

И. Г. ВЕНЕЦКИЙ, Г. С. КИЛЬДИШЕВ

ПОСОБИЕ  
ПО  
МАТЕМАТИЧЕСКОЙ  
СТАТИСТИКЕ

ГОССТАТИЗДАТ  
МОСКВА 1956

И. Г. ВЕНЕЦКИЙ, Г. С. КИЛЬДИШЕВ

# ПОСОБИЕ ПО МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

3

ГОСУДАРСТВЕННОЕ СТАТИСТИЧЕСКОЕ ИЗДАТЕЛЬСТВО  
Москва 1956



Настоящее пособие предназначено прежде всего для студентов экономических вузов, изучающих курс математической статистики. Имеется в виду, что оно поможет им в приобретении практических навыков при производстве разного рода математико-статистических расчетов, при использовании различных таблиц и т. п. В соответствии с этим теоретический материал сведен к минимуму и служит лишь для того, чтобы напомнить читателю самые основные положения и формулы без отсылки к теоретическим курсам.

Пособие может быть также рекомендовано широкому кругу читателей, которые в практической работе используют приемы и методы математической статистики.

Исходя из вопросов, обычно включаемых в программы экономических вузов, авторы соответственно ограничили объем пособия и не претендуют на полный охват всех проблем математической статистики.

Разделы I, II и III написаны И. Г. Венецким, разделы IV и V, задания и приложения — Г. С. Кильдишевым.

Научную редакцию осуществлял проф. А. Я. Боярский.

---

## РАЗДЕЛ I

### ВАРИАЦИОННЫЙ РЯД И ЕГО ХАРАКТЕРИСТИКИ

§ 1<sup>1</sup>. При измерении величины явления мы получаем определенные результаты. Статистическая обработка этих результатов путем группировки позволяет произвести подсчет числа единиц, обладающих данным значением признака.

Если даны отдельные значения признака (варианты —  $x$ ) в возрастающем или убывающем порядке и абсолютные количества единиц, обладающих данным значением признака (частоты —  $m$ ), то говорят, что дан упорядоченный (ранжированный) *вариационный ряд* или упорядоченное *распределение признака*.

Условились отдельные значения признака обозначать  $x_1, x_2, x_3, \dots x_n$  и называть *вариантами*, а абсолютное число случаев для каждого из вариантов обозначать  $m_1, m_2, m_3, \dots m_n$  и называть частотами.

Пример 1. Наблюдение за весом 50 деталей дало следующие результаты (в г): 83, 85, 81, 82; 84, 82, 79, 84, 80, 81, 82, 82, 80, 82, 80, 82, 83, 84, 79, 79, 83, 82, 83, 85, 82, 82, 81, 80, 82, 82, 83, 80, 82, 85, 81, 83, 81, 81, 83, 82, 81, 85, 83, 79, 81, 85, 81, 84, 81, 82.

Чтобы построить по этим данным вариационный ряд, располагаем полученные варианты (вес детали) и частоты (число деталей) следующим образом:

Таблица 1

Отдельные значения признака (Вес одной детали— $x$ )	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	Объем совокупности (Итого деталей)
	79	80	81	82	83	84	85	
Частоты (число деталей— $m$ )	4	5	10	14	8	4	5	50

<sup>1</sup> Нумерация параграфов, примеров, формул и таблиц дана для каждого раздела в отдельности, начиная с первого номера.

§ 2. Вариация признака может быть дискретной и непрерывной. *Дискретной вариацией* признака называется такая, при которой отдельные значения признака (варианты) отличаются друг от друга на некоторую конечную величину (обычно — целое число). *Непрерывной вариацией* называется вариация, при которой значения признака могут отличаться одно от другого на сколь угодно малую величину. В качестве примера можно привести: для дискретной вариации признака — число станков, обслуживаемых одним рабочим, число семян в 1 кг и т. д.; для непрерывной вариации признака — процент выполнения рабочим нормы выработки, вес одного семени и т. д.

§ 3. При непрерывной вариации распределение признака называется *интервальным*. Частоты относятся не к отдельному значению признака, как это бывает при дискретной вариации, а ко всему интервалу. Часто за значение интервала принимают его середину, т. е. *центральное значение*.

В качестве примера можно привести интервальный вариационный ряд по проценту выполнения норм выработки.

Пр и м е р 2. Распределение рабочих по проценту выполнения норм:

Т а б л и ц а 2

Процент выполнения норм (интервалы)	95—100	100—105	105—110	110—115	115—120	120—125	125—130	130—135	135 и выше
Число рабочих (частоты)	3	17	28	64	127	71	40	11	4
Центральное значение интервала	97,5	102,5	107,5	112,5	117,5	122,5	127,5	132,5	Условно принимают 137,5

§ 4. Вместо приведенных выше абсолютных количеств, т. е. частот, можно вычислить и использовать относительную величину — долю частоты того или иного варианта или интервала в сумме всех частот, которая называется *частостью*:

$$w_i = \frac{m_i}{\Sigma m_i}, \quad (1)$$

где  $w_i$  — частость  $i$  варианта или интервала;

$\Sigma$  — знак суммирования;

$m_i$  — частота  $i$  варианта или интервала.

Так:

$$w_1 = \frac{m_1}{\Sigma m}; \quad w_2 = \frac{m_2}{\Sigma m}; \quad w_3 = \frac{m_3}{\Sigma m}; \dots w_n = \frac{m_n}{\Sigma m}.$$

Сумма всех частотей равна единице:

$$\sum w_i = 1. \quad (2)$$

Частоты можно выражать и в процентах, тогда сумма всех частотей равна 100%.

Пример 3.

Т а б л и ц а 3

Варианты ( $x_i$ )	10	12	14	16	18	Итого
Частоты ( $m_i$ )	3 ( $m_1$ )	5 ( $m_2$ )	8 ( $m_3$ )	2 ( $m_4$ )	2 ( $m_5$ )	20 ( $\Sigma m$ )
Частоты ( $w_i$ ) (в виде доли)	0,15	0,25	0,40	0,10	0,10	1,00
Частоты ( $w_i$ ) (в виде процентов)	15	25	40	10	10	100

#### Вычисление частотей

Частоты  $w_i$  (в виде доли):

$$w_1 = \frac{3}{20} = 0,15; \quad w_2 = \frac{5}{20} = 0,25; \quad w_3 = \frac{8}{20} = 0,40;$$

$$w_4 = \frac{2}{20} = 0,10; \quad w_5 = \frac{2}{20} = 0,10.$$

Частоты  $w_i$  (в виде процентов):

$$w_1 = \frac{3}{20} \cdot 100 = 15\%; \quad w_2 = \frac{5}{20} \cdot 100 = 25\% \text{ и т. д.}$$

§ 5. В интервальном вариационном ряду, в каждом интервале различают *нижнюю и верхнюю границы интервала*:

нижняя граница интервала  $x_{i(\min)}$ ;

верхняя граница интервала  $x_{i(\max)}$ ;

величина интервала  $k = x_{i(\max)} - x_{i(\min)}$ .

§ 6. Интервальные вариационные ряды бывают с одинаковыми и неодинаковыми интервалами. В последнем случае чаще всего встречаются *интервалы последовательно увеличивающиеся*.

Пример 4а. Вариационный ряд с равными интервалами:

Т а б л и ц а 4а

Размеры деталей в см (интервалы)	5,40— 5,45	5,45— 5,50	5,50— 5,55	5,55— 5,60	5,60— 5,65	Итого
Количество деталей (частоты)	500	1200	3000	1700	300	6700

6. Вариационный ряд с последовательно увеличивающимися интервалами:

Т а б л и ц а 46

Группы рабочих по возрасту (интервалы в годах)	17—18	18—20	20—25	25—35	35 и старше	Итого
Количество рабочих (частоты)	7	11	32	40	25	115

§ 7. В качестве характеристики ряда распределения применяют *плотность распределения*, которую вычисляют как отношение частот к величине интервала.

Различают абсолютную плотность распределения:

$$P_A = \frac{m_i}{k_i} \quad (3)$$

и относительную плотность распределения:

$$P_o = \frac{w_i}{k_i}, \quad (4)$$

где  $P_A$  и  $P_o$  — плотности распределения, абсолютная (со значком  $A$ ) и относительная (со значком  $o$ ).

Пример 5. По данным примера 4а вычисляем абсолютную плотность распределения. Получаем: для первого интервала

абсолютная плотность  $P_{A(1)} = \frac{500}{0,05} = 10\,000$ ; для второго ин-

тервала  $P_{A(2)} = \frac{1200}{0,05} = 24\,000$  и т. д.

§ 8. Большое значение для наглядного представления вариационного ряда имеют графические методы его изображения. Вариационный ряд графически может быть изображен в виде полигона, гистограммы, кумуляты и огивы.

*Полигон распределения* строится в прямоугольной системе координат. Величина признака откладывается на оси абсцисс, частоты или частости (точнее — плотности распределения) — по оси ординат. Вершины ординат соединяются прямыми линиями. Чаще всего полигоны применяются для изображения дискретных вариационных рядов, но могут быть применены и для интервальных рядов. В этом случае ординаты, пропорциональные частоте или



частоты интервала, восстанавливаются перпендикулярно оси абсцисс в точке, соответствующей середине данного интервала. Для замыкания крайние ординаты соединяются с серединой интервалов, в которых частоты или частости равны нулю.

**Пример 6.** По данным примера 2 строим полигон.

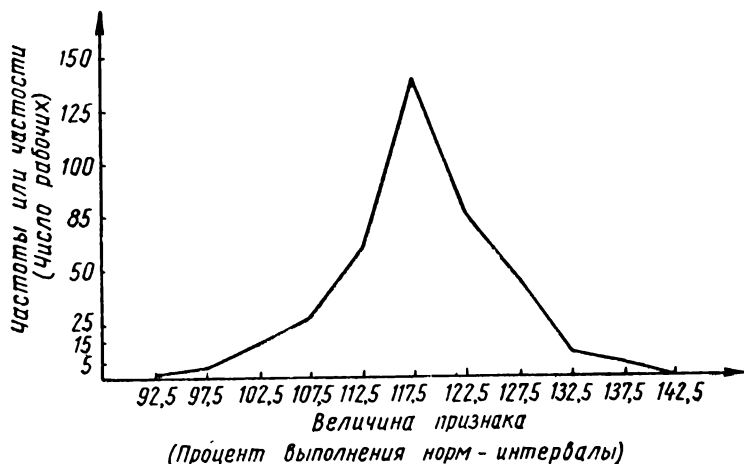


График 1. Полигон.

*Гистограмма распределения* строится аналогично полигону в прямоугольной системе координат. Отличие гистограммы от полигона состоит в том, что на оси абсцисс берутся не точки, а отрезки, изображающие интервал, а вместо ординат, соответствующих частотам или частостям отдельных вариантов, строят прямоугольники с высотой, пропорциональной частотам или частостям интервала.

В случае неравенства интервалов гистограмма распределения строится не по частотам или частостям, а по плотности интервалов (абсолютной или относительной). При этом общая площадь гистограммы равна численности совокупности, если построение производится по абсолютной плотности, или единице — если гистограмма построена по относительной плотности.

Если соединить прямыми линиями середины верхних сторон прямоугольников, то получим полигоны распределения.

Разбивая интервалы на несколько частей и исходя из того, что вся площадь гистограммы остается неизменной, можно получить мелкоступенчатую гистограмму, которая в пределе (за счет уменьшения величины интервала) перейдет в плавную кривую, называемую «кривой распределения».

Пример 7. Имеются данные о диаметре 200 валиков.

Таблица 5  
Размеры партии валиков<sup>1</sup>

№ п,п	Диаметр (в мм)	№ п/п	Диаметр (в мм)	№ п/п	Диаметр (в мм)	№ п,п	Диаметр (в мм)	№ п/п	Диаметр (в мм)
1	49,945	41	49,941	81	49,930	121	49,941	161	49,943
2	946	42	925	82	936	2	936	2	924
3	961	43	934	83	951	3	955	3	947
4	957	44	939	84	940	4	948	4	945
5	950	45	941	85	950	5	934	5	947
6	946	46	940	86	936	6	935	6	937
7	934	47	949	87	930	7	934	7	944
8	945	48	939	88	950	8	933	8	934
9	944	49	931	89	938	9	934	9	940
10	951	50	944	90	928	130	945	170	937
11	942	51	939	91	950	1	949	1	951
12	941	52	947	92	948	2	933	2	939
13	946	53	950	93	934	3	934	3	938
14	921	54	933	94	938	4	943	4	933
15	936	55	945	95	934	5	931	5	938
16	939	56	955	96	951	6	935	6	958
17	943	57	949	97	932	7	937	7	940
18	949	58	948	98	934	8	945	8	948
19	934	59	934	99	943	9	953	9	950
20	927	60	944	100	923	140	945	180	943
21	932	61	952	1	946	1	959	1	943
22	944	62	942	2	926	2	923	2	937
23	945	63	938	3	942	3	946	3	939
24	939	64	938	4	936	4	941	4	948
25	950	65	938	5	942	5	942	5	918
26	945	66	936	6	943	6	934	6	937
27	940	67	949	7	941	7	951	7	939
28	947	68	930	8	944	8	934	8	943
29	940	69	928	9	928	9	929	9	946
30	945	70	939	110	947	150	940	190	926
31	946	71	950	1	941	1	951	1	927
32	952	72	951	2	958	2	940	2	932
33	942	73	944	3	951	3	944	3	938
34	944	74	943	4	942	4	946	4	932
35	942	75	940	5	948	5	940	5	932
36	948	76	930	6	922	6	929	6	939
37	945	77	935	7	930	7	936	7	957
38	935	78	938	8	954	8	950	8	942
39	948	79	945	9	939	9	940	9	950
40	958	80	935	120	948	160	943	200	941

<sup>1</sup> А. Н. Журавлев, Статистический метод контроля в машиностроении, 1948.

Чтобы по этим данным построить гистограмму, а затем превратить ее в мелкоступенчатую, производим группировку данных с равными интервалами и получаем интервальный вариационный ряд, в котором интервальная разность  $k=0,01$  мм.

