_S(\gamma, \theta_\gamma), \quad (7.6.17)$$

где

$$\Lambda_S(\gamma, \theta) = (\Lambda_1^S(\gamma, \theta), \dots, \Lambda_q^S(\gamma, \theta))^T,$$

$$\Lambda_j^S(\gamma, \theta) = \sum_{i=0}^{\infty} \delta_i(\theta) \mathbf{E}_\beta \text{sign} [(y_0 - \theta^T \tilde{y}_{-1})(y_{i+j} - \theta^T \tilde{y}_{i+j-1})], \quad (7.6.18)$$

$$j = 1, \dots, q.$$

Функция $\Lambda_S(\gamma, \theta)$ определена при любом $0 \leq \gamma \leq 1$ и θ из некоторой окрестности $|\theta - \beta| \leq \varepsilon$, так как при достаточно малом ε величины $\delta_i(\theta)$ экспоненциально убывают к нулю вполне аналогично величинам $\delta_i(\beta)$ (см. неравенства (7.1.4) и (7.3.3)). Доказательство соотношения (7.6.17) при произвольном q принципиально ничем не отличается от доказательства аналогичного соотношения (6.7.19) для $q = 1$, которое мы привели со всеми подробностями в п. 6.7.3.

Из (7.6.18) и (7.6.17) следует, что

$$\Lambda_S(0, \beta) = 0, \quad \theta_0^{*S} = \beta. \quad (7.6.19)$$

Аналогично доказательству леммы 6.7.1 устанавливается, что при сделанных предположениях частные производные от $\Lambda_S(\gamma, \theta)$ существуют и непрерывны по (γ, θ) при $|\theta - \beta| \leq \varepsilon$, $0 \leq \gamma \leq 1$. В частности,

$$\frac{\partial \Lambda_S(0, \beta)}{\partial \theta} = \lambda_S(\beta), \quad (7.6.20)$$

$$\frac{\partial \Lambda_S(0, \beta)}{\partial \gamma} = \Delta_S(\beta), \quad (7.6.21)$$

где

$$\begin{aligned} \Delta_S(\beta) &= (\Delta_1^S(\beta), \dots, \Delta_q^S(\beta))^T, \\ \Delta_j^S(\beta) &= \sum_{i=0}^{q-j} \delta_i(\beta) \sum_{k=0}^{q-j-i} E_{\beta}(1 - 2G(\beta_k \xi_1))(1 - 2G(\beta_{k+j+i} \xi_1)), \\ \beta_0 &= 1, \quad j = 1, \dots, q. \end{aligned}$$

Теперь (7.6.17), непрерывная дифференцируемость $\Lambda(\gamma, \theta)$ и непрерывность θ_{γ}^e в нуле влекут

$$\begin{aligned} \theta_{\gamma}^{*S} &= \theta_{\gamma} - (\lambda(\beta) + o(1))^{-1} \left(\Lambda_S(0, \beta) + \right. \\ &\quad \left. + \frac{\partial \Lambda_S(0, \beta)}{\partial \gamma} \gamma + \frac{\partial \Lambda_S(0, \beta)}{\partial \theta} (\theta_{\gamma} - \beta) + o(\gamma) \right). \end{aligned}$$

Последнее соотношение, (7.6.19)–(7.6.21) и существование функционала влияния у предварительной оценки влекут

$$IF(\theta_{\gamma}^{*S}, \mu_{\xi}) := \lim_{\gamma \rightarrow 0} \frac{\theta_{\gamma}^{*S} - \beta}{\gamma} = \lambda_S^{-1}(\beta) \Delta_S(\beta). \quad (7.6.22)$$

При $q = 1$ (7.6.22) совпадает с функционалом влияния непараметрической знаковой оценки $\hat{\beta}_{n,S}$ (6.7.13). Легко видеть, $GES(\mathcal{M}_1, \theta_{\gamma}^{*S}) < \infty$.

Результаты этого параграфа были опубликованы в [54].

§ 7.7. Эмпирическая функция распределения и эмпирические процессы, построенные по остаткам

Рассмотрим авторегрессионную схему (7.1.1) в предположении, что выполнены условия 6.5.2, 6.5.3. Ради удобства напомним их здесь:

$E\varepsilon_1 = 0$, $M|\varepsilon_1|^{1+\Delta} < \infty$ при некотором $\Delta > 0$;

$G(x)$ имеет плотность $g(x)$ и $\sup_x g(x) < \infty$, $g(x)$ удовлетворяет условию Липшица.

Пусть u_{1-q}, \dots, u_n — наблюдения из строго стационарного решения (7.1.5) уравнения (7.1.1).

Одной из задач настоящего параграфа будет построение по этим наблюдениям тестов типа Колмогорова и омега-квадрат для проверки гипотез о виде функции распределения $G(x)$. Мы будем рассматривать различные гипотезы, в частности,

$$H_0: G(x) = G_0(x), \quad G_0(x) \text{ полностью известна}; \quad (7.7.1)$$

а также гипотезу о принадлежности $G(x)$ параметрическому семейству

$$H_G: G(x) \in \{G(x, \lambda), \lambda = (\lambda_1, \dots, \lambda_r)^T \in \Lambda\}, \quad (7.7.2)$$

Λ — область в \mathbb{R}^r .

Наши тесты будут основываться на эмпирической функции распределения $\hat{G}_n(x)$, построенной по оценкам ненаблюдаемых случайных величин $\varepsilon_1, \dots, \varepsilon_n$.

Исследование асимптотических свойств функции $\hat{G}_n(x)$ составляет вторую и основную задачу настоящего параграфа.

Итак, пусть для $\theta = (\theta_1, \dots, \theta_q)^T \in \mathbb{R}^q$

$$\varepsilon_k(\theta) := u_k - \theta_1 u_{k-1} - \dots - \theta_q u_{k-q}, \quad k = 1, \dots, n,$$

$$G_n(x, \theta) := n^{-1} \sum_{k=1}^n I(\varepsilon_k(\theta) < x).$$

Статистика $G_n(x, \theta)$ есть подобие эмпирической функции распределения, построенной по зависимым при $\beta \neq \theta$ величинам $\varepsilon_1(\theta), \dots, \varepsilon_n(\theta)$. Поскольку $\varepsilon_k(\beta) = \varepsilon_k$, то $G_n(x, \beta)$ есть истинная эмпирическая функция распределения, построенная по $\varepsilon_1, \dots, \varepsilon_n$. Вектор β неизвестен и построить $G_n(x, \beta)$ можно только гипотетически. Мы установим равномерную по x близость $G_n(x, \beta)$ и $G_n(x, \theta)$ при $|\theta - \beta| \leq n^{-1/2} \Theta$, $0 < \Theta < \infty$. Отсюда, пользуясь хорошо известными свойствами $G_n(x, \beta)$ (о которых скажем дальше), мы получим утверждения относительно $\hat{G}_n(x) := G_n(x, \hat{\beta}_n)$, где $\hat{\beta}_n$ — \sqrt{n} -состоятельная оценка β .

Основным результатом параграфа является

Теорема 7.7.1. *Предположим, что выполнены условия 6.5.2, 6.5.3.*

Тогда при любом $0 < \Theta < \infty$

$$\sup_x \sup_{|\theta| \leq \Theta} \sqrt{n} |G_n(x, \beta + n^{-1/2} \theta) - G_n(x, \beta)| = o_p(1).$$

Теорема 7.7.1 доказана в приложении, § 7.8. Для случая конечной дисперсии $E\varepsilon_1^2 < \infty$ она была опубликована в [6]. В [55] теорема 7.7.1 обобщена на случай $|\theta| \leq \Theta n^\alpha$, $\alpha < \min\left(\frac{1}{6}, \frac{\Delta}{2(1+\Delta)}\right)$.

Обозначим ради краткости $G_n(x, \beta) = G_n(x)$.