Т а б л и ц а 6

## Размеры партии валиков (1-я группировка)

Диаметр (в мм) (интервалы)	Количество вали- ков (частоты)	Центр интервала
49,918 — 49,928	14	49,923
49,928 — 49,938	58	49,933
49,938 — 49,948	92	49,943
49,948 — 49,958	34	49,953
49,958 — 49,968	2	49,963
И т о г о . . .	200	—

По полученному вариационному ряду строим гистограмму распределения (см. график 2).

Соединяем прямыми линиями середины верхних сторон прямоугольников и получаем полигон. Для получения мелкоступенчатой гистограммы разбиваем интервалы на две равные части и получаем:

Т а б л и ц а 7

## Размеры партии валиков (2-я группировка)

Диаметр (в мм) (интервалы)	Количество вали- ков (частоты)	Центр интервала
1	2	3
49,918 — 49,923	5	49,9205
49,923 — 49,928	9	49,9255
49,928 — 49,933	18	49,9305
49,933 — 49,938	40	49,9355
49,938 — 49,943	49	49,9405
49,943 — 49,948	43	49,9455
49,948 — 49,953	26	49,9505
49,953 — 49,958	8	49,9555
49,958 — 49,963	2	49,9605
49,963 — 49,968	0	49,9655
И т о г о . . .	200	—

Если построить гистограмму по новому вариационному ряду, с уменьшенными интервалами, то получим гистограмму с более мелкими ступенями. Учет требования о неизменности площади гистограммы приводит к необходимости увеличить масштаб оси ординат вдвое (см. график 3).

Можно продолжить процесс расчленения интервалов и дальше, получая все более и более мелкоступенчатую гистограмму.

*Кумулятивная кривая* (кривая сумм — кумюлята) получается при изображении вариационного ряда с накопленными частотами

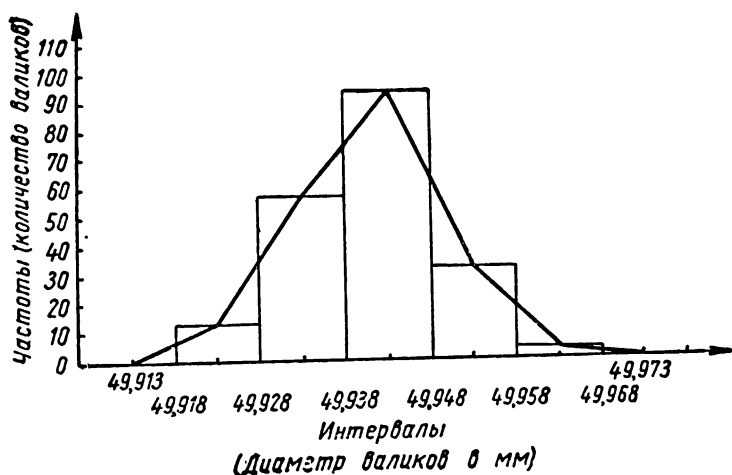


График 2. Гистограмма.

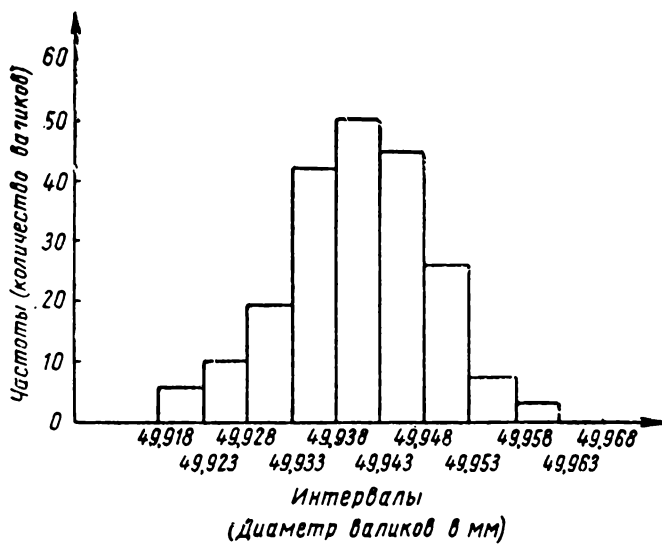


График 3. Гистограмма с уменьшенными ступенями.

или частотами в прямоугольной системе координат. При построении кумуляты дискретного признака на ось абсцисс наносятся значения признака (варианты). Ординатами служат вертикальные отрезки, длина которых пропорциональна накопленной частоте или частоте того или иного варианта. Соединением вершин ординат прямыми линиями получаем ломаную (кривую) кумуляту.

**Пример 8.** По данным табл. 5 построить кумуляту.

Составляем упорядоченный дискретный вариационный ряд с накопленными частотами (при наличии частостей можно для построения кумуляты пользоваться ими).

Накопленная частота определенного варианта получается суммированием всех частот вариантов, предшествующих данному, с частотой этого варианта.

Таблица 8

**Упорядоченный, дискретный вариационный ряд  
(размеров валиков) с накопленными частотами**

Диаметр (в мм)	Количество валиков	Накопленные частоты	Диаметр (в мм)	Количество валиков	Накопленные частоты	Диаметр (в мм)	Количество валиков	Накопленные частоты
49,918	1	1	49,933	4	32	49,948	9	164
919	—	1	934	14	46	949	5	169
920	—	1	935	5	51	950	10	179
921	1	2	936	7	58	951	8	187
922	1	3	937	5	63	952	2	189
923	2	5	938	9	72	953	1	190
924	1	6	939	11	83	954	1	191
925	1	7	940	11	94	955	2	193
926	2	9	941	8	102	956	—	193
927	2	11	942	9	111	957	2	195
928	3	14	943	10	121	958	3	198
929	2	16	944	9	130	959	1	199
930	5	21	945	12	142	960	—	199
931	2	23	946	8	150	961	1	200
932	5	28	947	5	155	Итого	200	—

Кумулята для этого примера — см. график 4.

При построении кумуляты интервального вариационного ряда нижней границе первого интервала соответствует частота, равная нулю, а верхней границе — вся частота интервала. Верхней границе второго интервала соответствует накопленная частота первых двух интервалов (т. е. сумма частот этих интервалов) и т. д. Верхней границе последнего (максимального) интервала соответствует накопленная частота, равная сумме всех частот.

**Пример 9.** По данным табл. 7 построить кумуляту.

Составляем интервальный вариационный ряд с накопленными частотами.

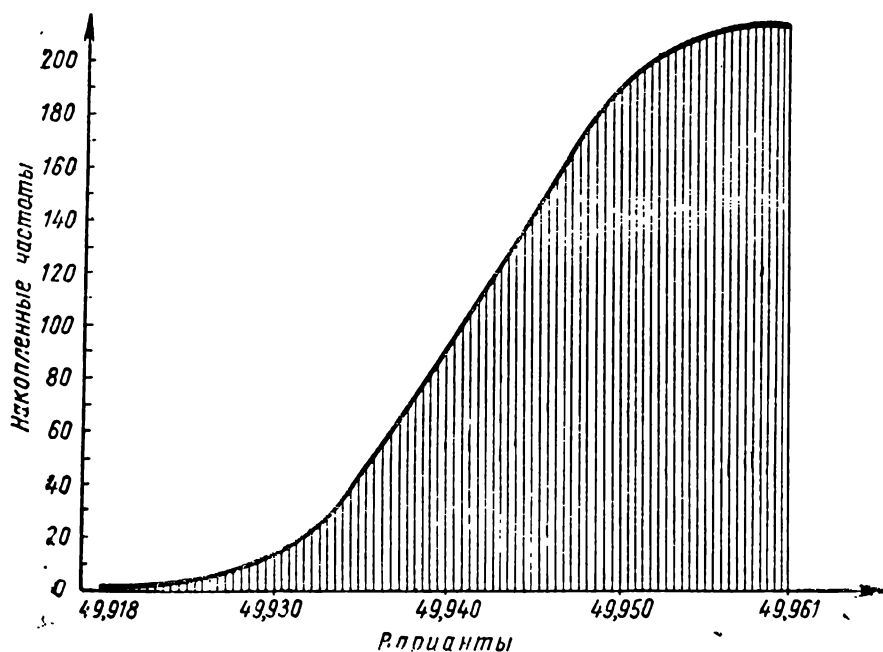
Таблица 9

**Упорядоченный, интервальный вариационный ряд  
с накопленными частотами**

Диаметр (в мм) (интервалы)	Количество вали- ков (частоты)	Накопленные частоты
49,918 — 49,923	5	5
49,923 — 49,928	9	14
49,928 — 49,933	18	32
49,933 — 49,938	40	72
49,938 — 49,943	49	121
49,943 — 49,948	43	164
49,948 — 49,953	26	190
49,953 — 49,958	8	198
49,958 — 49,963	2	200
Итого . . .	200	—

По полученным накопленным частотам строим кумуляту (см. график 5).

*Огива* строится аналогично кумуляте с той лишь разницей, что



**График 4. Кумулята дискретного вариационного ряда.**

на ось абсцисс наносят накопленные частоты, а на ось ординат — значения признака. Если лист бумаги, на котором изображена

кумулята, повернуть на  $90^\circ$  и посмотреть на него с обратной стороны на свет, то можно увидеть огиву.

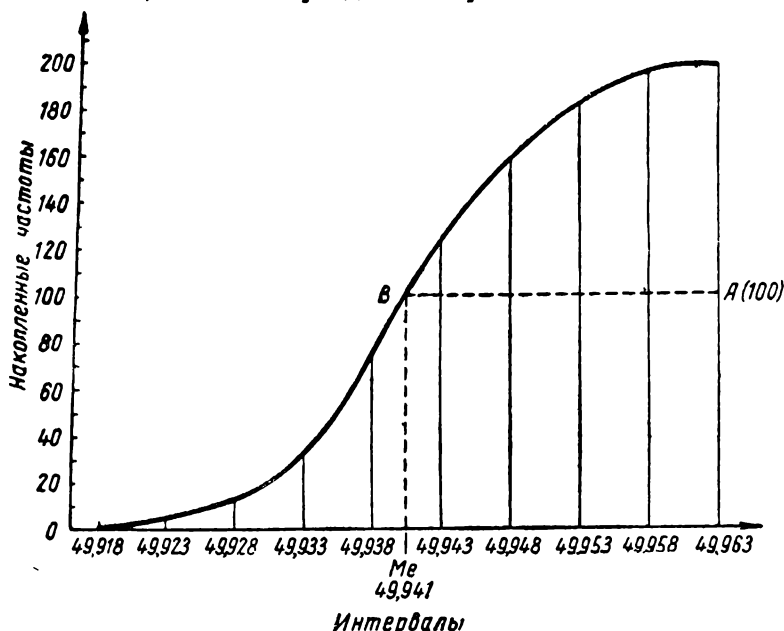


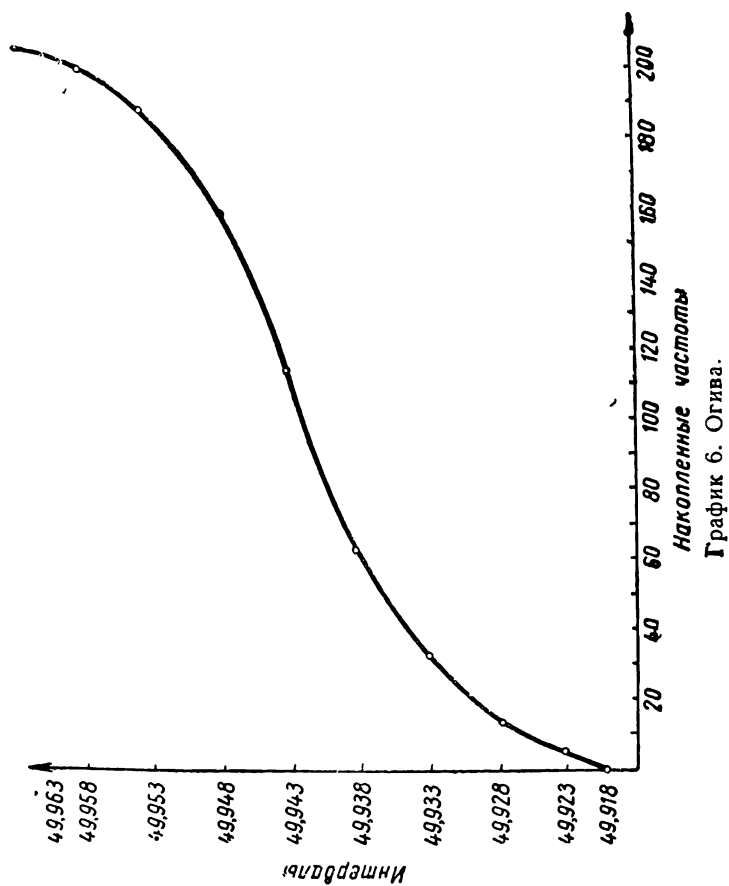
График 5. Кумулята интервального вариационного ряда.

Пример 10. По данным табл. 9 построить огиву (см. график 6).

§ 9. В качестве одной из важнейших характеристик вариационного ряда применяют среднюю величину. Математическая статистика различает ряд типов средних величин: гармоническую, геометрическую, арифметическую, квадратическую, кубическую и др. Все перечисленные типы средних могут быть исчислены для случаев, когда каждый из вариантов вариационного ряда встречается только по одному разу, — тогда средняя называется простой или невзвешенной, — и для случаев, когда варианты или интервалы повторяются различное число раз. При этом число повторений вариантов или интервалов называют частотой или статистическим весом, а среднюю вычисленную с учетом статистического веса называют взвешенной средней.

Выбор одного из перечисленных типов средних для характеристики вариационного ряда производится не произвольно, а в зависимости от особенностей изучаемого явления и цели, для которой средняя исчисляется.

Практически при выборе того или другого типа средней следует исходить из принципа осмысленности результата, при суммировании или при взвешивании. Если в результате взвешивания получаются величины, не имеющие смыслового значения, то выбор типа средней возможно произведен неверно.