С л е д с т в и е 7.7.1. Если выполнены условия 6.5.2, 6.5.3 и $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$, то

$$\sup_x \sqrt{n} |\hat{G}_n(x) - G_n(x)| = o_p(1). \quad (7.7.3)$$

Д о к а з а т е л ь с т в о. Для любых $\varepsilon, \delta > 0$ и $0 < \Theta < \infty$

$$\begin{aligned} P_\beta \left\{ \sup_x \sqrt{n} |\hat{G}_n(x) - G_n(x)| > \varepsilon \right\} &\leq \\ &\leq P_\beta \left\{ \sup_x \sqrt{n} |\hat{G}_n(x) - G_n(x)| > \varepsilon, |\sqrt{n}(\hat{\beta}_n - \beta)| \leq \Theta \right\} + \\ &\quad + P_\beta \{ |\sqrt{n}(\hat{\beta}_n - \beta)| > \Theta \} \leq \\ &\leq P_\beta \left\{ \sup_x \sup_{|\theta| \leq \Theta} \sqrt{n} |G_n(x, \beta + n^{-1/2}\theta) - G_n(x, \beta)| > \varepsilon \right\} + \\ &\quad + P_\beta \{ |\sqrt{n}(\hat{\beta}_n - \beta)| > \Theta \} < \delta \end{aligned}$$

при всех достаточно больших n и Θ в силу теоремы 7.7.1 и условия $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$. \square

Следствие 7.7.1 открывает возможность для асимптотических проверок гипотез о функции распределения случайных ошибок в схеме авторегрессии подобно тому, как это делается в схеме повторных независимых наблюдений. Начнем с проверки гипотезы H_0 из (7.7.1). Пусть $\mathbb{D}[0, 1]$, как и в § 5.7, будет метрическое пространство Скорохода функций без разрывов 2-го рода, определенных на отрезке $[0, 1]$, а $w(t)$, $t \in [0, 1]$, — броуновский мост. Обозначим $G^{-1}(t)$ обратную к $G(x)$ функцию,

$$G^{-1}(t) := \sup\{x: G(x) \leq t\}.$$

Введем случайные процессы

$$\begin{aligned} \hat{w}_n(t) &= \sqrt{n} [\hat{G}_n(G^{-1}(t)) - t], \\ w_n(t) &= \sqrt{n} [G_n(G^{-1}(t)) - t], \\ t &\in [0, 1]. \end{aligned}$$

В силу (7.7.3)

$$\sup_t |\hat{w}_n(t) - w_n(t)| = o_p(1). \quad (7.7.4)$$

Теперь (7.7.4) и сходимость $w_n(t)$ к $w(t)$ в $\mathbb{D}[0, 1]$ влекут

С л е д с т в и е 7.7.2. Пусть выполнены условия 6.5.2, 6.5.3, $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$.

Тогда $\hat{w}_n(t)$ слабо сходится к $w(t)$ в $\mathbb{D}[0, 1]$.

Введем статистики

$$\hat{D}_n = \sup_x |\sqrt{n}[\hat{G}_n(x) - G_0(x)]|, \quad \hat{w}_n^2 = \int_{-\infty}^{\infty} [\sqrt{n}(\hat{G}_n(x) - G_0(x))]^2 dG_0(x).$$

В силу следствия 7.7.2 при гипотезе H_0 из (7.7.1)

$$P_{\beta}\{\hat{D}_n < \lambda\} = P_{\beta}\left\{\sup_t |\hat{w}_n(t)| < \lambda\right\} \rightarrow P\left\{\sup_t |w(t)| < \lambda\right\} = K(\lambda),$$

$$P_{\beta}\{\hat{w}_n^2 < \lambda\} = P_{\beta}\left\{\int_0^1 \hat{w}_n^2(t) dt < \lambda\right\} \rightarrow P\left\{\int_0^1 w^2(t) dt < \lambda\right\} = S(\lambda).$$

Здесь $K(\lambda)$ и $S(\lambda)$ — известные и табулированные функции распределения Колмогорова и Смирнова (см., например, [10, гл. 6]).

Итак, асимптотические распределения статистик \hat{D}_n и \hat{w}_n^2 при H_0 свободны и известны, поэтому эти статистики можно использовать для проверки гипотез относительно $G(x)$.

Отметим еще, что если $k_{1-\alpha}$ — квантиль $K(\lambda)$ уровня $1 - \alpha$, то в силу следствия 7.7.1

$$P\left\{\sup_x \sqrt{n}|\hat{G}_n(x) - G(x)| < k_{1-\alpha}\right\} \rightarrow 1 - \alpha.$$

Отсюда следует, что асимптотический равномерный по x доверительный интервал для $G(x)$ надежности $1 - \alpha$ имеет вид

$$\hat{G}_n(x) - n^{-1/2}k_{1-\alpha} < G(x) < \hat{G}_n(x) + n^{-1/2}k_{1-\alpha}.$$

Рассмотрим теперь проверку параметрической гипотезы H_G из (7.7.2). В схеме повторной выборки проверка таких гипотез была объектом многих исследований, так что трудно перечислить всех, кто внес вклад в эту область. Да мы и не ставим такой задачи. Ограничимся несколькими избранными результатами. Итак, обозначим

$$f(x, \lambda) = \left(\frac{\partial G(x, \lambda)}{\partial \lambda_1}, \dots, \frac{\partial G(x, \lambda)}{\partial \lambda_r}\right)^T.$$

Пусть в предположении, что H_G верна, $\hat{\lambda}_n$ — оценка истинного значения параметра λ_0 , при котором $G(x) = G(x, \lambda_0)$, а $g(x, \lambda)$ — плот-

ность $G(x, \lambda)$. Пусть при H_G из (7.7.2) выполнены следующие предположения.

Условие 7.7.1. При любом $\lambda \in \Lambda$ $\sup_x g(x, \lambda) < \infty$, $g(x, \lambda)$ удовлетворяет условию Липшица по переменной x , $\int_{-\infty}^{\infty} xg(x, \lambda)dx = 0$, $\int_{-\infty}^{\infty} |x|^{1+\Delta}g(x, \lambda)dx < \infty$ при некотором $\Delta > 0$.

Условие 7.7.2. Функция $f(G^{-1}(t, \lambda), \lambda)$ непрерывна по (t, λ) при $t \in [0, 1]$, $\lambda \in \Lambda$.

Условие 7.7.3. Справедливо разложение

$$\sqrt{n}(\hat{\lambda}_n - \lambda_0) = n^{-1/2} \sum_{k=1}^n l(\varepsilon_k, \lambda_0) + o_p(1), \quad (7.7.5)$$

где

$$\int_{-\infty}^{\infty} l(x, \lambda)g(x, \lambda)dx = 0, \quad L(\lambda) := \int_{-\infty}^{\infty} l(x, \lambda)l^T(x, \lambda)g(x, \lambda)dx < \infty.$$

Типичная ситуация, когда верно условие 7.7.3, следующая. Пусть $\tilde{\lambda}_n = \tilde{\lambda}_n(\varepsilon_1, \dots, \varepsilon_n)$ — какая-нибудь «хорошая» оценка λ_0 , построенная (гипотетически) по $\varepsilon_1, \dots, \varepsilon_n$. Тогда обычно

$$\sqrt{n}(\tilde{\lambda}_n - \lambda_0) = n^{-1/2} \sum_{k=1}^n l(\varepsilon_k, \lambda_0) + o_p(1), \quad (7.7.6)$$

причем слагаемые в (7.7.6) имеют среднее нуль и конечную ковариационную матрицу. Например, если $\tilde{\lambda}_n$ есть оценка максимального правдоподобия, то при стандартных условиях регулярности

$$l(x, \lambda) = I^{-1}(\lambda) \frac{\partial \log g(x, \lambda)}{\partial \lambda}, \quad (7.7.7)$$

где $I(\lambda)$ — информация Фишера о параметре λ , содержащаяся в $\varepsilon_1, \dots, \varepsilon_n$. В качестве оценки $\tilde{\lambda}_n$ естественно брать $\hat{\lambda}_n := \tilde{\lambda}_n(\varepsilon_1(\hat{\beta}_n), \dots, \varepsilon_n(\hat{\beta}_n))$, т.е. обращаться с остатками $\varepsilon_1(\hat{\beta}_n), \dots, \varepsilon_n(\hat{\beta}_n)$ как с самими ненаблюдаемыми $\varepsilon_1, \dots, \varepsilon_n$. Если

$$\sqrt{n} [\tilde{\lambda}_n(\varepsilon_1, \dots, \varepsilon_n) - \tilde{\lambda}_n(\varepsilon_1(\hat{\beta}_n), \dots, \varepsilon_n(\hat{\beta}_n))] = o_p(1), \quad (7.7.8)$$

то (7.7.6) и (7.7.8) влекут (7.7.5). Мы приведем далее пример такой оценки.

Обозначим ради краткости $L = L(\lambda_0)$, $f(t) = f(G^{-1}(t), \lambda_0)$, $l(t) = l(G^{-1}(t), \lambda_0)$, $e(t) = \int_0^t l(s) ds$ и введем случайный процесс

$$\hat{u}_n[G(x, \hat{\lambda}_n)] = \sqrt{n} [\hat{G}_n(x) - G(x, \hat{\lambda}_n)].$$

После замены $G(x, \hat{\lambda}_n) = t$ получим процесс $\hat{u}_n(t)$, определенный на отрезке $[0, 1]$.

С л е д с т в и е 7.7.3. Пусть выполнены условия 7.7.1–7.7.3 и $\sqrt{n}(\hat{\beta}_n - \beta) = O_p(1)$.

Тогда при гипотезе H_G $\hat{u}_n(t)$ слабо сходится в $\mathbb{D}[0, 1]$ к гауссовскому процессу $u(t)$ со средним нуль и ковариацией

$$\min(t, s) - ts - f^T(t)e(s) - f^T(s)e(t) + f^T(t)Lf(s).$$

Д о к а з а т е л ь с т в о. В силу (7.7.3)

$$\sup_x |\hat{u}_n[G(x, \hat{\lambda}_n)] - \sqrt{n}[G_n(x) - G(x, \hat{\lambda}_n)]| = o_p(1). \quad (7.7.9)$$

Обозначим

$$u_n[G(x, \hat{\lambda}_n)] = \sqrt{n}[G_n(x) - G(x, \hat{\lambda}_n)].$$

Тогда $u_n(t)$ слабо сходится в $\mathbb{D}[0, 1]$ к $u(t)$ в силу [62, теорема 1]. (В частном случае (7.7.7) сходимость $u_n(t)$ в $\mathbb{D}[0, 1]$ к $u(t)$ была установлена в [29].) Этот факт и соотношение (7.7.9) влекут утверждение следствия 7.7.3. \square

Рассмотрим, в частности, гипотезу

$$H_\Phi: G(x) \in \{\Phi(x/\sigma), 0 < \sigma < \infty\}$$

о нормальности $G(x)$ (которая, кстати, эквивалентна нормальности процесса $\{u_i\}$ из (7.1.5)).

Оценкой дисперсии σ_0^2 возьмем

$$\hat{s}_n^2 = n^{-1} \sum_{k=1}^n \varepsilon_k^2(\hat{\beta}_n).$$

Тогда, легко проверить, при H_Φ выполнены (7.7.6) и (7.7.8) с $l(x, \sigma) = x^2 - \sigma^2$ и, следовательно, выполнено условие 7.7.3. Очевидно, условия 7.7.1, 7.7.2 также выполнены. Обозначим

$$\hat{u}_{1n}(t) = \sqrt{n} [\hat{G}_n(\hat{s}_n \Phi^{-1}(t)) - t].$$

В силу следствия 7.7.3 с помощью простых вычислений получаем, что при H_Φ $\hat{u}_{1n}(t)$ слабо сходится в $\mathbb{D}[0, 1]$ к гауссовскому процессу $u_1(t)$ со средним нуль и ковариацией

$$\min(t, s) - ts - \frac{1}{2} \Phi^{-1}(t) \varphi(\Phi^{-1}(t)) \Phi^{-1}(s) \varphi(\Phi^{-1}(s)),$$

где $\varphi(x) = \Phi'(x)$. Распределение $\int_0^1 u_1^2(t) dt$ табулировано, например, в [21, 22] или [50], и потому H_Φ можно проверять критерием типа омега-квадрат со статистикой

$$\hat{w}_{1n}^2 := \int_{-\infty}^{\infty} \hat{u}_{1n}^2[\Phi(x/\hat{s}_n)] d\Phi(x/\hat{s}_n) = \int_0^1 \hat{u}_{1n}^2(t) dt.$$

Приведенный пример нуждается в комментарии. Оказывается, можно построить более мощный на естественном классе альтернатив тест. А именно, следует основывать интегральную тестовую статистику на процессе

$$\hat{u}_{2n}(t) := \sqrt{n} [\hat{G}_n(\tilde{\varepsilon}_n + \hat{s}_n \Phi^{-1}(t)) - t].$$

где $\tilde{\varepsilon}_n = n^{-1} \sum_{k=1}^n \varepsilon_k(\hat{\beta}_n)$, т.е. действовать так, как будто среднее значение у ε_i не известно.

Мы отсылаем читателя к работе [28], где подобный феномен отмечен и объяснен для случая повторной выборки. В силу следствия 7.7.1 рассматриваемые тестовые статистики в схеме авторегрессии асимптотически распределены так же, как статистики в схеме повторной выборки. С другой стороны, на практике естественнее сразу предполагать неизвестное среднее у ε_i и включать его в число оцениваемых параметров. (О применении тестов типа Колмогорова и омега-квадрат в схемах авторегрессии с ненулевым средним и в более общих ситуациях см. [7].)

Процесс $u(t)$ зависит от гипотетического семейства и истинного параметра λ_0 (кроме нескольких исключений, например, когда не известны параметры сдвига и масштаба). Поэтому и функционалы типа Колмогорова и омега-квадрат, основанные на $\hat{u}_n(t)$, также асимптотически несвободны при H_G . Пример тому — процесс $u_1(t)$ и статистика \hat{w}_{1n}^2 : распределения процесса $u_1(t)$ зависят от нормального семейства, но не зависят от σ_0 . Можно преобразовать $u(t)$ (и, аналогично, $\hat{u}_n(t)$) к новому процессу, распределения которого свободны. Такие преобразования изучали многие авторы, в частности см. [29, 39, 40]. Тестовые статистики можно основывать на преобразованном процессе $\hat{u}_n(t)$, они будут асимптотически свободны.

В некоторых случаях удается избежать необходимости преобразовывать процесс $\hat{u}_n(t)$. А именно, для статистик Колмогорова—Смирнова, основанных на $\varepsilon_1, \dots, \varepsilon_n$, можно найти постоянные a и b

такие, что при гипотезе H_G

$$\lim_{n \rightarrow \infty} P \left\{ \sup_x |u_n[G(x, \tilde{\lambda}_n)]| > z \right\} = a e^{-bz^2} (1 + o(1)), \quad z \rightarrow \infty.$$

Последнее соотношение позволяет находить квантили статистики $\sup_x |u_n[G(x, \tilde{\lambda}_n)]|$ высокого уровня. Подобные результаты можно найти, например, в [24, 31]. В силу следствия 7.7.3 статистика $\sup_x |\hat{u}_n[G(x, \hat{\lambda}_n)]|$ будет иметь те же асимптотические квантили. Например, при H_Φ

$$\lim_{n \rightarrow \infty} P_\beta \left\{ \sup_x |\hat{u}_{1n}[\Phi(x, \hat{s}_n)]| > z \right\} = \frac{2\sqrt{6}}{3} \exp(-2z^2) (1 + o(1)), \quad z \rightarrow \infty.$$

Уместно отметить, что теоремы типа 7.7.1 (и следствия типа 7.7.1-7.7.3 из них) получены и для схем наблюдений, отличных от рассматриваемых здесь. Выделим, в частности, следующие работы: [7], где рассматривается схема линейной регрессии с автокоррелированными погрешностями; [8], где рассматривается схема скользящего среднего; [88], где рассматривается авторегрессия 1-го порядка с коэффициентом β таким, что $|\beta| > 1$; [93], где рассматриваются схема авторегрессии-скользящего среднего и более общие линейные процессы; [87], где изучаются взвешенные эмпирические процессы.

§ 7.8. Приложение. Доказательство теоремы 7.7.1

Докажем теорему при $q = 1$. Рассуждение в общем случае принципиальных отличий не имеет, но более громоздко. Приводимое здесь доказательство является модификацией доказательства из [6].

Итак, дальше β и θ — скаляры. Обозначим

$$\Delta_k(x) = \begin{cases} 1, & \varepsilon_k < x, \\ 0, & \varepsilon_k \geq x. \end{cases}$$

Тогда $G_n(x, \beta) = n^{-1} \sum_{k=1}^n \Delta_k(x)$, а так как $\varepsilon_k(\theta) = \varepsilon_k - (\theta - \beta)u_{k-1}$, то

$$G_n(x, \theta) = n^{-1} \sum_{k=1}^n \Delta_k(x + (\theta - \beta)u_{k-1}).$$

Поэтому

$$\sqrt{n}[G_n(x, \beta + n^{-1/2}\theta) - G_n(x, \beta)] = z_{1n}(x, \theta) + z_{2n}(x, \theta),$$

где

$$z_{1n}(x, \theta) = n^{-1/2} \sum_{k=1}^n [\Delta_k(x + n^{-1/2} \theta u_{k-1}) - G(x + n^{-1/2} \theta u_{k-1}) - \Delta_k(x) + G(x)],$$

$$z_{2n}(x, \theta) = n^{-1/2} \sum_{k=1}^n [G(x + n^{-1/2} \theta u_{k-1}) - G(x)].$$

Плотность вероятности $g(x)$ ограничена и удовлетворяет условию Липшица, а потому удовлетворяет и условию Гельдера любого порядка $0 < \delta \leq 1$ с некоторой постоянной L . Пусть дальше $\delta = \min(\Delta, 1)$.