Следует иметь в виду, что средняя только в том случае является обобщающей характеристикой, если она применяется к однородной совокупности. В случаях использования средней для неоднородных совокупностей можно прийти к неверным выводам. Научной основой статистического анализа является метод статистических группировок, т. е. расчленения совокупности на качественно однородные группы.

Все указанные типы средних величин могут быть получены из формул степенной средней. Если имеются варианты  $x_1^z; x_2^z; x_3^z; \dots x_n^z$ , то средняя из вариантов может быть исчислена по формуле *простой невзвешенной степенной средней*:

$$\bar{x} = \sqrt[z]{\frac{\sum x^z}{n}}. \quad (5)$$

При наличии соответствующих частот  $m_1; m_2; m_3; \dots m_n$  средняя исчисляется по формуле *взвешенной степенной средней*:

$$\bar{x} = \sqrt[z]{\frac{\sum x^z m}{\sum m}}, \quad (6)$$

где  $\bar{x}$  — степенная средняя;

$z$  — показатель степени, определяющий тип средней;

$x$  — варианты;

$m$  — частоты или статистические веса вариантов.

*Средняя арифметическая* получается из формулы степенной средней при подстановке  $z=1$ :

$$\bar{x}_a = \frac{\sum x}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (7)$$

— средняя арифметическая невзвешенная и

$$\bar{x}_a = \frac{\sum x m}{\sum m} = \frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{m_1 + m_2 + \dots + m_n} \quad (8)$$

— средняя арифметическая взвешенная.

**П р и м е р 11.** Измерение 20 единиц продукции дали следующие результаты (колонки 1 и 2):

Т а б л и ц а 10

Размер единицы продукции (в см) ( $x$ )	Количество единиц ( $m$ )	$xm$
1	2	3
10,3	1	10,3
10,5	3	31,5
10,7	9	96,3
10,9	5	54,5
11,1	2	22,2
Итого . . .	20	214,8

Вычислить средний размер единицы продукции.

Находим среднюю арифметическую по формуле 8. Для этого исчисляем в табл. 10 колонку 3.

$$\bar{x}_a = \frac{\sum x m}{\sum m} = \frac{214,8}{20} = 10,74.$$

Здесь умножение значения признака на вес и суммирование этих произведений дает общий размер продукции, т. е. имеет реальный смысл.

Средняя гармоническая получается при подстановке в формулу степенной средней значения  $z = -1$ .

Средняя гармоническая простая:

$$\bar{x}_{\text{гарм}} = \sqrt[n]{\frac{\sum x^{-1}}{n}} = \frac{1}{\frac{\sum \frac{1}{x}}{n}} = \frac{n}{\sum \frac{1}{x}};$$

получаем:

$$\bar{x}_{\text{гарм}} = \frac{n}{\sum \frac{1}{x}}. \quad (9)$$

Средняя гармоническая взвешенная:

$$\bar{x}_{\text{гарм}} = \frac{\sum m}{\sum \frac{m}{x}}. \quad (10)$$

Средняя гармоническая применяется в тех случаях, когда суммируемый признак выражен обратной величиной данного признака, т. е.:

$$\frac{1}{x_1}; \frac{1}{x_2}; \frac{1}{x_3}; \dots \frac{1}{x_n}$$

или

$$x_1^{-1}; x_2^{-1}; x_3^{-1}; \dots x_n^{-1}.$$

Пример 12. По следующим данным о работе 22 рабочих в течение 8 час. вычислить среднюю гармоническую взвешенную:

Т а б л и ц а 11

Затраты времени на изготовление одной дета- ли (в мин.) ( $x$ )	Количество рабо- чих ( $m$ )	$\frac{m}{x}$
10	2	0,20
12	9	0,75
14	7	0,50
16	4	0,25
Итого . . .	22	1,70

$$\bar{x}_{гарм} = \frac{\Sigma m}{\Sigma \frac{m}{x}} = \frac{22}{1,7} \approx 12,95.$$

В данном случае взвешивание состоит в делении по каждой группе количества рабочих ( $m$ ) на затраты времени по изготовлению одной детали ( $x$ ). Для проверки правильности выбора типа средней осмыслим результат взвешивания. Исходя из того, что все рабочие работали по 8 час., количество рабочих можно рассматривать как величину, определяющую общие затраты времени. Тогда результат деления представит вполне осмысленную величину:

$$\frac{\text{Количество рабочих}}{\text{Затраты времени на 1 деталь}} = \frac{\text{Общие затраты времени}}{\frac{\text{Общие затраты времени}}{\text{Количество изготовленных деталей}}} =$$

= Количество изготовленных деталей за 8 час.

Таким образом, применение средней гармонической является правильным.

Средняя квадратическая получается из формулы степенной средней при подстановке  $z=2$ :

$$\bar{x}_{квadr} = \sqrt[n]{\frac{\Sigma x^2}{n}} = \sqrt[n]{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} \quad (11)$$

— средняя квадратическая невзвешенная и

$$\bar{x}_{квadr} = \sqrt{\frac{\Sigma x^2 m}{\Sigma m}} = \sqrt{\frac{x_1^2 m_1 + x_2^2 m_2 + \dots + x_n^2 m_n}{m_1 + m_2 + \dots + m_n}} \quad (12)$$

— средняя квадратическая взвешенная.

Пример 13. Измерения сторон 10 квадратных площадок дали следующие результаты (колонки 1 и 2):

Т а б л и ц а 12

Длина стороны квадратной площадки (в м) ( $x$ )	Количество площадок ( $m$ )	$x^2$	$x^2 \cdot m$
1	2	3	4
5	1	25	25
6	3	36	108
7	4	49	196
8	1	64	64
9	1	81	81
Итого . . .	10	—	474

Вычислить среднюю сторону квадратных площадок.

Находим среднюю квадратическую взвешенную по формуле 12, для этого исчисляем в табл. 12 колонки 3 и 4; получаем:

$$\bar{x}_{\text{квдр}} = \sqrt{\frac{\sum x^2 m}{\sum m}} = \sqrt{\frac{474}{10}} = \sqrt{47,4} = 6,88.$$

В данном примере взвешивание квадрата стороны площадки по количеству площадок дает осмысленный результат, т. е. общую площадь всех площадок группы.

Средняя кубическая получается из формулы степенной средней при подстановке  $z=3$ :

$$\bar{x}_{\text{куб}} = \sqrt[3]{\frac{\sum x^3}{n}} = \sqrt[3]{\frac{x_1^3 + x_2^3 + \dots + x_n^3}{n}} \quad (13)$$

— средняя кубическая невзвешенная и

$$\bar{x}_{\text{куб}} = \sqrt[3]{\frac{\sum x^3 m}{\sum m}} = \sqrt[3]{\frac{x_1^3 m_1 + x_2^3 m_2 + \dots + x_n^3 m_n}{m_1 + m_2 + \dots + m_n}} \quad (14)$$

— средняя кубическая взвешенная.

Средняя геометрическая получается из формулы степенной средней при подстановке  $z=0$  и раскрытия неопределенности вида  $1^{\infty}$ :

$$\bar{x}_{\text{геом}} = \sqrt[n]{\prod(x)} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}, \text{ т. е.} \quad (15)$$

— средняя геометрическая невзвешенная и

$$\bar{x}_{\text{геом}} = \sqrt[n]{\prod(x^m)} = \sqrt[n]{x_1^{m_1} \cdot x_2^{m_2} \cdot \dots \cdot x_n^{m_n}}, \text{ т. е.} \quad (16)$$

— средняя геометрическая взвешенная,

где  $\Pi$  — знак произведения;

$n$  — число вариантов;

$m$  — частота каждого варианта.

Вычисления по формуле 16 в значительной мере упрощаются применением логарифмирования. Для невзвешенной средней геометрической:

$$\begin{aligned} \bar{x}_{\text{геом}} &= \sqrt[n]{\prod(x)} \\ \lg \bar{x}_{\text{геом}} &= \frac{\sum \lg x}{n}. \end{aligned} \quad (17)$$

Для взвешенной средней геометрической:

$$\bar{x}_{гeом} = \sqrt[m]{\prod (x^m)}$$

$$\lg \bar{x}_{гeом} = \frac{\sum (\lg x) m}{\sum m} . \quad (18)$$

Таким образом, логарифм средней геометрической есть средняя арифметическая из логарифмов вариантов (см. формулы 7 и 8).

Средняя геометрическая используется главным образом при изучении динамики.

Расчет средних коэффициентов и темпов роста производится по формулам средней геометрической.

Пример 14. Выпуск промышленной продукции производился предприятием в следующих размерах.

Т а б л и ц а 13

Месяцы	Январь I	Февраль II	Март III	Апрель IV
Выпуск промышленной продукции (в млн. руб.)	10,2	11,1	11,3	12,0

Чтобы найти средний месячный коэффициент и темп роста промышленной продукции, определяем помесячные коэффициенты роста  $K_p$ , которые в данном случае и являются вариантами.

$$K_p(\text{II к I}) = \frac{11,1}{10,2} = 1,08824;$$

$$K_p(\text{III к II}) = \frac{11,3}{11,1} = 1,01802;$$

$$K_p(\text{IV к III}) = \frac{12,0}{11,3} = 1,06195.$$

Из найденных трех помесячных коэффициентов роста (вариантов) определяем средний месячный коэффициент роста ( $\bar{K}_p$ ) по формуле 15. Для этого нужно найденные коэффициенты роста перемножить и из произведения извлечь корень третьей степени.

$$\bar{K}_p = \sqrt[3]{1,08824 \cdot 1,01802 \cdot 1,06195} = \sqrt[3]{1,17647} = 1,056.$$

Из разобранный примера замечаем: во-первых, что произведение трех найденных коэффициентов роста можно получить без их предварительного исчисления путем деления апрельского выпуска (12,0) на январский выпуск (10,2):

$$1,08824 \cdot 1,01802 \cdot 1,06195 = \frac{12,0}{10,2}$$

и, во-вторых, что показатель степени корня, равный трем (число коэффициентов роста), можно получить вычитанием единицы из числа приведенных в примере месяцев (четыре).

Таким образом, наиболее удобной для исчисления среднего коэффициента роста следует считать формулу:

$$\bar{K}_p = \sqrt[n-1]{\frac{y_n}{y_1}}, \quad (18a)$$

где  $n$  — число приведенных дат или периодов;

$y_n$  — последний член ряда;

$y_1$  — первый член ряда.

§ 10. Из вышеуказанных средних средняя арифметическая наиболее часто применяется. Знание свойств средней арифметической позволяет вычислять ее упрощенно.

Математические свойства средней арифметической:

*Средняя постоянной величины равна постоянной величине:*

$$\bar{c} = c. \quad (19)$$

Пример 15. Все варианты признака одинаковы.

Т а б л и ц а 14

$x$	5	5	5	5	5	5
-----	---	---	---	---	---	---

Находим среднюю:

$$\bar{x} = \frac{30}{6} = 5,$$

т. е. средняя равна этой же постоянной величине.

*Сумма отклонений от средней, соответствующим образом взвешенных, равна нулю:*

$$\Sigma (x - \bar{x})m = 0. \quad (20)$$

Пример 16. Вычислить среднюю (по колонкам 1 и 2) и убедиться в правильности формулы 20.

Т а б л и ц а 15

$x$	$m$	$xm$	$x - \bar{x}$	$(x - \bar{x})m$
1	2	3	4	5
3	1	3	-4,4	-4,4
5	3	15	-2,4	-7,2
7	9	63	-0,4	-3,6
9	5	45	+1,6	+8,0
11	2	22	+3,6	+7,2
Итого..	20	148	—	0

$$\bar{x} = \frac{\sum x m}{\sum m} = \frac{148}{20} = 7,4;$$

$$\sum (x - \bar{x}) m = -15,2 + 15,2 = 0.$$

Если из всех вариантов ( $x$ ) вычесть постоянную величину ( $x_0$ ) и из результатов вычитания, т. е. из отклонений вариантов от этой постоянной величины ( $x - x_0 = x'$ ), вычислить среднюю ( $\bar{x}'$ ), то она окажется меньше искомой средней, поэтому, чтобы получить среднюю из вариантов ( $x$ ), нужно к найденной средней ( $\bar{x}'$ ) прибавить ту же постоянную величину:

$$\bar{x} = \bar{x}' + x_0, \quad (21)$$

если  $x' = x - x_0$ .

Пример 17. Вычислить среднюю путем вычитания 1000 из всех вариантов по следующим данным (колонки 1 и 2):

Т а б л и ц а 16

$x_0 = 1000$			
$x$	$m$	$x - x_0 = x'$	$x' \cdot m$
1	2	3	4
995	2	-5	-10
998	3	-2	-6
1001	5	1	5
1004	9	4	36
1007	2	7	14
1010	1	10	10
И т о г о . .	22	-	+ 39

$$\bar{x} = \bar{x}' + x_0;$$

$$\bar{x}' = \frac{\sum x' m}{\sum m} = \frac{39}{22} = 1,77;$$

$$\bar{x} = 1,77 + 1000 = 1001,77,$$

Если все варианты ( $x$ ) уменьшить в одно и то же число раз, т. е. разделить на постоянную величину ( $k$ ), и из частных ( $\frac{x}{k} = x'$ ) вычислить среднюю, то она окажется уменьшенной в такое же число раз, а поэтому, чтобы получить среднюю из вариантов ( $x$ ), нужно найденную среднюю ( $\bar{x}'$ ) умножить на ту же постоянную величину ( $k$ ):

$$\bar{x} = \bar{x}' \cdot k, \quad (22)$$

если  $x' = \frac{x}{k}$ .