Применяя формулу Тейлора к $z_{2n}(x, \theta)$, получаем

$$z_{2n}(x, \theta) = \theta g(x) n^{-1} \sum_{k=1}^n u_{k-1} + \varepsilon_n(x, \theta),$$

где

$$\sup_x \sup_{|\theta| \leq \Theta} |\varepsilon_n(x, \theta)| \leq \Theta^{1+\delta} L n^{-(1+\delta/2)} \sum_{k=1}^n |u_{k-1}|^{1+\delta} = o_p(1).$$

Поскольку еще и $n^{-1} \sum_{k=1}^n u_{k-1} = o_p(1)$, то

$$\sup_x \sup_{|\theta| \leq \Theta} |z_{2n}(x, \theta)| = o_p(1). \tag{7.8.1}$$

Докажем теперь, что

$$\sup_x \sup_{|\theta| \leq \Theta} |z_{1n}(x, \theta)| = o_p(1). \tag{7.8.2}$$

Соотношения (7.8.1) и (7.8.2) повлекут утверждение теоремы.

Для доказательства (7.8.2) перейдем от верхней грани по континууму значений x и θ в (7.8.2) к верхней грани по конечным множествам аргументов. Для этого разобьем отрезок $[-\Theta n^{-1/2}, \Theta n^{-1/2}]$ на 3^{m_n} частей точками $\eta_{sn} = -\Theta n^{-1/2} + 2\Theta n^{-1/2} 3^{-m_n} s$, $s = 0, 1, \dots, 3^{m_n}$, и пусть $3^{m_n} \sim \log n$. Разобьем также действительную ось на N_n частей точками $-\infty = x_0 < x_1 < \dots < x_{N_n-1} < x_{N_n} = +\infty$, где $G(x_i) = iN_n^{-1}$, и пусть $N_n \sim \sqrt{n} \log n$. Введем, кроме того, величины

$$\tilde{u}_{ks} = u_k [1 - 2\Theta n^{-1/2} 3^{-m_n} \eta_{sn}^{-1} I(u_k \geq 0)],$$

$$\hat{u}_{ks} = u_k [1 - 2\Theta n^{-1/2} 3^{-m_n} \eta_{sn}^{-1} I(u_k \leq 0)],$$

$$k = 1, \dots, n. \tag{7.8.3}$$

Тогда

$$\max(|\tilde{u}_{ks}|, |\hat{u}_{ks}|) \leq 3|u_{k-1}|. \quad (7.8.4)$$

Если η_{jn} таково, что

$$0 \leq \eta_{jn} - n^{-1/2}\theta \leq 2\Theta n^{-1/2}3^{-m_n},$$

то справедливы неравенства

$$\eta_{jn}\tilde{u}_{k-1,j} \leq \theta n^{-1/2}u_{k-1} \leq \eta_{jn}\hat{u}_{k-1,j}, \quad k = 1, \dots, n.$$

Поэтому для таких x_r и x_{r+1} , что $x_r \leq x \leq x_{r+1}$, имеем

$$x_r + \eta_{jn}\tilde{u}_{k-1,j} \leq x + n^{-1/2}\theta u_{k-1} \leq x_{r+1} + \eta_{jn}\hat{u}_{k-1,j}, \quad (7.8.5)$$

$$k = 1, \dots, n.$$

Обозначим $\hat{U}_s = (\hat{u}_{0s}, \dots, \hat{u}_{n-1,s})$. Пусть

$$z_n(x, \eta_{sn}, \hat{U}_s) = n^{-1/2} \sum_{k=1}^n [\Delta_k(x + \eta_{sn}\hat{u}_{k-1,s}) - G(x + \eta_{sn}\hat{u}_{k-1,s}) - \Delta_k(x) + G(x)]$$

и пусть $\tilde{U}_s, Z_n(x, \eta_{sn}, \tilde{U}_s)$ определены аналогично по \tilde{u}_{ks} . Вследствие монотонности $\Delta_k(x)$ и $G(x)$ и (7.8.5) получаем (см. аналогичные рассуждения при доказательстве леммы 5.8.1)

$$\sup_x \sup_{|\theta| \leq \Theta} |z_{1n}(x, \theta)| \leq$$

$$\leq \sup_{i \leq N_n - 1} \sup_{s \leq 3^{m_n}} |z_n(x_{i+1}, \eta_{sn}, \hat{U}_s)| + \quad (7.8.6)$$

$$+ \sup_{i \leq N_n} \sup_{s \leq 3^{m_n}} |z_n(x_i, \eta_{sn}, \tilde{U}_s)| + \quad (7.8.7)$$

$$+ \sup_{|t_1 - t_2| \leq N_n^{-1}} \left| n^{-1/2} \sum_{k=1}^n [\Delta_k(G^{-1}(t_1)) - t_1 - \Delta_k(G^{-1}(t_2)) + t_2] \right| + \quad (7.8.8)$$

$$+ \sup_{i \leq N_n} \sup_{s \leq 3^{m_n}} n^{-1/2} \sum_{k=1}^n [G(x_{i+1} + \eta_{sn}\hat{u}_{k-1,s}) - G(x_i + \eta_{sn}\tilde{u}_{k-1,s})]. \quad (7.8.9)$$

Покажем, что выражения (7.8.6)–(7.8.9) есть $o_p(1)$. Для (7.8.8) это верно при $N_n^{-1} = o(1)$ в силу доказательства теоремы 13.1 [4, 49].

Рассмотрим (7.8.9). Выражение под знаком супремумов с помощью формулы Тейлора перепишем в виде

$$\sqrt{n}[G(x_{i+1}) - G(x_i)] + g(x_{i+1})\eta_{sn}n^{-1/2} \times \\ \times \sum_{k=1}^n \hat{u}_{k-1,s} - g(x_i)\eta_{sn}n^{-1/2} \sum_{k=1}^n \tilde{u}_{k-1,s} + \varepsilon_n, \quad (7.8.10)$$

где

$$|\varepsilon_n| \leq 2L\Theta^{1+\delta}n^{-(1+\delta/2)} \sum_{k=1}^n |\hat{u}_{k-1,s}|^{1+\delta} = o_p(1)$$

в силу (7.8.4). Третье слагаемое в (7.8.10) не больше

$$\Theta \sup_x g(x) \left[\left| n^{-1} \sum_{k=1}^n u_{k-1} \right| + 2 \cdot 3^{-m_n} n^{-1} \sum_{k=1}^n |u_{k-1}| \right] = o_p(1),$$

поскольку $3^{-m_n} = o(1)$ и $n^{-1} \sum_{k=1}^n u_{k-1} = o_p(1)$. Аналогично и второе

слагаемое в (7.8.10) есть $o_p(1)$, а первое равно $\sqrt{n}N_n^{-1} = o(1)$ в силу выбора N_n . Итак, (7.8.9) есть $o_p(1)$.

Рассмотрим теперь (7.8.7). Здесь рассуждения довольно кропотливы. Прежде всего аппроксимируем величины

$$u_{k-1} = \sum_{j=0}^{\infty} \beta^j \varepsilon_{k-1-j}$$

величинами

$$u_{k-1}^* = \sum_{j=0}^{l_n-1} \beta^j \varepsilon_{k-1-j},$$

зависящими от конечного числа ε_j . Пусть $l_n \sim -r \log_b n$, где $|\beta| < b < 1$, а $r > 2(1+\delta)/\delta$ натуральное. Аналогично $\tilde{u}_{k,s}$ из (7.8.3) введем величины

$$\tilde{u}_{k,s}^* = u_k^* [1 - 2\Theta n^{-1/2} 3^{-m_n} \eta_{sn}^{-1} I(u_k^* \geq 0)].$$

Перепишем теперь $z_n(x_i, \eta_{sn}, \tilde{U}_s)$ тождественно в виде

$$z_n(x_i, \eta_{sn}, \tilde{U}_s) = n^{-1/2} \sum_{k=1}^n \nu_k(i, s) + n^{-1/2} \sum_{k=1}^n \xi_k(i, s),$$