**Пример 18.** Вычислить среднюю путем деления всех вариантов на 100 по следующим данным (колонки 1 и 2):

Т а б л и ц а 17

$$k = 100$$

$x$	$m$	$\frac{x}{k} = x'$	$x'm$
1	2	3	4
1000	2	10	20
1100	3	11	33
1200	5	12	60
1300	9	13	117
1400	2	14	28
1500	1	15	15
Итого . . . .	22	—	273

$$\bar{x} = \bar{x'}k;$$

$$\bar{x'} = \frac{\Sigma x'm}{\Sigma m} = \frac{273}{22} = 12,4091;$$

$$\bar{x} = 12,4091 \cdot 100 = 1240,91.$$

При вычислении средней вместо абсолютных значений весов ( $m$ ) можно использовать относительные величины структуры (частоты), т. е. удельные веса отдельных частот в общей сумме всех частот (см. § 4), или координации (отношение частот всех вариантов к одной из частот, принятой за единицу).

$$\bar{x} = \frac{\Sigma xw}{\Sigma w} = \Sigma xw; \quad (23)$$

$$\bar{x} = \frac{\Sigma x \cdot O_{\text{координ}}}{\Sigma O_{\text{координ}}}, \quad (23a)$$

где  $w$  — частоты, т. е. доля частоты варианта в общей сумме частот;  
 $O_{\text{координ}}$  — отношение частот всех вариантов к частоте одного из вариантов.

**Пример 19.** Вычислить средний размер детали по следующим данным (колонки 1 и 2):



Таблица 18

Размер детали ( <i>x</i> )	Коли- чество деталей ( <i>m</i> )	Удельный вес дета- лей во всей сово- купности ( <i>w</i> )	Отношение частот всех вариантов к частоте 2-го вари- анта <i>O</i> <sub>координ</sub>	<i>xw</i>	<i>x · O</i> <sub>координ</sub>	<i>xm</i>
1	2	3	4	5	6	7
10	850	0,170	0,825	1,700	8,250	8 500
11	1030	0,206	1,000	2,266	11,000	11 330
12	1721	0,344	1,671	4,128	20,052	20 652
13	627	0,125	0,609	1,625	7,917	8 151
14	553	0,111	0,537	1,554	7,518	7 742
15	219	0,044	0,213	0,660	3,195	3 285
Итого	5000	1,000	4,855	11,933	57,932	59 660

Предварительно найдем относительные величины структуры и координации (колонки 3 и 4), а затем по формулам 23 и 23а вычислим средний размер детали:

$$\bar{x} = \frac{\sum xw}{\sum w} = \frac{11,933}{1,0} \approx 11,93;$$

$$\bar{x} = \frac{\sum x \cdot O_{\text{координ}}}{\sum O_{\text{координ}}} = \frac{57,932}{4,855} \approx 11,93.$$

Если теперь вычислить средний размер детали по формуле 8, то получим:

$$\bar{x} = \frac{\sum xm}{\sum m} = \frac{59\,660}{5000} \approx 11,93,$$

что согласуется с результатами, полученными ранее.

Очень часто встречаются случаи, когда абсолютные веса вариантов отсутствуют, тогда используют данное свойство.

Если в частотах (*m*) имеется общий множитель (*A*), то его можно при вычислении средней не принимать во внимание, т. е. взвешивание производить по сокращенным частотам  $\frac{m}{A} = m'$ . Численное значение средней от замены частот (*m*) на сокращенные частоты (*m'*) не изменится.

$$\bar{x} = \frac{\sum xm'}{\sum m'}. \quad (24)$$

Пример 20. Вычислить среднюю по данным нижеследующей таблицы (колонки 1 и 2), произведя взвешивание вариантов по сокращенным весам:

Таблица 19

$x$	$m$	$\frac{m}{30} = m'$	$xm'$
1	2	3	4
5	60	2	10
6	150	5	30
7	270	9	63
8	180	6	48
9	90	3	27
Итого . .	750	25	178

Вычисляем среднюю по формуле (24) предварительно сокращая веса и заполняя колонки 3 и 4.

Получаем:

$$\bar{x} = \frac{\sum xm'}{\sum m'} = \frac{178}{25} = 7,12.$$

Общая средняя равна средней из частных средних, соответствующим образом взвешенных:

$$\bar{x} = \frac{\sum \bar{x}_i m_i}{\sum m_i}, \quad (25)$$

где  $\bar{x}_i$  — частные средние, т. е. средние для отдельных групп совокупности;

$\bar{x}_1$  — средняя из вариантов первой группы;

$\bar{x}_2$  — средняя из вариантов второй группы и т. д.;

$m_i$  — частоты отдельных групп;

$m_1$  — частота первой группы;

$m_2$  — частота второй группы и т. д.

Пример 21. В трех партиях продукции, численностью 1000, 2000 и 500 единиц, вычислен средний вес детали в кг: 3,3; 3,1; 3,7. Вычислить средний вес детали во всех трех партиях деталей:

$$\bar{x} = \frac{\sum \bar{x}_i m_i}{\sum m_i} = \frac{3,3 \cdot 1000 + 3,1 \cdot 2000 + 3,7 \cdot 500}{3500} = \frac{11350}{3500} = 3,24.$$

Сумма квадратов отклонений от средней меньше суммы квадратов отклонений от произвольной величины ( $B$ ) на величину поправки  $C$ , равной произведению объема совокупности на квадрат разности между средней и данной произвольной величиной:

$$\sum (x - B)^2 - \sum (x - \bar{x})^2 = n (\bar{x} - B)^2, \quad (26)$$

или

$$\Sigma(x - B)^2 m - \Sigma(x - \bar{x})^2 m = (\bar{x} - B)^2 \cdot \Sigma m, \quad (27)$$

где  $B$  — произвольная величина.

Пример 22. По нижеследующим данным (колонки 1 и 2) убедиться в правильности соотношений, указанных в формуле (27):

Таблица 20

$B = 2$								
$x$	$m$	$xm$	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 m$	$x - 2$	$(x - 2)^2$	$(x - 2)^2 m$
1	2	3	4	5	6	7	8	9
2	1	2	-2	4	4	0	0	0
3	2	6	-1	1	2	1	1	2
4	4	16	0	0	0	2	4	16
5	2	10	1	1	2	3	9	18
6	1	6	2	4	4	4	16	16
Итого	10	40	—	—	12	—	—	52

Вычисляем колонки 3, 4, 5, 6, 7, 8 и 9 и находим:

$$\bar{x} = \frac{\Sigma xm}{\Sigma m} = \frac{40}{10} = 4.$$

По формуле (27) имеем:  $52 - 12 = (4 - 2)^2 \cdot 10 = 40$ .

§ 11. Упрощенное вычисление средней, состоящее в использовании ряда свойств средней, называется методом отсчета от условного нуля и предполагает:

- 1) вычитание из всех вариантов начала отсчета ( $x_0$ );
- 2) деление всех вариантов или отклонений вариантов от начала отсчета на общий множитель, содержащийся в них ( $k$ );
- 3) условное принятие центра интервала за значение признака у всех единиц в данном интервале.

Кроме того, в качестве весов используют сокращенные частоты ( $m'$ ) или относительные величины (структуры или координации).

Формула исчисления средней методом отсчета от условного нуля

$$\bar{x} = \bar{x}' \cdot k + x_0, \quad (28)$$

где  $x' = \frac{x - x_0}{k}$ , т. е. отклонение от начала отсчета делится на общий множитель, а исчисление средней из  $x'$  в зависимости от наличия весов производится по одной из следующих формул:

$$\bar{x}' = \frac{\Sigma x' m'}{\Sigma m'}; \quad \bar{x}' = \frac{\Sigma x' w}{100}; \quad \bar{x}' = \frac{\Sigma x' O_{\text{координ}}}{\Sigma O_{\text{координ}}}.$$

Пример 23. Вычислить средний вес зерен (на  $m^2$ ) по данным колонок 1 и 2 нижеследующей таблицы, используя метод отсчета от условного нуля:

$$x_0 = 95 \\ k = 10$$

Таблица 21

Группы по весу зерен (в г на $m^2$ )	Число метровок в процентах к итогу ( $w$ )	Центр групп $x$	$x - x_0$	$\frac{x - x_0}{k} = x'$	$x'w$
1	2	3	4	5	6
40—50	1	45	— 50	— 5	— 5
50—60	3	55	— 40	— 4	— 20
60—70	11	65	— 30	— 3	— 33
70—80	14	75	— 20	— 2	— 28
80—90	18	85	— 10	— 1	— 18
90—100	25	95	0	0	0
100—110	15	105	+ 10	1	15
110—120	7	115	+ 20	2	14
120—130	4	125	+ 30	3	12
130—140	2	135	+ 40	4	8
Итого . . . .	100	—	—	—	— 55

Используем формулу (28), предварительно заполняя колонки 3, 4, 5 и 6 табл. 21:

$$\bar{x} = \bar{x}'k + x_0;$$

$$\bar{x}' = \frac{\Sigma x'w}{\Sigma w} = - \frac{55}{100} = -0,55;$$

$$\bar{x} = -0,55 \cdot 10 + 95 = -5,5 + 95 = 89,5.$$

§ 12. Если произвести вычисление различных типов средних для одного и того же вариационного ряда, то численные значения средних будут отличаться друг от друга. При этом средние по своей величине расположатся в определенном порядке.

Наименьшей из перечисленных средних окажется средняя гармоническая, затем геометрическая и т. д. до наибольшей — средней кубической. Порядок средней при таком расположении определяется показателем степени  $z$  в формуле степенной средней и вытекает из «правила мажорантности».

Так, при  $z = -1$  получаем среднюю гармоническую

- |           |   |   |                |
|-----------|---|---|----------------|
| • $z = 0$ | • | • | геометрическую |
| • $z = 1$ | • | • | арифметическую |
| • $z = 2$ | • | • | квадратическую |
| • $z = 3$ | • | • | кубическую     |

$$\bar{x}_{\text{гарм}} < \bar{x}_{\text{геом}} < \bar{x}_{\text{арифм}} < \bar{x}_{\text{квадр}} < \bar{x}_{\text{куб}}. \quad (29)$$

Пример 24. Вычислить различные типы средних по следующим данным (колонки 1 и 2) и убедиться в правильности порядка возрастания средних:

Т а б л и ц а 22

$x$	$m$	$\frac{m}{x}$	$\lg x$	$(\lg x)m$	$xm$	$x^2$	$x^2m$	$x^3$	$x^3m$
1	2	3	4	5	6	7	8	9	10
10	1	0,10000	1,00000	1,00000	10	100	100	1 000	1 000
14	2	0,14286	1,14613	2,29226	28	196	392	2 744	5 488
18	2	0,11111	1,25527	2,51054	36	324	648	5 832	11 664
22	4	0,18182	1,34242	5,36968	88	484	1 936	10 648	42 592
26	1	0,03846	1,41497	1,41497	26	676	676	17 576	17 576
Итого	10	0,57425	—	12,58745	188	—	3 752	—	78 320

Заполняем колонки с 3 по 10 и по соответствующим формулам исчисляем средние взвешенные:

$$\bar{x}_{арм} = \frac{\Sigma m}{\Sigma \frac{m}{x}} = \frac{10}{0,57425} = 17,41;$$

$$\lg \bar{x}_{геом} = \frac{\Sigma (\lg x) m}{\Sigma m} = \frac{12,58745}{10} = 1,258745;$$

$$\bar{x}_{геом} = 18,14;$$

$$\bar{x}_{арифм} = \frac{\Sigma xm}{\Sigma m} = \frac{188}{10} = 18,8;$$

$$\bar{x}_{квадр} = \sqrt{\frac{\Sigma x^2 m}{\Sigma m}} = \sqrt{\frac{3752}{10}} = \sqrt{375,2};$$

$$\lg \bar{x}_{квадр} = \frac{1}{2} \lg 375,2 = \frac{1}{2} \cdot 2,57426 = 1,28713;$$

$$\bar{x}_{квадр} = 19,37;$$

$$\bar{x}_{куб} = \sqrt[3]{\frac{\Sigma x^3 m}{\Sigma m}} = \sqrt[3]{\frac{78320}{10}} = \sqrt[3]{7832};$$

$$\lg \bar{x}_{куб} = \frac{1}{3} \lg 7832 = \frac{1}{3} \cdot 3,89387 = 1,29796;$$

$$\bar{x}_{куб} = 19,86.$$

Порядок средних определен в соответствии с правилом:

$$17,41 < 18,14 < 18,8 < 19,37 < 19,86.$$

§ 13. В качестве характеристики вариационного ряда применяется *медиана* ( $M_e$ ), т. е. такое значение варьирующего признака, которое приходится на середину упорядоченного вариационного ряда. Если в вариационном ряду  $2m+1$  случаев, то значение признака у случая  $m+1$  будет медианным. Если в ряду четное число  $2m$  случаев, то медиана равна средней арифметической из двух срединных значений.

Формулы для исчисления медианы в случае нечетного и четного числа случаев:

$$M_e = x_{m+1}, \quad (30)$$

$$M_e = \frac{x_m + x_{m+1}}{2}. \quad (31)$$

**Пример 25.** Дано девять вариантов признака  $x$ , расположенного в возрастающем порядке:

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \dots x_{2m+1}$$

$$8, \quad 9, \quad 11, \quad 12, \quad 15, \quad 16, \quad 18, \quad 19, \quad 21.$$

Находим медиану.

Имеем нечетное число случаев:

$$2m + 1 = 9;$$

$$2m = 8;$$

$$m = 4.$$

По формуле (30) получаем:  $M_e = x_{4+1} = x_5 = 15$ .

**Пример 26.** Дано 12 вариантов признака  $x$ , расположенного в возрастающем порядке:

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad x_{2m}$$

$$8 \quad 9 \quad 11 \quad 12 \quad 15 \quad 16 \quad 18 \quad 19 \quad 21 \quad 23 \quad 24 \quad 26$$

Ищем медиану.

Имеем четное число случаев:

$$2m = 12;$$

$$m = 6.$$

По формуле 31 получаем:

$$M_e = \frac{x_6 + x_7}{2} = \frac{16 + 18}{2} = 17.$$

§ 14. При исчислении медианы интервального вариационного ряда сначала находят интервал, содержащий медиану, путем использования накопленных частот или частостей. Медианному интервалу соответствует первая из накопленных частот или частостей, превышающая половину всего объема совокупности.

Для нахождения медианы при постоянстве плотности внутри интервала, содержащего медиану, используют следующую формулу:

$$M_e = x_{M_e(\min)} + k \frac{\frac{\Sigma m}{2} - v_{M_e-1}}{m_{M_e}}, \quad (32)$$

где  $x_{M_e(\min)}$  — нижняя граница медианного интервала;

$k$  — интервальная разность;

$v_{M_e-1}$  — накопленная частота интервала, предшествующего медианному;

$m_{M_e}$  — частота медианного интервала.

**Пример 27.** По данным табл. 7 вычислить медиану.

Используем табл. 9, в которой дана колонка накопленных частот. Так как вариационный ряд содержит 200 единиц, то медиана будет 100-й единицей, входящей в интервал 49,938—49,943 (определяется из колонки 3 табл. 9 по накопленной частоте 121, первой из накопленных частот, которая превышает половину всего объема вариационного ряда). Следовательно:

$$x_{M_e(\min)} = 49,938;$$

$$k = 0,005;$$

$$\frac{\Sigma m}{2} = 100;$$

$$v_{M_e-1} = 72;$$

$$m_{M_e} = 49.$$

По формуле (32) имеем:

$$\begin{aligned} M_e &= 49,938 + 0,005 \cdot \frac{100 - 72}{49} = 49,938 + 0,005 \cdot \frac{28}{49} = \\ &= 49,938 + 0,003 = 49,941. \end{aligned}$$

§ 15. Медиана может быть определена и графически по кумуляте или оживе. Для определения медианы по кумуляте последнюю ординату, пропорциональную сумме всех частот или частостей, делят пополам. Из полученной точки восстанавливают перпендикуляр до пересечения с кумулятой. Абсцисса точки пересечения и дает значение медианы.

**Пример 28.** По графику 5 определить медиану.

Последняя ордината, как видно из графика, равна 200. Деление этой ординаты пополам дает точку  $A$  (100). Перпендикуляр из точки  $A$  до пересечения с кумулятой дает точку  $B$ . Абсцисса точки  $B$ , равная 49,941, и будет медианой.

§ 16. Характеристикой вариационного ряда является и *мода* ( $M_0$ ). Модой называется вариант, наиболее часто встречающийся в данном вариационном ряду. Для дискретного ряда мода определяется по частотам вариантов и соответствует варианту с наибольшей частотой.

В случае интервального распределения с равными интервалами модальный интервал (т. е. содержащий моду) определяется по наибольшей частоте, а при неравных интервалах — по наибольшей плотности.

Вычисление моды производится по следующей формуле:

$$M_0 = x_{M_0 (mln)} + k \frac{m_{M_0} - m_{M_0-1}}{(m_{M_0} - m_{M_0-1}) + (m_{M_0} - m_{M_0+1})}, \quad (33)$$

где  $x_{M_0 (mln)}$  — нижняя граница модального интервала;

$k$  — интервальная разность;

$m_{M_0}$  — частота модального интервала;

$m_{M_0-1}$  — частота интервала, предшествующего модальному;

$m_{M_0+1}$  — частота интервала, последующего за модальным.

**Пример 29.** По данным табл. 7 находим моду.

Наибольшая частота, равная 49 (колонка 2, табл. 7), соответствует интервалу 49,938—49,943, который и будет модальным.

Следовательно:

$$x_{M_0 (mln)} = 49,938;$$

$$k = 0,005;$$

$$m_{M_0} = 49;$$

$$m_{M_0-1} = 40;$$

$$m_{M_0+1} = 43.$$

Имеем:

$$\begin{aligned} M_0 &= 49,938 + 0,005 \cdot \frac{49 - 40}{(49 - 40) + (49 - 43)} = 49,938 + \\ &+ 0,005 \cdot \frac{9}{15} = 49,938 + 0,005 \cdot 0,6 = 49,941. \end{aligned}$$

Как видно из разобранных примера и примера 27, для данного вариационного ряда мода и медиана очень близки друг к другу.

§ 17. **Симметричные вариационные ряды.** Особенностью симметричных вариационных рядов является равенство трех характеристик: средней арифметической, моды и медианы<sup>1</sup>:

$$\bar{x} = M_0 = M_e. \quad (34)$$

<sup>1</sup> Это необходимое условие симметричности вариационного ряда, но не совсем достаточное.



Пример 30. По данным табл. 7 вычислить среднюю и сопоставить с модой и медианой, вычисленными по этим же данным в примерах 27 и 29.

Вычисляем среднюю.

Таблица 23

Центр интервала ( $x$ )	Частоты ( $m$ )	$\frac{x - x_0}{k} = x'$	$x' m$
1	2	3	4
49,9205	5	— 4	— 20
49,9255	9	— 3	— 27
49,9305	18	— 2	— 36
49,9355	40	— 1	— 40
49,9405	49	0	0
49,9455	43	1	43
49,9505	26	2	52
49,9555	8	3	24
49,9605	2	4	8
Итого . . . .	200	—	+ 5

$$x_0 = 49,9405$$

$$k = 0,005$$

$$\bar{x} = \frac{5}{200} \cdot 0,005 + 49,9405 = 49,9406.$$

Имеем:

$$M_e = 49,941 \text{ (из примера 27);}$$

$$M_o = 49,941 \text{ (из примера 29);}$$

$$\bar{x} = 49,941.$$

Найденные характеристики по своей величине близки друг к другу, что дает нам основание считать данный вариационный ряд симметричным.

§ 18. Средние величины, характеризующая вариационный ряд одним числом, не учитывают вариацию признака, между тем эта вариация существует. Для измерения вариации признака математическая статистика применяет ряд способов.

*Вариационный размах (R)* представляет собой величину неустойчивую, чрезвычайно зависящую от случайных обстоятельств, применяется в качестве приблизительной оценки вариации. В последнее время вариационный размах стал применяться в ряде отраслей промышленности при статистическом изучении качества продукции.

$$R = x_{\max} - x_{\min}, \quad (35)$$

где  $x_{\max}$  — наибольший вариант вариационного ряда;

$x_{\min}$  — наименьший вариант вариационного ряда.

*Среднее линейное отклонение* или простое среднее отклонение ( $\rho$ ) представляет собой среднюю арифметическую из абсолютных значений отклонений вариантов от средней.

Среднее линейное отклонение в зависимости от отсутствия или наличия частот вычисляют невзвешенное или взвешенное:

$$\rho = \frac{\sum |x - \bar{x}|}{n}; \quad (36)$$

$$\rho = \frac{\sum |x - \bar{x}| m}{\sum m}, \quad (37)$$

где прямые скобки, в которых заключены разности между вариантами и средней, показывают, что непосредственное суммирование и суммирование после взвешивания производится без учета знаков.

*Средний квадрат отклонения* — дисперсия ( $\alpha^2$  или  $\sigma^2$ ) наиболее часто применяется и в теории и на практике в качестве меры колеблемости признака. Если дисперсию вычисляют для всей совокупности, то обозначают  $\alpha^2$ :

$$\alpha^2 = \frac{\sum (x - \bar{x})^2}{n}; \quad (38)$$

Дисперсия (невзвешенная)

$$\alpha^2 = \frac{\sum (x - \bar{x})^2 m}{\sum m}. \quad (39)$$

Дисперсия (взвешенная)

*Среднее квадратическое отклонение* ( $\alpha$ ) представляет собой квадратный корень из дисперсии:

$$\alpha = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}; \quad (40)$$

Среднее квадратическое отклонение (невзвешенное)

$$\alpha = \sqrt{\frac{\sum (x - \bar{x})^2 m}{\sum m}}. \quad (41)$$

Среднее квадратическое отклонение (взвешенное)

Учитывая, что среднее линейное отклонение и среднее квадратическое отклонение представляют собой абсолютные величины, выраженные в тех же единицах измерения, что и варианты, для характеристики колеблемости признака используют коэффициенты вариации ( $V$ ), представляющие собой отношение среднего линей-

ного отклонения или среднего квадратического отклонения к средней, выраженное в процентах (или в долях единицы):

$$V_p = \frac{p}{\bar{x}} \cdot 100; \quad (42)$$

Коэффициент вариации  
по среднему линейному  
отклонению

$$V_a = \frac{a}{\bar{x}} \cdot 100. \quad (43)$$

Коэффициент вариации  
по среднему квадрати-  
ческому отклонению

*Показатель неровноты* применяется в текстильной промышленности в качестве меры колеблемости при изучении неровноты пряжи (по толщине, по весу и другим показателям).

Неровнота ( $H$ ) представляет собой видоизмененный показатель коэффициента вариации по среднему линейному отклонению ( $V_p$ )

$$H = \frac{2n_{<}}{n} \cdot \frac{\bar{x} - \bar{x}_{<}}{\bar{x}}; \quad (44)$$

Показатель неровноты  
невзвешенный

$$H = \frac{2 \cdot \Sigma m_{<}}{\Sigma m} \cdot \frac{\bar{x} - \bar{x}_{<}}{\bar{x}}, \quad (45)$$

Показатель неровноты  
взвешенный

где  $\bar{x}$  — общая средняя;

$n_{<}$  — количество вариантов меньших, чем общая средняя;

$n$  — объем вариационного ряда;

$\bar{x}_{<}$  — средняя из вариантов, меньших, чем общая средняя;

$\Sigma m_{<}$  — сумма частот вариантов, меньших общей средней;

$\Sigma m$  — сумма частот всех вариантов.

**Пример 31.** По нижеследующим данным о крепости одиночной нити (в г) вычисляются показатели вариации признака: вариационный размах, показатель неровноты, коэффициенты вариации по среднему линейному отклонению и среднему квадратическому отклонению:

## Для вычисления

Крепость одиночной нити (в з)	Центр интер- вала (x)	Частота (m)	средней (x)		Показателя неровности (H)		Коэффициента вариации по среднему линейному отклонению ( $V_p$ )			Коэффициента вариации по среднему квадратическому отклонению ( $V_a$ )	
			xm		m <		x - $\bar{x}$	x - $\bar{x}$   m	$(x - \bar{x})^2$	$(x - \bar{x})^2 m$	
1	2	3	4	5	6	7	8	9	10		
120—130	125	4	500	4	500	— 67,16	— 268,64	4510,4656	18041,8624		
130—140	135	16	2160	16	2160	— 57,16	— 914,56	3267,2656	52276,2496		
140—150	145	15	2175	15	2175	— 47,16	— 707,40	2224,0656	33360,9840		
150—160	155	31	4805	31	4805	— 37,16	— 1151,96	1380,8656	42806,8336		
160—170	165	51	8415	51	8415	— 27,16	— 1385,16	737,6656	37620,9456		
170—180	175	58	10150	58	10150	— 17,16	— 995,28	294,4656	17079,0048		
180—190	185	65	12025	65	12025	— 7,16	— 465,40	51,2656	3332,2640		
190—200	195	69	13455	15	2866	—	—	8,0656	556,5264		
200—210	205	50	10250	—	—	—	—	164,8656	8243,2800		
210—220	215	57	12255	—	—	—	—	521,6656	29734,9392		
220—230	225	33	7425	—	—	—	—	1078,4656	35589,3648		
230—240	235	21	4935	—	—	—	—	1835,2656	38540,5776		
240—250	245	16	3920	—	—	—	—	2792,0656	44673,0496		
250—260	255	10	2550	—	—	—	—	3948,8656	39488,6560		
260—270	265	4	1060	—	—	—	—	5305,6656	21222,6624		
Итого . .	—	500	96080	255	43096	—	5888,40	—	—	422567,2000	

Находим среднюю:  $\bar{x} = \frac{96080}{500} = 192,16$ .

Вычисляем  $R$  по формуле (35):

$$R = x_{\max} - x_{\min} = 270 - 120 = 150.$$

Находим  $H$  по формуле 45. Интервал 190—200 расчленим на две части: 190—192,16 и 192,16—200.

Аналогично поступаем с частотами: так как вся частота данного интервала равна 69, то, предполагая равномерное распределение признака внутри интервала, получим, что на величину, равную единице интервала, приходится 6,9 единицы частот (абсолютная плотность); на новый интервал (190—192,16), в котором интервальная разность равна 2,16, придется  $6,9 \cdot 2,16 = 14,9$ . Для простоты возьмем 15. Суммируя частоты вариантов, меньших общей средней, получим 255 (см. колонку 5 табл. 24). Суммируя произведения  $x_{<} \cdot m_{<}$  и учитывая произведение центра нового интервала 191,08 на вычисленную частоту 15, получим 43 096 (итог колонки 6 табл. 24). Используя полученные данные, находим:

$$\bar{x}_{<} = \frac{\Sigma x_{<} m_{<}}{\Sigma m_{<}} = \frac{43\,096}{255} = 169,004;$$

$$H = \frac{2\Sigma m_{<}}{\Sigma m} \cdot \frac{\bar{x} - \bar{x}_{<}}{\bar{x}} = \frac{2 \cdot 255}{500} \cdot \frac{192,16 - 169,004}{192,16} = 0,1229.$$

Вычисляем  $\rho$  и  $V_{\rho}$  по формулам 37 и 42.

Учитывая одно из свойств средней, а именно, что сумма отклонений от средней, соответствующим образом взвешенных, равна нулю (§ 10), практически поступают следующим образом. В колонке 7 табл. 24, несмотря на знак прямых скобок, указывающих на абсолютную величину отклонений, для отрицательных отклонений от средней знак минус оставляют и ведут вычисление только до перемены знака на плюс. Взвешивают отрицательные отклонения от средней (колонка 8 табл. 24) и так как сумма взвешенных положительных отклонений от средней должна быть равна сумме взвешенных отрицательных отклонений от средней, для определения общей суммы взвешенных отклонений найденную сумму удваивают.