где

$$\begin{aligned} \nu_k(i, s) &= \Delta_k(x_i + \eta_{sn} \tilde{u}_{k-1, s}) - G(x_i + \eta_{sn} \tilde{u}_{k-1, s}) - \\ &\quad - \Delta_k(x_i + \eta_{sn} \tilde{u}_{k-1, s}^*) + G(x_i + \eta_{sn} \tilde{u}_{k-1, s}^*), \\ \xi_k(i, s) &= \Delta_k(x_i + \eta_{sn} \tilde{u}_{k-1, s}^*) - G(x_i + \eta_{sn} \tilde{u}_{k-1, s}^*) - \Delta_k(x_i) + G(x_i). \end{aligned}$$

Тогда выражение в (7.8.7) не больше

$$\sup_{i \leq N_n} \sup_{s \leq 3^{m_n}} \left(\left| n^{-1/2} \sum_{k=1}^n \nu_k(i, s) \right| + \left| n^{-1/2} \sum_{k=1}^n \xi_k(i, s) \right| \right).$$

Поэтому $\forall \varepsilon > 0$ в силу неравенства Чебышева

$$\begin{aligned} \mathbf{P}_\beta \left\{ \sup_{i \leq N_n} \sup_{s \leq 3^{m_n}} (|z_n(x_i, \eta_{sn}, \tilde{U}_s)| > 2\varepsilon) \right\} &\leq \\ &\leq \sum_{i=0}^{N_n} \sum_{s=0}^{3^{m_n}} \left[\mathbf{P}_\beta \left\{ \left| n^{-1/2} \sum_{k=1}^n \nu_k(i, s) \right| > \varepsilon \right\} + \right. \\ &\quad \left. + \mathbf{P}_\beta \left\{ \left| n^{-1/2} \sum_{k=1}^n \xi_k(i, s) \right| > \varepsilon \right\} \right] \leq \\ &\leq \sum_{i=0}^{N_n} \sum_{s=0}^{3^{m_n}} \left[\varepsilon^{-2} \mathbf{E}_\beta \left| n^{-1/2} \sum_{k=1}^n \nu_k(i, s) \right|^2 + \right. \\ &\quad \left. + \varepsilon^{-4} \mathbf{E}_\beta \left| n^{-1/2} \sum_{k=1}^n \xi_k(i, s) \right|^4 \right]. \quad (7.8.11) \end{aligned}$$

Следующие две леммы устанавливают скорости сходимости к нулю средних в (7.8.11).

$$\text{Л е м м а 7.8.1. } \sup_{i \leq N_n} \sup_{s \leq 3^{m_n}} \mathbf{E}_\beta \left| n^{-1/2} \sum_{k=1}^n \nu_k(i, s) \right|^2 = O(n^{-3/2}).$$

Д о к а з а т е л ь с т в о. Прежде всего покажем, что

$$\sup_{k, s} \mathbf{E}_\beta |\tilde{u}_{ks} - \tilde{u}_{ks}^*| = O(n^{-1}). \quad (7.8.12)$$

Имеем в силу определения u_{k-1}^* и выбора l_n

$$\sup_k \mathbf{E}_\beta |u_{ks} - u_{ks}^*| \leq \sup_k \mathbf{E}_\beta \sum_{j \geq l_n} b^j |\varepsilon_{k-j}| = O(b^{l_n}) = O(n^{-r}). \quad (7.8.13)$$

Теперь из определений \tilde{u}_{ks} , \tilde{u}_{ks}^* и неравенства $|2\Theta n^{-1/2} 3^{-m_n} \eta_{sn}^{-1}| \leq 2$

следует

$$\sup_{k,s} \mathbf{E}_\beta |\tilde{u}_{ks} - \tilde{u}_{ks}^*| \leq 3 \mathbf{E}_\beta |u_1 - u_1^*| + \\ + 2 \mathbf{E}_\beta |u_1 [I(u_1^* \geq 0) - I(u_1 \geq 0)]|. \quad (7.8.14)$$

Второе слагаемое в правой части (7.8.14) в силу неравенства Гёльдера не больше

$$2(\mathbf{E}_\beta |u_1|^{1+\delta})^{\frac{1}{1+\delta}} [\mathbf{E}_\beta |I(u_1^* \geq 0) - I(u_1 \geq 0)|^{\frac{1+\delta}{\delta}}]^{\frac{\delta}{1+\delta}} \leq \\ \leq c [\mathbf{E}_\beta |I(u_1^* \geq 0) - I(u_1 \geq 0)|]^{\frac{\delta}{1+\delta}} = \\ = c [P_\beta \{u_1^* \geq 0, u_1 < 0\} + P_\beta \{u_1^* < 0, u_1 \geq 0\}]^{\frac{\delta}{1+\delta}}.$$

Здесь c и далее c_1, c_2, \dots не зависят от i, r, s, n . Теперь в силу неравенства Чебышева

$$P_\beta \{u_1^* \geq 0, u_1 < 0\} \leq P_\beta \{|u_1^* - u_1| \geq n^{-\frac{1+\delta}{\delta}}\} + \\ + P_\beta \{-n^{\frac{1+\delta}{\delta}} < u_1 < 0\} \leq \\ \leq n^{\frac{1+\delta}{\delta}} \mathbf{E}_\beta |u_1^* - u_1| + n^{-\frac{1+\delta}{\delta}} \sup_x g(x) \leq c_1 n^{-\frac{1+\delta}{\delta}}.$$

Оценивая аналогично $P_\beta \{u_1^* < 0, u_1 \geq 0\}$ получаем, что второе слагаемое в (7.8.14) есть $O(n^{-1})$. Отсюда и из (7.8.14), (7.8.13) следует (7.8.12). Осталось заметить, что величины $\{\nu_k(i, s), k = 1, 2, \dots, n\}$ центрированы, некоррелированы и

$$\mathbf{E}_\beta \nu_k^2(i, s) \leq \mathbf{E}_\beta |\eta_{sn} \tilde{u}_{k-1, s} - \eta_{sn} \tilde{u}_{k-1, s}^*| \leq c_2 n^{-3/2}$$

в силу (7.8.12) (см. аналогичное более подробное рассуждение при доказательстве леммы 5.6.1). Отсюда следует утверждение леммы 7.8.1. \square

$$\text{Л е м м а 7.8.2. } \sup_{i,s} \mathbf{E}_\beta \left| n^{-1/2} \sum_{k=1}^n \xi_k(i, s) \right|^4 = O(n^{-1} \log^3 n).$$

Д о к а з а т е л ь с т в о. Величины $\{\xi_k(i, s), k = 1, \dots, n\}$ p -зависимы с $p = l_n \sim -r \log_\delta n$, центрированы, некоррелированы и

$$|\xi_k(i, s)| \leq 2, \quad \mathbf{E}_\beta \xi_k^2(i, s) \leq \mathbf{E}_\beta |\eta_{sn} \tilde{u}_{k-1}^*| \leq 3 \Theta n^{-1/2} \mathbf{E}_\beta |u_{k-1}|$$

в силу (7.8.4). Отсюда следует, что

$$\begin{aligned} \mathbf{E}_\beta \left| n^{-1/2} \sum_{k=1}^n \xi_k(i, s) \right|^4 &\leq c_3 \left[n^{-1} t_n^3 \sup_{k \leq n} \mathbf{E}_\beta \xi_k^2(i, s) + \right. \\ &\quad \left. + \sup_{k \leq n} (\mathbf{E}_\beta \xi_k^2(i, s))^2 \right] = O(n^{-1} \log^3 n) \end{aligned}$$

равномерно по i, s . \square

Из лемм 7.8.1 и 7.8.2 следует, что (7.8.11) есть

$$O(N_n 3^{m_n} n^{-1} \log^3 n) = O(n^{-1/2} \log^5 n) = o(1)$$

в силу выбора последовательностей N_n и 3^{m_n} . Следовательно, (7.8.7) есть $o_p(1)$. Аналогично показывается, что и (7.8.6) есть $o_p(1)$. Тем самым установлено соотношение (7.8.2), что совместно с (7.8.1) влечет утверждение теоремы 7.7.1. \square