Получаем:

$$\Sigma |x - \bar{x}| m = 5888,40 \cdot 2 = 11776,8;$$

$$\rho = \frac{11776,8}{500} = 23,5536;$$

$$V_{\rho} = \frac{23,5536}{192,16} = 0,1226.$$

Вычисляем  $V_\alpha$  по формулам (41) и (43):

$$\alpha = \sqrt{\frac{\sum (x - \bar{x})^2 m}{\sum m}} = \sqrt{\frac{422567,2}{500}} = \sqrt{845,1344} = 29,07;$$

$$V_\alpha = \frac{\alpha}{\bar{x}} = \frac{29,07}{192,16} = 0,152.$$

§ 19. Между средним квадратическим отклонением ( $\alpha$ ) и средним линейным отклонением ( $\rho$ ) существует определенное соотношение (такое же соотношение между  $V_\alpha$  и  $V_\rho$ ). По правилу мажорантности  $\alpha$  всегда больше  $\rho$ .

Если объем совокупности достаточно большой и распределение признака в вариационном ряде близко к нормальному (см. раздел IV), то связь между  $\alpha$  и  $\rho$  определяется по формуле:

$$\alpha = \rho \cdot \sqrt{\frac{\pi}{2}} \approx \rho \cdot 1,25. \quad (46)$$

Отклонения от 1,25 в обе стороны зависят от близости распределения к нормальному.

**Пример 32.** По данным примера 31 найти соотношение между  $\alpha$  и  $\rho$ .

Имеем:

$$\alpha = 29,07;$$

$$\rho = 23,554;$$

$$\frac{\alpha}{\rho} \approx 1,23.$$

Отношение ненамного отличается от теоретического (1,25), что косвенно свидетельствует о близости взятого распределения к нормальному.

§ 20. Средний квадрат отклонения — дисперсия — обладает рядом свойств, которые позволяют упростить вычисления:

*Дисперсия постоянной величины равна нулю:*

$$\alpha_c^2 = 0, \quad (47)$$

где  $c$  — постоянная величина;

$\alpha_c^2$  — дисперсия постоянной величины.

*Если все значения вариантов признака  $x$  уменьшить на постоянную величину, то дисперсия не изменится.* Это позволяет вычислять дисперсию вариационного ряда путем вычитания из вариантов начала отсчета ( $x_0$ ):

$$\alpha_{(x-x_0)}^2 = \alpha_x^2, \quad (48)$$

где  $\alpha_x^2$  — дисперсия вариантов  $x$ ;

$\alpha_{(x-x_0)}^2$  — дисперсия вариантов, уменьшенных вычитанием  $x_0$ .

Дисперсия алгебраической суммы независимых случайных величин (см. раздел II, § 23) равна сумме их дисперсий:

$$\alpha_{(x+y+\dots+w)}^2 = \alpha_x^2 + \alpha_y^2 + \dots + \alpha_w^2. \quad (49)$$

Если все значения вариантов  $x$  уменьшить в « $k$ » раз, то дисперсия уменьшится в « $k^2$ » раз:

$$\alpha_x^2 = \alpha_{\frac{x}{k}}^2 \cdot k^2, \quad (50)$$

где  $\alpha_{\frac{x}{k}}^2$  — дисперсия из частных, полученных в результате деления вариантов на постоянную величину  $k$ .

Дисперсия суммы двух случайных величин, связанных корреляционной зависимостью, равна сумме их дисперсий плюс удвоенное произведение среднеквадратических отклонений на коэффициент корреляции между этими случайными величинами.

$$\alpha_{(x+y)}^2 = \alpha_x^2 + \alpha_y^2 + 2\alpha_x \cdot \alpha_y \cdot R_{y/x}, \quad (51)$$

где  $R_{y/x}$  — коэффициент корреляции между величинами  $y$  и  $x$  (см. раздел V).

Пример 33. Даны случайные величины  $y$  и  $x$ , связанные корреляционной зависимостью так, что  $R_{y/x} = 0,5$ .

Т а б л и ц а 25

$x$	1	3	5
$y$	3	5	4

Найти дисперсию суммы этих случайных величин (для простоты дан пример без взвешивания):

Находим средние:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{9}{3} = 3;$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{12}{3} = 4.$$

Определяем дисперсии по формуле (38):

$$\alpha_x^2 = \frac{(-2)^2 + (2)^2}{3} = \frac{8}{3};$$

$$\alpha_y^2 = \frac{(-1)^2 + (1)^2}{3} = \frac{2}{3}.$$

Используя формулу 51, имеем:

$$\alpha_{(x+y)}^2 = \frac{2}{3} + \frac{8}{3} + 2 \sqrt{\frac{2}{3} \cdot \frac{8}{3}} \cdot 0,5 = \frac{2}{3} + \frac{8}{3} + 2 \frac{4}{3} \cdot 0,5 = \\ = \frac{14}{3} = 4 \frac{2}{3}.$$

Убедимся, что если  $x + y = z$ , то получаем три значения  $z$ : 4, 8 и 9.

Находим: среднюю

$$\bar{z} = \frac{21}{3} = 7;$$

дисперсию

$$\alpha_z^2 = \frac{(-3)^2 + (1)^2 + (2)^2}{3} = \frac{14}{3} = 4 \frac{2}{3},$$

т. е.

$$\alpha_{(x+y)}^2 = 4 \frac{2}{3}.$$

Результаты совпадают.

*Дисперсия суммы двух случайных величин, связанных линейной функциональной зависимостью (см. раздел V), равна сумме их дисперсий плюс или минус удвоенное произведение средних квадратических отклонений:*

$$\alpha_{(x+y)}^2 = \alpha_x^2 + \alpha_y^2 \pm 2\alpha_x \cdot \alpha_y. \quad (52)$$

В данной формуле знак «+» или «—» определяется характером связи. При прямолинейной связи  $y$  с  $x$   $y = a_0 + a_1x$  знак зависит от величины  $a_1$ . Если  $a_1 > 0$ , то в формуле берем знак «+», если  $a_1 < 0$ , то берем знак «—».

**Пример 34.** Даны две случайные величины  $x$  и  $y$ , связанные уравнением  $y = 2 + 3x$ .

Т а б л и ц а 26

$x$	1	3	5
$y$	5	11	17

Найти дисперсию суммы этих случайных величин.



Находим средние:

$$\bar{x} = \frac{9}{3} = 3; \quad \bar{y} = \frac{33}{3} = 11.$$

Определяем дисперсии по формуле:

$$\alpha_x^2 = \frac{(-2)^2 + (2)^2}{3} = \frac{8}{3};$$

$$\alpha_y^2 = \frac{(-6)^2 + (6)^2}{3} = \frac{72}{3} = 24.$$

Используем формулу 52. В данном случае берем знак „+“:

$$\alpha_{(x+y)}^2 = \frac{8}{3} + 24 + 2\sqrt{\frac{8}{3} \cdot 24} = 42 \frac{2}{3}.$$

Убеждаемся, что если  $x + y = z$ , получаем три значения  $z$ :

$$6 \quad 14 \quad 22.$$

Находим: среднюю

$$\bar{z} = \frac{42}{3} = 14;$$

дисперсию

$$\alpha_z^2 = \frac{(-8)^2 + (8)^2}{3} = \frac{128}{3} = 42 \frac{2}{3},$$

т. е.

$$\alpha_{(x+y)}^2 = 42 \frac{2}{3}.$$

Учитывая свойства дисперсии, расчет ее производят по формуле, упрощающей вычисления:

$$\alpha^2 = \frac{\sum \left( \frac{x - x_0}{k} \right)^2 m}{\Sigma m} \cdot k^2 - (\bar{x} - x_0)^2. \quad (53)$$

**Пример 35.** По данным таблицы 24 (2 и 3 колонки) рассчитать дисперсию, используя формулу, упрощающую вычисления.

Располагаем данные в таблицу.

Т а б л и ц а 27

Центр интервала ( $x$ )	Частоты ( $m$ )	$x - x_0$	$\frac{x - x_0}{k} = x'$	$x' m$	$(x')^2$	$(x')^2 m$
1	2	3	4	5	6	7
125	4	-70	-7	-28	49	196
135	16	-60	-6	-96	36	576
145	15	-50	-5	-75	25	375
155	31	-40	-4	-124	16	496
165	51	-30	-3	-153	9	459
175	58	-20	-2	-116	4	232
185	65	-10	-1	-65	1	65
195	69	0	0	0	0	0
205	50	10	1	50	1	50
215	57	20	2	114	4	228
225	33	30	3	99	9	297
235	21	40	4	84	16	336
245	16	50	5	80	25	400
255	10	60	6	60	36	360
265	4	70	7	28	49	196
Итого...	500	—	—	-657 + 515 - 142	—	4266

$x_0 = 195$ ;  $k = 10$ .

$$\bar{x} = \bar{x}' \cdot k + x_0 = -\frac{142}{500} \cdot 10 + 195 = 192,16;$$

$$\begin{aligned} \alpha^2 &= \frac{\Sigma(x')^2 m}{\Sigma m} \cdot k^2 - (\bar{x} - x_0)^2 = \frac{4266}{500} \cdot 10^2 - (192,16 - 195)^2 = \\ &= 853,2 - (-2,84)^2 = 853,2 - 8,0656 = 845,1344. \end{aligned}$$

Величина дисперсии совпадает с полученной в примере 31, но в данном случае вычисления в значительной мере упрощены.

Из формулы 53 вытекает еще одна формула дисперсии.

При  $x_0 = 0$  и  $k = 1$ :

$$\alpha^2 = \bar{x}^2 - \bar{x}^2, \quad (54)$$

где  $\bar{x}^2$  — средняя из квадратов вариантов

$$\bar{x}^2 = \frac{\Sigma x^2 m}{\Sigma m};$$

$\bar{x}^2$  — квадрат средней

$$\bar{x}^2 = \left( \frac{\Sigma x m}{\Sigma m} \right)^2.$$

Так, если вычислить дисперсию по данным табл. 24 примера 31, пользуясь формулой 54, то получим:

Таблица 28

Центр интервала ( $x$ )	Частоты ( $m$ )	$xm$	$x^2$	$x^2m$
125	4	500	15 625	62 500
135	16	2 160	18 225	291 600
145	15	2 175	21 025	315 375
155	31	4 805	24 025	744 775
165	51	8 415	27 225	1 388 475
175	58	10 150	30 625	1 776 250
185	65	12 025	34 225	2 224 625
195	69	13 455	38 025	2 623 725
205	50	10 250	42 025	2 101 250
215	57	12 255	46 225	2 634 825
225	33	7 425	50 625	1 670 625
235	21	4 935	55 225	1 159 725
245	16	3 920	60 025	960 400
255	10	2 550	65 025	650 250
265	4	1 060	70 225	280 900
Итого . . .	500	96 080	—	18 885 300

$$\overline{x^2} = \frac{\Sigma x^2 m}{\Sigma m} = \frac{18\,885\,300}{500} = 37\,770,6;$$

$$\overline{x} = \frac{\Sigma x m}{\Sigma m} = \frac{96\,080}{500} = 192,16;$$

$$\overline{x^2} = 36\,925,4656;$$

$$s^2 = 37\,770,6 - 36\,925,4656 = 845,1344.$$

Результат совпадает с дисперсией, полученной по этим данным в примере 31.

§ 21. Для каждой группы вариантов вариационного ряда может быть исчислена наряду с частной средней и дисперсия, которая называется частной дисперсией  $\alpha_i^2$

$$\alpha_i^2 = \frac{\Sigma (x - \overline{x}_i)^2 m_i}{\Sigma m_i}, \quad (55)$$

где  $\overline{x}_i$  — частная средняя  $i$ -й группы;

$\alpha_i^2$  — частная дисперсия  $i$ -й группы.

§ 22. Из частных дисперсий может быть найдена средняя, которая обозначается  $\overline{\alpha_i^2}$

$$\overline{\alpha_i^2} = \frac{\Sigma \alpha_i^2 m_i}{\Sigma m_i}. \quad (56)$$

§ 23. Частные средние по группам ( $\bar{x}_i$ ) могут не совпадать с общей средней ( $\bar{x}$ ). Мерой колеблемости частных средних вокруг общей средней является межгрупповая дисперсия ( $\delta^2$  — дельта квадрат в среднем).

$$\bar{\delta}^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 m_i}{\sum m_i}. \quad (57)$$

§ 24. Между общей дисперсией, средней из частных дисперсий и межгрупповой дисперсией, существует связь:

$$\alpha^2 = \bar{\alpha}_i^2 + \bar{\delta}^2. \quad (58)$$

Это — правило сложения вариации.

Пример 36 (на § 21, 22, 23 и 24). Используя данные табл. 24 и расчленив вариационный ряд на две группы (1-я группа с интервала 120—130 до интервала 190—200 включительно, а 2-я группа с интервала 200—210 до интервала 260—270), исчислить частные дисперсии, среднюю из частных дисперсий и межгрупповую дисперсию.

Начинаем расчет с 1-й группы.

Таблица 29

Центр интервала ( $x_1$ )	Частоты ( $m$ )	$x_1 - x_0$	$\frac{x_1 - x_0}{k} = x_1'$	$x_1' m$	$(x_1')^2$	$(x_1')^2 m$
1	2	3	4	5	6	7
125	4	— 70	— 7	— 28	49	196
135	16	— 60	— 6	— 96	36	576
145	15	— 50	— 5	— 75	25	375
155	31	— 40	— 4	— 124	16	496
165	51	— 30	— 3	— 153	9	459
175	58	— 20	— 2	— 116	4	232
185	65	— 10	— 1	— 65	1	65
195	69	0	0	0	0	0
Итого .	309	—	—	— 657	—	2399

$$x_0 = 195; k = 10.$$

$$\bar{x}_1 = \bar{x}_1' k + x_0 = -\frac{657}{309} \cdot 10 + 195 = -21,26 + 195 = 173,74;$$

$$\alpha_1^2 = \frac{\sum (x_1')^2 m}{\sum m} \cdot k^2 - (\bar{x}_1 - x_0)^2 = \frac{2399}{309} \cdot 100 - (-21,26)^2 = 776,3754 - 451,9876 = 324,3878.$$

Для 2-й группы получаем аналогично:

$$\bar{x}_2 = \frac{515}{191} \cdot 10 + 195 = 26,96 + 195 = 221,96;$$

$$\alpha_2^2 = \frac{1867}{191} \cdot 100 - (26,96)^2 = 977,4869 - 726,8416 = 250,6453.$$

Вычисляем среднюю из частных дисперсий:

$$\begin{aligned} \bar{\alpha}_i^2 &= \frac{\sum \alpha_i^2 m_i}{\sum m_i} = \frac{324,3878 \cdot 309 + 250,6453 \cdot 191}{309 + 191} = \frac{100235,83 + 47873,25}{500} = \\ &= \frac{148108,08}{500} = 296,22. \end{aligned}$$

Находим межгрупповую дисперсию, используя общую среднюю для всего вариационного ряда, найденную в примере 31 и равную 192,16.

$$\begin{aligned} \bar{\delta}^2 &= \frac{\sum (\bar{x}_i - \bar{x})^2 \cdot m_i}{\sum m_i} = \frac{(173,74 - 192,16)^2 \cdot 309 + (221,96 - 192,16)^2 \cdot 191}{309 + 191} = \\ &= \frac{104842,59 + 169615,64}{500} = \frac{274458,23}{500} = 548,916. \end{aligned}$$

Для получения общей дисперсии используем правило сложения вариации:

$$\alpha^2 = \bar{\alpha}_i^2 + \bar{\delta}^2 = 296,216 + 548,916 \approx 845,13.$$

Результат совпадает с дисперсией, вычисленной в примере 31 по табл. 24 без расчленения вариационного ряда на две группы.

§ 25. Наряду с количественной вариацией признака может иметь место и качественная вариация. Если имеются два взаимно исключающих друг друга варианта, то вариация признака называется альтернативной.

Так, например, рассмотрение выпущенной продукции с точки зрения ее качества, как пригодной к дальнейшему использованию или непригодной, дает альтернативный признак. Обозначая наличие признака — 1, а отсутствие — 0 и долю вариантов, обладающих данным признаком, —  $p$ , а долю вариантов, не обладающих им, —  $q$ , получаем формулу дисперсии альтернативного признака:

$$\alpha^2 = pq. \quad (59)$$

Из дисперсии альтернативного признака извлечением корня находим среднее квадратическое отклонение:

$$\alpha = \sqrt{pq}. \quad (60)$$

**Пример 37.** Совокупность состоит из 10 000 электрических лампочек, включающих в свой состав 20 бракованных. Найти дисперсию признака.

Находим долю брака и долю доброкачественных лампочек:

$$p = \frac{20}{10\,000} = 0,002;$$

$$q = \frac{9\,880}{10\,000} = 0,998.$$

По формуле (59) имеем:

$$\alpha^2 = pq = 0,998 \cdot 0,002 = 0,001996$$

и по формуле (60) имеем:

$$\alpha = \sqrt{0,001996} = 0,0447.$$

§ 26. Как уже было показано в § 13, медиана — это вариант, который делит упорядоченный вариационный ряд на две равные по объему группы. Аналогично можно в каждой группе найти вариант, делящий ее на две подгруппы. Такие варианты называются квантилями.

Различают нижний и верхний квантили. Иногда вычисляют и децили, т. е. такие варианты, которые делят вариационный ряд на 10 равных по объему групп.

При отношении объема двух подгрупп как  $\frac{1}{4}$  к  $\frac{3}{4}$  имеем нижний квантиль ( $Q_1$ ), при отношении объемов подгрупп  $\frac{3}{4}$  к  $\frac{1}{4}$  — верхний квантиль ( $Q_3$ ), а при отношениях объемов групп  $\frac{1}{10}$  к  $\frac{9}{10}$ ;  $\frac{2}{10}$  к  $\frac{8}{10}$  и т. д. первый, второй и т. д. децили.

Формулы для расчетов в интервальном ряду:  
нижнего квантиля

$$Q_1 = x_{Q_1(\min)} + k \frac{\frac{\Sigma m}{4} - v_{Q_1-1}}{m_{Q_1}}; \quad (61)$$

верхнего квартиля

$$Q_3 = x_{Q_3(\min)} + k \cdot \frac{\frac{3}{4} \Sigma m - v_{Q_3-1}}{m_{Q_3}}, \quad (62)$$

где  $x_{Q_1(\min)}$  — минимальная граница интервала, содержащего нижний квартиль (определяется по накопленным частотам);

$x_{Q_3(\min)}$  — то же, для верхнего квартиля;

$k$  — интервальная разность;

$v_{Q_1-1}$  — накопленная частота интервала, предшествующего интервалу, содержащему нижний квартиль;

$v_{Q_3-1}$  — то же, для верхнего квартиля;

$m_{Q_1}$  — частота интервала, содержащего нижний квартиль;

$m_{Q_3}$  — то же, для верхнего квартиля.

Вычисление децилей ничем принципиально не отличается от вычисления медианы и квартилей. Так, первый и второй децили могут быть вычислены по формулам:

$$D_1 = x_{D_1(\min)} + k \frac{\frac{1}{10} \Sigma m - v_{D_1-1}}{m_{D_1}};$$

$$D_2 = x_{D_2(\min)} + k \frac{\frac{2}{10} \Sigma m - v_{D_2-1}}{m_{D_2}} \text{ и т. д.}$$

**Пример 38.** По данным табл. 7 вычислить нижний и верхний квартили (рекомендуется предварительно вспомнить вычисление медианы).

Используем табл. 9, в которой дана колонка накопленных частот. Расчет для нижнего квартиля — по формуле (61). Из итога колонки 2 табл. 9 видно, что численность совокупности для этого ряда равна 200 единицам. Следовательно, нижний квартиль отвечает 50-й единице. По колонке накопленных частот (3) видим, что нижний квартиль содержится в интервале 49,933—49,938, потому что первая из накопленных частот, превышающих 50, это накопленная частота данного интервала.

Следовательно:

$$x_{Q_1(\min)} = 49,933;$$

$$v_{Q_1-1} = 32;$$

$$m_{Q_1} = 40;$$

$$k = 0,005.$$

Имеем:

$$\begin{aligned} Q_1 &= 49,933 + 0,005 \cdot \frac{50 - 32}{40} = 49,933 + 0,005 \cdot \frac{18}{40} = \\ &= 49,933 + 0,005 \cdot 0,45 = 49,933 + 0,002 = 49,935. \end{aligned}$$

Расчет для верхнего квартиля по формуле (62).

Верхний квартиль отвечает 150-й единице и содержится в интервале 49,943—49,948 (так как первая из накопленных частот, превышающая 150, равна 164 и соответствует данному интервалу).

Имеем:

$$\begin{aligned} Q_3 &= 49,943 + 0,005 \cdot \frac{150 - 121}{43} = 49,943 + 0,005 \cdot \frac{29}{43} = \\ &= 49,943 + 0,005 \cdot 0,67 = 49,943 + 0,003 = 49,946. \end{aligned}$$

§ 27. В качестве характеристики колеблемости вариационного ряда применяется относительный показатель, подобный коэффициенту вариации, но для вычисления которого используются нижний и верхний квартили и медиана. Этот показатель называют квартилем ( $Q_0$ ) без добавления слова нижний или верхний:

$$Q_0 = \frac{Q_3 - Q_1}{2M_e} \cdot 100, \quad (63)$$

где  $\frac{Q_3 - Q_1}{2}$  — половина межквартильного расстояния.

Пример 39. По результатам исчисления медианы, а также нижнего и верхнего квартилей по табл. 7 (см. примеры 27 и 38) найти квартиль.

Имеем:

$$M_e = 49,941;$$

$$Q_1 = 49,935;$$

$$Q_3 = 49,946;$$

$$Q_0 = \frac{49,946 - 49,935}{2 \cdot 49,941} \cdot 100 = \frac{0,011}{99,882} \cdot 100 \cong 1,1\%.$$

Интересно, что исчисление коэффициента вариации по данным табл. 7 дает:

$$V_a = \frac{\alpha}{\bar{x}} \cdot 100 = \frac{0,8}{49,94} \cdot 100 = 1,6\%.$$

§ 28. Характеристиками вариационного ряда являются моменты распределения.

Средняя из  $k$ -х степеней отклонений вариантов  $x$  от некоторой постоянной величины  $A$  называется моментом  $k$ -го порядка:

$$M_k = \overline{(x - A)^k}. \quad (64)$$

При исчислении средней в качестве весов могут быть использованы частоты, частости или вероятности (см. раздел II). При



использовании в качестве весов частот или частостей моменты называются эмпирическими, а при использовании вероятностей — теоретическими.

Порядок момента определяется величиной  $k$ . В соответствии с определением, данным выше, эмпирический момент  $k$ -го порядка находится как отношение суммы произведений  $k$ -х степеней отклонений вариантов от постоянной величины  $A$  на частоты к сумме частот:

$$M_k = \frac{\sum (x - A)^k m}{\sum m}. \quad (65)$$

В зависимости от выбора постоянной величины  $A$  различают моменты:

Если постоянная величина  $A$  равна нулю ( $A=0$ ), то моменты называются начальными. Приводим их формулу:

$$M_k = \overline{(x - 0)^k} = \frac{\sum x^k m}{\sum m}. \quad (66)$$

Тогда:

при  $k=0$  получаем

$$M_0 = \frac{\sum x^0 m}{\sum m} = 1; \quad (66a)$$

Начальный момент  
нулевого порядка

при  $k=1$  получаем

$$M_1 = \frac{\sum x m}{\sum m} = \bar{x}; \quad (66б)$$

Начальный момент  
первого порядка

при  $k=2$  получаем

$$M_2 = \frac{\sum x^2 m}{\sum m} = \overline{x^2}; \quad (66в)$$

Начальный момент  
второго порядка

при  $k=3$  получаем

$$M_3 = \frac{\sum x^3 m}{\sum m} = \overline{x^3}; \quad (66г)$$

Начальный момент  
третьего порядка

при  $k=4$  получаем

$$M_4 = \frac{\sum x^4 m}{\sum m} = \overline{x^4}. \quad (66д)$$

Начальный момент  
четвертого порядка

Пример 40. Вычислить начальные моменты первых четырех порядков, если варианты  $x'$  имеют отрицательные и положительные значения.

Располагаем все расчеты в таблицу:

Т а б л и ц а 30

$x'$	Частоты $m$	$x'm$	$(x')^2 m$	$(x')^3 m$	$(x')^4 m$
-2	1	-2	4	-8	16
-1	3	-3	3	-3	3
0	4	0	0	0	0
1	1	1	1	1	1
2	1	2	4	8	16
Итого..	10	-2	12	-2	36

Имеем:  $M_0 = 1$ ;

$$M_1 = -\frac{2}{10} = -0,2$$

$$M_2 = \frac{12}{10} = 1,2;$$

$$M_3 = -\frac{2}{10} = -0,2;$$

$$M_4 = \frac{36}{10} = 3,6.$$

Если постоянная величина равна некоторой произвольной величине  $x_0$  (начало отсчета), то моменты называются начальными относительно  $x_0$  и обозначаются  $M'_k$ :

$$M'_k = \overline{(x - x_0)^k} = \frac{\Sigma (x - x_0)^k m}{\Sigma m}. \quad (67)$$

Тогда получаем начальные моменты относительно  $x_0$ :  
при  $k = 0$

$$M'_0 = \frac{\Sigma (x - x_0)^0 m}{\Sigma m} = 1; \quad (67a)$$

нулевого порядка

при  $k = 1$

$$M'_1 = \frac{\Sigma (x - x_0) m}{\Sigma m} = \bar{x} - x_0; \quad (67b)$$

первого порядка

при  $k = 2$

$$M'_2 = \frac{\Sigma (x - x_0)^2 m}{\Sigma m}; \quad (67b)$$

второго порядка

при  $k=3$

$$M'_3 = \frac{\Sigma (x - x_0)^3 m}{\Sigma m}; \quad (67\Gamma)$$

третьего порядка

при  $k=4$

$$M'_4 = \frac{\Sigma (x - x_0)^4 m}{\Sigma m} \quad (67Д)$$

четвертого порядка

и т. д.

Из формулы 67б  $\bar{x} = x_0 + M'_1$ , т. е. средняя арифметическая равна началу отсчета плюс начальный момент первого порядка относительно начала отсчета; далее, если отклонения  $x$  от  $x_0$  имеют общий множитель  $c$ , то на него можно разделить отклонения, а по окончании вычислений полученный момент умножить на этот множитель в соответствующей степени, т. е.:

$$M'_k = \frac{\Sigma \left( \frac{x - x_0}{c} \right)^k m}{\Sigma m} c^k. \quad (68)$$

Отсюда следует, что  $\bar{x} = x_0 + M'_1 c$  (сравните с вычислением средней методом отсчета от условного нуля), формула 28.

**Пример 41.** Вычислить начальные моменты относительно  $x_0$  первых четырех порядков по следующим данным (колонки 1 и 2 табл. 31). Располагаем все расчеты в таблице:

Таблица 31

$x$	$m$	$x - x_0$	$\frac{x - x_0}{c} = x'$
1	2	3	4
16	1	-4	-2
18	3	-2	-1
20	4	0	0
22	1	2	1
24	1	4	2
Итого . . .	10	—	—

Возьмем в качестве  $x_0$  вариант, равный 20, тогда получим колонку 3, разделим все отклонения от начала отсчета на общий множитель  $c$ , равный 2, и получим значения  $x'$  в колонке 4, для которых вычисления начальных моментов произведены в примере 40.

Для получения  $M'_k$  нужно найденные в примере 40 начальные моменты умножить на  $c$ , равное 2, в соответствующей степени.

Получаем:

$$M'_0 = 1 \cdot 2^0 = 1;$$

$$M'_1 = -0,2 \cdot 2^1 = -0,4;$$

$$M'_2 = 1,2 \cdot 2^2 = 4,8;$$

$$M'_3 = -0,2 \cdot 2^3 = -1,6;$$

$$M'_4 = 3,6 \cdot 2^4 = 57,6.$$

Поэтому практически при нахождении начальных моментов относительно  $x_0$  поступают следующим образом:

из всех вариантов вычитают начало отсчета и находят отклонения  $(x - x_0)$ ;

делят эти отклонения на общий множитель  $\frac{x - x_0}{c} = x'$ ;

находят начальные моменты для  $x'(M_k)$ ;

путем умножения найденных начальных моментов на  $c^k$  получают начальные моменты относительно  $x_0(M'_k)$ .