СПИСОК ЛИТЕРАТУРЫ

1. *Альтшулер С.В.* Методы оценки параметров процессов авторегрессии-скользящего среднего // *А и Т.*—1982.—Т. 8.—С. 5–18.
2. *Андерсон Т.* Статистический анализ временных рядов /Пер. с англ. под ред. Ю.К.Беляева.—М.: Мир, 1976.
3. *Бюкел П., Доксам К.* Математическая статистика. Вып. 1, 2 /Пер. с англ. Ю.А.Давылова.—М.: Финансы и статистика, 1983.
4. *Биллингсли П.* Сходимость вероятностных мер /Пер. с англ. под ред. В.В.Сазонова.—М.: Наука, 1977.
5. *Бокс Дж., Дженкинс Г.* Анализ временных рядов. Прогноз и управление. Вып. 1, 2 /Пер. с англ.—М.: Мир, 1974.
6. *Болдин М.В.* Оценка распределения возмущений в схеме авторегрессии // *Теория вероятн. и ее применен.*—1982.—Т. 27, № 4.—С. 905–910.
7. *Болдин М.В.* Проверка гипотез в схемах авторегрессии критериями Колмогорова и омега-квадрат // *ДАН СССР.*—1983.—Т. 273, № 1.—С. 19–22.
8. *Болдин М.В.* О проверке гипотез в схеме скользящего среднего критериями Колмогорова—Смирнова и омега-квадрат // *Теория вероятн. и ее применен.*—1989.—Т. 34, № 4.—С. 758–764.
9. *Болдин М.В.* Об оценке наименьших модулей в нестационарной авторегрессии и проверке стационарности // *Теория вероятн. и ее применен.*—1996.—Т. 41, № 2.—С. 409–417.
10. *Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики.—М.: Наука, 1983.
11. *Гаск Я., Шадак З.* Теория ранговых критериев /Пер. с англ. под ред. Л.Н.Большева.—М.: Наука, 1971.
12. *Джапаридзе К.О.* Критерии проверки сложных гипотез для случайных переменных и случайных процессов // *Теория вероятн. и ее применен.*—1977.—Т. 22, № 2.—С. 104–118.
13. *Дженкинс Г., Ваттс Д.* Спектральный анализ и его приложения. Вып. 1, 2 /Пер. с англ.—М.: Мир, 1971, 1972.
14. *Дубровин Б.А., Новиков С.П., Фоменко А.Т.* Современная геометрия: Методы и приложения.—М.: Наука, 1986.
15. *Ибрагимов И.А., Линник Ю.В.* Независимые и стационарно связанные величины.—М.: Наука, 1965.
16. *Ибрагимов И.А., Хасьминский Р.З.* Асимптотическая теория оценивания.—М.: Наука, 1979.
17. *Ивченко Г.И., Медведев Ю.И.* Математическая статистика.—М.: Высшая школа, 1992.
18. *Кокс Д., Хинкли Д.* Теоретическая статистика /Пер. с англ. под ред. Ю.К.Беляева.—М.: Мир, 1978.
19. *Леман Э.* Теория точечного оценивания /Пер. с англ.—М.: Наука, 1991.
20. *Мартин Р.Д.* Устойчивый авторегрессионный анализ временных рядов // *Устойчивые статистические методы оценки данных* /Под ред. Р.Л.Донера, Г.Н.Уилкинсона; пер. с англ. под ред. Н.Г.Волкова.— М.: Машиностроение, 1984.—С. 121–146.

21. Мартынов Г.В. Вычисление предельных распределений статистик критериев нормальности типа омега-квадрат // Теория вероятн. и ее применен.—1976.—Т. 21, № 1.—С. 3–15.
22. Мартынов Г.В. Критерии омега-квадрат.—М.: Наука, 1978.
23. Мудров В.И., Кушко В.Л. Методы обработки измерений. Квазиправдоподобные оценки.—М.: Радио и связь, 1983.
24. Питербарг В.И. Асимптотические методы в теории гауссовских случайных процессов и полей.—М.: Изд-во МГУ, 1988.
25. Поляк Б.Т., Цыпкин Я.З. Оптимальные методы оценивания коэффициентов авторегрессии при неполной информации // Изв. АН СССР. Техн. кибернетика.—1983.—Т. 6, № 1.—С. 118–126.
26. Симонова Г.И. Использование пакета «SIGN» для обучения новым методам статистического анализа данных // Международная конференция «Статистическое образование в современном мире: идеи, ориентации, технологии», 3–5 июля 1996 г.: Тезисы докладов.—С.-П.: Изд-во Санкт-Петербург. ун-та экономики и финансов, 1996.—С. 224–225.
27. Тюлки Дж. Анализ результатов наблюдений. Разведочный анализ / Пер. с англ. под ред. В.Ф.Писаренко.—М.: Мир, 1981.
28. Тюлгали А.Н. Об асимптотической допустимости критериев согласия // ДАН СССР. Сер. мат.—1982.—Т. 264, № 2.—С. 309–313.
29. Тюрик Ю.Н. О проверке параметрических гипотез непараметрическими методами // Теория вероятн. и ее применен.—1970.—Т. 15, № 4.—С. 745–749.
30. Тюрик Ю.Н. Линейная модель в многомерной непараметрической статистике // Многомерный статистический анализ в социально-экономических исследованиях.—М.: Наука, 1974.—С. 7–24.
31. Тюрик Ю.Н. О предельном распределении статистики Колмогорова—Смирнова для сложных гипотез // Изв. АН СССР. Техн. кибернетика.—1984.—Т. 48, № 6.—С. 1314–1343.
32. Тюрик Ю.Н. Статистические выводы для наблюдений с групповой структурой // Методы анализа данных, оценивания и выбора. Вып. 11.—М.: ВНИИСИ, 1984.—С. 3–28.
33. Тюрик Ю.Н. Знаковый статистический линейный анализ // Методы анализа данных, оценивания и выбора в системных исследованиях. Вып. 14.—М.: ВНИИСИ, 1986.—С. 4–16.
34. Тюрик Ю.Н. Непараметрические методы для линейных моделей временных рядов // Системные исследования в области компьютеризации и статистики.—М.: ИСА РАН, 1992.—С. 54–61.
35. Тюрик Ю.Н., Симонова Г.И. Знаковый анализ линейных моделей // Обзорные прикладной и промышленной математики. Вероятность и статистика.—1994.—Т. 1, вып. 2.—С. 214–278.
36. Хампель Ф., Рочетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния / Пер. с англ. под ред. В.М.Золотарева.—М.: Мир, 1989.
37. Хемман Э. Многомерные временные ряды / Пер. с англ. под ред. Ю.А.Розанова.—М.: Мир, 1974.
38. Хеттманспергер Т. Статистические выводы, основанные на рангах / Пер. с англ.—М.: Финансы и статистика, 1987.
39. Хмаладзе Э.В. Применение критериев типа омега-квадрат для проверки параметрических гипотез // Теория вероятн. и ее применен.—1979.—Т. 24, № 2.—С. 280–297.
40. Хмаладзе Э.В. Мартингалльный подход к теории критериев согласия // Теория вероятн. и ее применен.—1981.—Т. 26, № 2.—С. 240–257.
41. Холлендер М., Вульф М. Непараметрические методы статистики / Пер. с англ. под ред. Ю.П.Адлера, Ю.Н.Тюрпина.—М.: Финансы и статистика, 1983.
42. Хьюбер П. Робастность в статистике / Пер. с англ.—М.: Мир, 1984.
43. Шаров А.С., Новиков И.Д. Человек, открывший взрыв Вселенной: Жизнь и труд Эдвина Хаббла.—М.: Наука, 1989.

44. *Шуреев А.Н.* Вероятность.—М.: Наука, 1980.
45. *Anderson T.W.* The Statistical Analysis of Time Series.—New York: Wiley, 1971.
46. *Bahr B. von, Essen C.G.* Inequalities for the r th absolute moment of a sum of random variables, $1 \leq r \leq 2$ // *Ann. Math. Statist.*—1965.—V. 36.—P. 299–303.
47. *Basset G., Koenker R.* Asymptotic theory of least absolute error regression // *J. Amer. Statist. Assoc.*—1978.—V. 73, № 363.—P. 618–622.
48. *Bickel P.J., Doksum K.A.* Mathematical Statistics. Basic Ideas and Selected Topics.—San Francisco: Holden-Day, 1977.
49. *Billingsley P.* Convergence of Probability Measures.—New York: Wiley, 1968.
50. *Biometrika Tables for Statisticians. II* /Eds. E.S.Pearson and H.O.Hartley.—Cambridge: Cambridge Univ. Press, 1972.
51. *Bloomfield P., Steiger W.L.* Least Absolute Deviations. Theory, Applications and Algorithms.—Boston: Birkhäuser, 1983.
52. *Boldin M.V.* On median estimates and tests in autoregressive models // *Math. Methods Statist.*—1994.—V. 3, № 2.—P. 114–129.
53. *Boldin M.V.* On nonparametric sign tests in multiparameter autoregression // *Math. Methods Statist.*—1995.—V. 4, № 4.—P. 435–448.
54. *Boldin M.V.* On nonparametric sign estimators in multiparameter autoregression // *Math. Methods Statist.*—1996.—V. 5, № 2.—P. 154–172.
55. *Boldin M.V.* On residual empirical distribution function and rank estimators in autoregression // *Math. Methods Statist.*—1997.—V. 6, № 1.—P. 70–91.
56. *Boldin M.V., Tyurin Ya.N.* On nonparametric sign procedures in autoregression models // *Math. Methods Statist.*—1994.—V. 3, № 4.—P. 279–305.
57. *Box G.E.P., Jenkins G.M.* Time Series Analysis Forecasting and Control.—San Francisco: Holden-Day, 1970.
58. *Bustos O., Yohai V.J.* Robust estimates for ARMA models // *J. Amer. Statist. Assoc.*—1986.—V. 81.—P. 155–168.
59. *Cox D.R., Hinkley D.V.* Theoretical Statistics.—London: Chapman and Hall, 1974.
60. *Denoël E., Solway J.P.* Linear prediction of speech with a least absolute error criterion // *IEEE Trans. on ASSP.*—1985.—V. 33.—P. 1337–1403.
61. *Dufour J.M.* Rank tests for serial dependence // *J. Time Ser. Anal.*—1981.—V. 2.—P. 117–128.
62. *Durbin J.* Weak convergence of the sample distribution function when parameters are estimated // *Ann. Statist.*—1973.—V. 1.—P. 279–290.
63. *Eiker F.* Asymptotic normality consistency of the least squares estimators for families of linear regressions // *Ann. Math. Statist.*—1963.—V. 34.—P. 447–456.
64. *Evans G.B., Savin N.E.* The calculation of the least square estimator of the parameter in a random walk model // *Ann. Statist.*—1981.—V. 9.—P. 1114–1118.
65. *Ferretti N.E., Kelmansky D.M., Yohai V.J.* Estimators based on ranks for ARMA models // *Commun. Statist. Theory Meth.*—1991.—V. 20, № 12.—P. 3879–3907.
66. *Freedman W.L., Madore B.F., Mould J.R. et al.* Distance to the Virgo cluster galaxy M100 from Hubble Space Telescope observations of Cepheids // *Nature.*—1994.—V. 371.—P. 757–762.
67. *Hajek J., Šidak Z.* Theory of Rank Tests.—Prague: Academia, 1967.
68. *Hallin M., Ingenbleek J.Fr., Puri M.L.* Linear serial rank tests for randomness against ARMA alternatives // *Ann. Statist.*—1985.—V. 13, № 3.—P. 1156–1181.
69. *Hallin M., Ingenbleek J.Fr., Puri M.L.* Linear and quadratic rank tests for randomness against serial dependence // *J. Time Ser. Anal.*—1987.—V. 8.—P. 409–424.
70. *Hallin M., Puri M.L.* Optimal rank-based procedure for time-series analysis: Testing an ARMA model against other ARMA models // *Ann. Statist.*—1988.—V. 16.—P. 402–432.
71. *Hallin M., Puri M.L.* Aligned rank test for linear models with autocorrelated error terms // *J. Multivar. Anal.*—1994.—V. 50, № 2.—P. 175–237.
72. *Hampel F.R.* The influence curve and its role in robust estimation // *J. Amer. Statist. Assoc.*—1974.—V. 69.—P. 383–393.

73. *Hampel F.R., Ronchetti E.M., Rousseeuw R.J., Stahel W.A.* Robust Statistics. The Approach Based on Influence Functions.—New York: Wiley, 1986.
74. *Hannan E.J.* Multiple Time Series.—New York: Wiley, 1970.
75. *Heiler S., Willers R.* Asymptotic Normality of R -estimates in the linear model // *Statistics*.—1988.—V. 19.—P. 173–184.
76. *Hettmansperger Th.P.* Statistical Inference Based on Ranks.—New York: Wiley, 1984.
77. *Hodges J.L.Jr., Lehmann E.L.* Rank methods for combinations of independent experiments in analysis of variance // *Ann. Math. Statist.*—1962.—V. 33.—P. 482–497.
78. *Hodges J.L.Jr., Lehmann E.L.* Estimates of location based on rank tests // *Ann. Math. Statist.*—1963.—V. 34.—P. 598–611.
79. *Hollander M., Wolfe D.A.* Nonparametric Statistical Methods.—New York: Wiley, 1973.
80. *Hubble E.* A relation between distance and radial velocity among extra-galactic nebulae // *Astronomy*.—1929.—V. 15.—P. 168–173.
81. *Huber P.J.* Robust Statistics.—New York: Wiley, 1981.
82. *Jakoud J., Shiryaev A.N.* Stochastic Processes.—New York: Springer, 1987.
83. *Jenkins G.M., Watts D.G.* Spectral Analysis and Its Applications.—San Francisco: Holden-Day, 1968.
84. *Kashyap R.L., Rao A.R.* Dynamic Stochastic Models from Empirical Data.—New York: Academic Press, 1976.
85. *Koul H.L.* Asymptotic behavior of Wilcoxon type confidence regions in multiple linear regression // *Ann. Math. Statist.*—1969.—V. 40.—P. 1950–1979.
86. *Koul H.L.* A class of ADF tests for subhypothesis in the multiple linear regression // *Ann. Math. Statist.*—1970.—V. 41.—P. 1273–1281.
87. *Koul H.L.* A weak convergence result useful in robust autoregression // *Journal of Statistical Planning and Inference*.—1991.—V. 29, № 3.—P. 291–308.
88. *Koul H.L., Levental S.* Weak convergence of the residual empirical process in explosive autoregression // *Ann. Statist.*—1989.—V. 17, № 4.—P. 1784–1794.
89. *Koul H.L., Pflug J.C.* Weakly adaptive estimators in explosive autoregression // *Ann. Statist.*—1990.—V. 18, № 2.—P. 939–960.
90. *Krasker W.S., Welsh R.E.* Efficient bounded-influence regression estimation // *J. Amer. Statist. Assoc.*—1982.—V. 77.—P. 595–604.
91. *Kreiss J.P.* On adaptive estimation in stationary ARMA processes // *Ann. Statist.*—1987.—V. 15.—P. 112–133.
92. *Kreiss J.P.* Testing linear hypotheses in autoregressions // *Ann. Statist.*—1990.—V. 18.—P. 1470–1482.
93. *Kreiss J.P.* Estimation of the distribution function of noise in stationary processes // *Metrika*.—1991.—V. 38.—P. 285–297.
94. *Lehmann E.L.* Nonparametrics Statistical Methods Based on Ranks.—San Francisco: Holden-Day, 1975.
95. *Lehmann E.L.* Theory of Point Estimation.—New York: Wiley, 1983.
96. *Maldelbrot B.* The variation of certain speculative prices // *J. Business*.—1963.—V. 36.—P. 394–419.
97. *Maldelbrot B.* The variation of some other speculative prices // *J. Business*.—1967.—V. 40.—P. 383–413.
98. *Martin R.D.* Robust estimation for time series autoregression // *Robustness in Statistics* /Eds. R.L.Launer and G.N.Wilkinson.—New York: Academic Press, 1979.
99. *Martin R.D.* The Cramer—Rao bound and robust M -estimates for autoregressions // *Biometrika*.—1982.—V. 69.—P. 437–442.
100. *Martin R.D., Yohai V.J.* Influence functionals for time series // *Ann. Statist.*—1986.—V. 14.—P. 781–818.
101. *Mokkadem A.* Mixing properties of ARMA processes // *Stochastic Processes Appl.*—1988.—V. 29.—P. 309–315.

102. *Pollard D.* Asymptotics for least absolute deviation regression estimators // *Econometrics Theory*.—1991.—V. 7.—P. 186–199.
103. *Puri M.L., Sen P.K.* Nonparametric Methods in General Linear Models.—New York: Wiley, 1985.
104. *Rao M.M.* Asymptotic distribution of an estimator of the boundary parameter of an unstable process // *Ann. Statist.*—1978.—V. 6.—P. 185–190.
105. *Roy S.N.* Some Aspects of Multivariate Analysis.—New York: Wiley, 1957.
106. *Shiryayev A.N., Spokoiny V.G.* Statistical Experiments and Decisions: Asymptotic Theory.—New York: Springer, 1995.
107. *Stigler S.M.* The History of Statistics. The Measurements of Uncertainty before 1900.—Cambridge: The Belknap Press of Harvard University Press, 1986.
108. *Tukey J.W.* Exploratory Data Analysis. Preliminary Edition.—Mass.: Addison Wesley, 1977.
109. *Vaucouleurs G., Vaucouleurs A., Corwin H.G. et al.* Third Reference Catalog of Bright Galaxies (RC3).—New York: Springer, 1991.
110. *Walden A.T.* Non-Gaussian reflectivity, entropy and deconvolution // *Geophysics*.—1985.—V. 50.—P. 2862–2888.
111. *Yohai V.J., Maronna R.A.* Asymptotic behaviour of least-squares estimates for autoregressive processes with infinite variances // *Ann. Statist.*—1977.—V. 5, № 3.—P. 554–560.

M. V. BOLDIN, G. I. SIMONOVA, Yu. N. TYURIN

**SIGN-BASED METHODS
IN LINEAR STATISTICAL MODELS**

*Physical and Mathematical Literature Publishing Company
Russian Academy of Sciences*

Moscow, 1997, 288 pages

M. V. BOLDIN was graduated from Moscow State University in 1974. Ph.D. in mathematics in 1984 (Moscow State University). Associate Professor of Moscow State University, Faculty of Mechanics and Mathematics, Department of Probability. He published about 50 scientific papers on mathematical statistics especially on nonparametrics. Most of them are devoted to linear models of time series.

G. I. SIMONOVA was graduated from Moscow State University in 1973. Ph.D. in mathematics in 1982 (Moscow State University). The author of about 30 publications on data analysis, mathematical statistics and programming. Senior researcher, Institute of Mechanics, Moscow State University.

Yu. N. TYURIN was graduated from Moscow State University in 1958. Ph.D. in mathematics in 1965. Doctor of Mathematical Sciences in 1986 (Moscow State University). Professor of the Theory of Probability and Statistics, Moscow State University. Member of Moscow Mathematical Society. Yu.N. Tyurin is the author of about 70 publications, mainly on theory of Statistics and its applications. He is the author of a textbook «Data analysis with PC», 1994 (in coauthorship with A. A. Makarov).

In the book a new nonparametric approach to the analysis of statistical data is exposed. It consists in using only the signs of observations or of certain functions of them depending on the data structure. Hence the method is referred to as the sign-based one. The book treats regression and autoregression models important for applications. For these models the sign-based method yields the solutions of the principal statistical problems (parameter estimation, hypothesis testing, etc.). Both finite-sample and large-sample properties of the sign procedures are studied. The sign procedures are shown to be robust with respect to gross errors in the data. Numerical algorithms to implement the sign analysis are proposed, and examples of their application to real and simulated data are given. The exposition evolves from elementary to advanced theory to make the book accessible for broader readership. The book is intended for those studying or applying mathematical statistics.

The book can be used as a textbook on nonparametrics and time series analysis by:

- 1) graduate and undergraduate students who study Theory of Probability and Mathematical Statistics;
- 2) graduate and undergraduate students who study statistics as an applied method of data analysis;
- 3) applied statisticians and data analysts;
- 4) professionals in Theory of Probability and Mathematical Statistics.

The mathematical background requirement includes the basics of Calculus, Linear Algebra, Theory of Probability and Mathematical Statistics.

This book was published in English in 1997 by the American Mathematical Society in a series *Translations of Mathematical Monographs*, vol. 162 as «*Sign-based Methods in Linear Statistical Models*».

The authors are contacted by:
tel./fax: 7 (095) 9391403
E-mail: yuri@petr.math.msu.su
simonova@inmech.msu.su

CONTENTS

PREFACE INTRODUCTION

Part I

LINEAR MODELS OF INDEPENDENT OBSERVATIONS

Chapter 1. Sign-based Analysis of One-parameter Linear Regression

- § 1.1. Hubble's Law: a historical overview
- § 1.2. Determination of the Hubble constant by the sign-based method
- § 1.3. Asymptotic results
 - 1.3.1. Consistency. 1.3.2. Asymptotic normality.
- § 1.4. Influence function

Chapter 2. Sign Tests

- § 2.1. General linear model
- § 2.2. Locally optimal sign tests in the regression problem
- § 2.3. Evaluation of critical values: Asymptotic theory
- § 2.4. Example: Two-way layout
- § 2.5. Computation of critical values: Finite samples

Chapter 3. Sign Estimators

- § 3.1. Sign estimators and their computation
- § 3.2. Sign estimation: asymptotic theory
 - 3.2.1. The role of asymptotic theory. 3.2.2. Consistency of sign estimators. 3.2.3. Asymptotic normality of sign estimators. 3.2.4. Asymptotic covariance of sign estimators. 3.2.5. Uniform law of large numbers. 3.2.6. Theorem on uniform linearity. 3.2.7. Asymptotic power of sign tests. 3.2.8. Sensitivity curve.
- § 3.3. Comparison of estimators
 - 3.3.1. How estimators are compared. 3.3.2. Rank estimation. 3.3.3. Least squares and least absolute deviations estimators. 3.3.4. Asymptotic efficiency of sign estimators.

Chapter 4. Testing Linear Hypotheses

- § 4.1. Sign procedures for testing linear hypotheses
- § 4.2. Asymptotic properties of sign tests for linear hypotheses
- § 4.3. Examples
- § 4.4. Testing linear hypotheses in one- and two-way layout problems
 - 4.4.1. One-way analysis (t samples which may differ by location).
 - 4.4.2. Two-way analysis.
- § 4.5. Computation of critical values in testing linear hypotheses
 - 4.5.1. One-way analysis. 4.5.2. Two-way analysis.

Part II

LINEAR MODELS OF TIME SERIES

Chapter 5. Least Squares and Least Absolute Deviations Procedures in the Simplest Autoregressive Model

- § 5.1. Introduction
- § 5.2. The simplest stationary autoregressive equation and its solutions

- § 5.3. Least squares procedures
 - 5.3.1. Least squares estimators. 5.3.2. Tests based on the LSE.
- § 5.4. Least squares estimator in nonstationary autoregression
- § 5.5. Least absolute deviations procedures
 - 5.5.1. Least absolute deviations estimator. 5.5.2. Tests based on the LAD estimator. 5.5.3. Weighted least absolute deviations estimators.
- § 5.6. Influence functionals of least squares and least absolute deviations estimators
 - 5.6.1. Influence functional of the least squares estimator. 5.6.2. Influence functional of the LAD estimator. 5.6.3. Influence functional of weighted LAD estimators.
- § 5.7. Testing for stationarity of the autoregression process
- § 5.8. Proofs

Chapter 6. Sign-based Analysis of One-parameter Autoregression

- § 6.1. Introduction to sign-based autoregression analysis
- § 6.2. Sign tests
- § 6.3. Sign tests in a nonstationary autoregression
- § 6.4. Uniform stochastic expansion: the power of sign tests under local alternatives
- § 6.5. Sign tests: comparison with other nonparametric tests
- § 6.6. Sign estimators
 - 6.6.1. Sign estimator $\hat{\beta}_{n,S}$. 6.6.2. Sign estimator $\beta_{n,S}^*$. 6.6.3. Sign estimator $\tilde{\beta}_{n,S}^*$.
- § 6.7. Influence functionals of sign estimators
 - 6.7.1. Influence functional of the sign estimator $\hat{\beta}_{n,S}$. 6.7.2. Influence functional of the sign estimator $\tilde{\beta}_{n,S}$. 6.7.3. Influence functional of the sign estimator $\beta_{n,S}^*$.
- § 6.8. Simulation results: evaluation of quantiles, confidence sets, contaminated samples
 - 6.8.1. Evaluation of quantiles. 6.8.2. Confidence estimation of β . 6.8.3. Sign estimation from contaminated samples.
- § 6.9. Proof of Theorem 6.4.1

Chapter 7. Sign-based Analysis of the Multiparameter Autoregression

- § 7.1. Introduction
- § 7.2. Test statistics and their null distributions
- § 7.3. Uniform stochastic expansion: the power of sign tests under local alternatives
- § 7.4. Testing linear hypotheses
- § 7.5. Sign-based estimators
 - 7.5.1. Sign estimator $\hat{\beta}_{n,S}$. 7.5.2. Sign estimator $\tilde{\beta}_{n,S}$. 7.5.3. Sign estimator $\beta_{n,S}^*$.
- § 7.6. Influence functionals of estimators in the multiparameter autoregression
 - 7.6.1. Influence functional of the least squares estimator $\hat{\beta}_{n,LS}$. 7.6.2. Influence functional of the least absolute deviations $\hat{\beta}_{n,LAD}$. 7.6.3. Influence functionals of weighted LAD $\hat{\beta}_{n,LADW}$. 7.6.4. Influence functional of the sign estimator $\tilde{\beta}_{n,S}$. 7.6.5. Influence functional of the sign estimator $\beta_{n,S}^*$.
- § 7.7. Empirical distribution function of residuals and related empirical processes
- § 7.8. Proof of Theorem 7.7.1

BIBLIOGRAPHY