Если за постоянную величину взять среднюю  $(\bar{x})$ , то моменты называются центральными и обозначаются  $\mu_k$ :

$$\mu_k = \overline{(x - \bar{x})^k} = \frac{\sum (x - \bar{x})^k m}{\sum m}. \quad (69)$$

Тогда:  
при  $k=0$

$$\mu_0 = \overline{(x - \bar{x})^0} = \frac{\sum (x - \bar{x})^0 m}{\sum m} = 1; \quad (69a)$$

Центральный момент нулевого  
порядка равен единице

при  $k=1$

$$\mu_1 = \overline{(x - \bar{x})} = \frac{\sum (x - \bar{x}) m}{\sum m} = 0; \quad (69б)$$

Центральный момент первого  
порядка равен нулю

при  $k=2$

$$\mu_2 = \overline{(x - \bar{x})^2} = \frac{\sum (x - \bar{x})^2 m}{\sum m}; \quad (69в)$$

Центральный момент второго  
порядка равен дисперсии и  
служит мерой колеблемости  
признака

при  $k=3$

$$\mu_3 = \overline{(x - \bar{x})^3} = \frac{\sum (x - \bar{x})^3 m}{\sum m}; \quad (69г)$$

Центральный момент третьего  
порядка служит мерой асиммет-  
рии распределения признака.  
Если распределение симметрич-  
но, то  $\mu_3 = 0$ .

при  $k = 4$

$$\mu_4 = \overline{(x - \bar{x})^4} = \frac{\Sigma (x - \bar{x})^4 m}{\Sigma m}. \quad (69д)$$

Центральный момент  
четвертого порядка

Пример 42. Вычислим центральные моменты первых четырех порядков по данным табл. 31 (колонки 1, 2).

Располагаем все расчеты в таблицу:

Т а б л и ц а 32

$x$	$m$	$xm$	$x - \bar{x}$	$(x - \bar{x})m$	$(x - \bar{x})^2 m$	$(x - \bar{x})^3 m$	$(x - \bar{x})^4 m$
1	2	3	4	5	6	7	8
16	1	16	-3,6	-3,6	12,96	-46,656	167,9616
18	3	54	-1,6	-4,8	7,68	-12,288	19,6608
20	4	80	0,4	1,6	0,64	0,256	0,1024
22	1	22	2,4	2,4	5,76	13,824	33,1776
24	1	24	4,4	4,4	19,36	85,184	374,8096
Итого . .	10	196	—	-8,4 +8,4 0	46,40	-58,944 +99,264 +40,320	595,7120

Получаем:

$$\bar{x} = \frac{\Sigma xm}{\Sigma m} = \frac{196}{10} = 19,6;$$

$$\mu_0 = 1;$$

$$\mu_1 = 0;$$

$$\mu_2 = \frac{\Sigma (x - \bar{x})^2 m}{\Sigma m} = \frac{46,4}{10} = 4,64 = \sigma^2 \text{ (дисперсия);}$$

$$\mu_3 = \frac{\Sigma (x - \bar{x})^3 m}{\Sigma m} = \frac{40,32}{10} = 4,032;$$

$$\mu_4 = \frac{\Sigma (x - \bar{x})^4 m}{\Sigma m} = \frac{595,712}{10} = 59,5712.$$

§ 29. Имеется связь между начальными моментами первых четырех порядков вариантов  $x'$  и начальным моментом 4-го порядка вариантов  $x''$  для случая, когда варианты  $x'$  меньше вариантов  $x''$  на единицу:

$$M_4^* = M_0 + 4M_1 + 6M_2 + 4M_3 + M_4, \quad (70)$$

где  $M_4^*$  — четвертый начальный момент вариантов  $x''$ .

В правой части формулы все начальные моменты (от нулевого порядка до четвертого порядка) — вариантов  $x'$ .

Практически данная формула используется для проверки вычисления начальных моментов первых четырех порядков вариантов  $x'$  путем вычисления начального момента 4-го порядка новых вариантов  $x''$ , полученных прибавлением к вариантам  $x$  единицы.

Если исчисления  $M_4^*$  непосредственно из данных по формуле 66 и по формуле связи между моментами 70 дают тождественные результаты, то это свидетельствует о правильности всех начальных моментов первых четырех порядков, вычисленных для вариантов  $x'$ .

**Пример 43.** Проверяется правильность начальных моментов первых четырех порядков, вычисленных в примере 40.

Располагаем все расчеты в таблицу:

Т а б л и ц а 33

$x'$	$m$	$x' + 1 = x''$	$(x'')^4$	$(x'')^4 m$
1	2	3	4	5
— 2	1	— 1	1	1
— 1	3	0	0	0
0	4	1	1	4
1	1	2	16	16
2	1	3	81	81
Итого . . .	10	—	—	102

В колонке 3 записываем новые варианты  $x''$  путем прибавления к старым вариантам  $x'$  единицы.

Получаем по формуле (66):

$$M_4^* = \frac{\sum (x'')^4 m}{\sum m} = \frac{102}{10} = 10,2.$$

Для расчетов  $M_4^*$  по формуле 70 привлекаем данные из примера 40:

$$M_0 = 1;$$

$$M_1 = -0,2;$$

$$M_2 = 1,2;$$

$$M_3 = -0,2;$$

$$M_4 = 3,6.$$

Получаем:

$$M_4^* = 1 + 4(-0,2) + 6 \cdot 1,2 + 4(-0,2) + 3,6 = 11,8 - 1,6 = 10,2.$$

Результаты совпадают, следовательно, начальные моменты первых четырех порядков в примере 40 вычислены правильно.

§ 30. Вычисление центральных моментов, привлекаемых в качестве характеристик вариационного ряда по формуле 69, с точки зрения вычислительной техники довольно громоздко. Поэтому сначала вычисляют начальные моменты, затем начальные относительно  $x_0$ , а для нахождения центральных моментов используют формулу перехода от начальных моментов, вычисленных относительно  $x_0$ , к центральным:

$$\mu_k = M'_k - C_k^1 \cdot M'_{k-1} \cdot M'_1 + C_k^2 M'_{k-2} \cdot (M'_1)^2 - C_k^3 M'_{k-3} \cdot (M'_1)^3 + \dots + (-M'_1)^k \quad (71)$$

Знаки в формуле чередуются.

$C_k^1, C_k^2$  и т. д. обозначают числа сочетаний из:  $k$  по 1;  $k$  по 2;  $k$  по 3 и т. д.

Полагая в этой формуле  $k$  равным 0, 1, 2, 3, 4 и т. д., можем получить:

$$\mu_0 = M'_0 = 1; \quad (71a)$$

$$\mu_1 = M'_1 - M'_1 = 0; \quad (71б)$$

$$\mu_2 = M'_2 - (M'_1)^2; \quad (71в)$$

$$\mu_3 = M'_3 - 3M'_2 M'_1 + 2(M'_1)^3; \quad (71г)$$

$$\mu_4 = M'_4 - 4M'_3 M'_1 + 6M'_2 (M'_1)^2 - 3(M'_1)^4 \quad (71д)$$

и т. д.

Для вычисления центральных моментов высших порядков по найденным центральным моментам низших порядков и начальным моментам относительно  $x_0$  подставляем в формулу третьего центрального момента 71г величину  $M'_2$ , найденную из формулы второго центрального момента 71в, и получаем:

$$M'_2 = \mu_2 + (M'_1)^2;$$

$$\begin{aligned} \mu_3 &= M'_3 - 3[\mu_2 + (M'_1)^2] M'_1 + 2(M'_1)^3 = M'_3 - 3\mu_2 \cdot M'_1 - \\ &- 3(M'_1)^2 + 2(M'_1)^3 = M'_3 - 3\mu_2 M'_1 - (M'_1)^3, \end{aligned}$$

т. е.

$$\mu_3 = M'_3 - 3\mu_2 \cdot M'_1 - (M'_1)^3 \quad (72)$$

и далее аналогично:

$$\mu_4 = M'_4 - 4\mu_3 M'_1 - 6\mu_2 \cdot (M'_1)^2 - (M'_1)^4. \quad (73)$$

Пример 44. Используя данные примера 41, где вычислены начальные моменты относительно  $x_0=20$ , ищем центральные моменты первых четырех порядков по формулам 71, 72 и 73 и сверяем полученные результаты с вычислением центральных моментов в примере 42.

Из примера 41 имеем:

$$M'_0 = 1; \quad M'_3 = -1,6;$$

$$M'_1 = -0,4; \quad M'_4 = 57,6.$$

$$M'_2 = 4,8;$$

По формулам 71а, б, в, г и т. д. получаем:

$$\mu_0 = 1;$$

$$\mu_1 = 0;$$

$$\mu_2 = 4,8 - (-0,4)^2 = 4,8 - 0,16 = 4,64;$$

$$\mu_3 = -1,6 - 3(4,8) \cdot (-0,4) + 2(-0,4)^3 = -1,6 + 5,76 - 0,128 = 4,032;$$

$$\mu_4 = 57,6 - 4(-1,6) \cdot (-0,4) + 6 \cdot 4,8(-0,4)^2 - 3(-0,4)^4 = 57,6 - 2,56 + 4,608 - 0,0768 = 59,5712.$$

Сравнивая центральные моменты первых четырех порядков, вычисленные по формулам 71а, б, в, г, д, с центральными моментами, вычисленными в примере 42 непосредственно по формуле 69, убеждаемся в сравнительной простоте исчисления центральных моментов по приведенным формулам.

Аналогично и использование формул 72, 73.

Так, вычисление третьего центрального момента по второму центральному моменту и начальным относительно  $x_0$  моментам:

$$\mu_3 = -1,6 - 3 \cdot 4,64 \cdot (-0,4) - (-0,4)^3 = -1,6 + 5,568 + 0,064 = 4,032.$$

Вычисление четвертого центрального момента по третьему и второму центральным моментам и начальным относительно  $x_0$  моментам:

$$\mu_4 = 57,6 - 4 \cdot 4,032(-0,4) - 6 \cdot 4,64(-0,4)^2 - (-0,4)^4 = 57,6 + 6,4512 - 4,4544 - 0,0256 = 59,5712.$$

§ 31. Исчисление центральных моментов сводится к: нахождению начальных моментов ( $M_k$ ) и их проверке; нахождению начальных моментов относительно произвольно выбранного начала отсчета ( $M'_k$ );

использованию формул перехода от начальных моментов относительно произвольно выбранного начала отсчета к центральным моментам ( $\mu_k$ ).



Пример 45. По нижеследующим данным (колонки 1, 2 и 3) вычислить центральные моменты первых четырех порядков:

Таблица 34

Интервальное значение при- знака	Частоты ( <i>m</i> )	Центральное значение ин- тервала ( <i>x</i> )	Отклонения от $x_0$ , делен- ное на $c$ $\frac{x - x_0}{c} = x'$	$x' m$	$(x')^2 m$	$(x')^3 m$	$(x')^4 m$	Для проверки на- чальных моментов первых четырех порядков	
								$x' + 1 = x''$	$(x'')^4 m$
1	2	3	4	5	6	7	8	9	10
22—25	1	23,5	—7	—7	49	—343	2 401	—6	1 296
25—28	2	26,5	—6	—12	72	—432	2 592	—5	1 250
28—31	1	29,5	—5	—5	25	—125	625	—4	256
31—34	4	32,5	—4	—16	64	—256	1 024	—3	324
34—37	4	35,5	—3	—12	36	—108	324	—2	64
37—40	6	38,5	—2	—12	24	—48	96	—1	6
40—43	12	41,5	—1	—12	12	—12	12	0	0
43—46	17	44,5	0	0	0	0	0	1	17
46—49	15	47,5	1	15	15	15	15	2	240
49—52	17	50,5	2	34	68	136	272	3	1 377
52—55	8	53,5	3	24	72	216	648	4	2 048
55—58	3	56,5	4	12	48	192	768	5	1 875
Итого	90	—	—	$\frac{-76}{+85} + 9$	485	$\frac{-1\,324}{+529} - 765$	8 777	—	8 753

$$x_0 = 44,5; c = 3.$$

Начнем с вычисления начальных моментов. Для этого выбираем  $x_0 = 44,5$ , находим отклонения вариантов  $x$  от  $x_0$  и делим эти отклонения на общий множитель  $c = 3$ .

Все действия производим в табл. 34 и получаем колонку  $x'$  (колонка 4). Далее, произведя расчет по формулам бба, б, в, г, д, находим начальные моменты, для этого рассчитываем колонки 5, 6, 7 и 8.

При этом следует заметить, что для простоты расчета числа колонки 5 получают перемножением чисел, расположенных в колонках 2 и 4, числа колонки 6 получают перемножением чисел колонок 2 и 5, числа колонки 7 — перемножением чисел колонок 2 и 6 и т. д.

$$M_0 = 1; \quad M_1 = -\frac{765}{90} = -8,5;$$

$$M_1 = \frac{9}{90} = 0,1; \quad M_4 = \frac{8777}{90} = 97,52222.$$

$$M_2 = \frac{485}{90} = 5,38889;$$

Проверяем вычисление начальных моментов первых четырех порядков. Для этого вычисляем колонки 9 и 10.

Числа колонки 9 получают прибавлением к числам четвертой колонки единицы.

Числа колонки 10 (а можно и 8) получают, используя таблицу, имеющую следующий вид:

Варианты $x'$ или $x''$	1	2	3	4	5	6	и т. д.
Частоты ( $m$ )							
1	1	16	81	256	625	1 296	
2	2	32	162	512	1 250	2 592	
и т. д.							

В первой колонке таблицы указаны частоты ( $m$ ) от 1 до 50, а в верхнем заголовке — числа  $x'$  или  $x''$ . Произведения  $(x')^4 m$  или  $(x'')^4 m$  находятся на пересечении соответствующей строки и столбца.

Так, если  $m = 1$ , а  $x'' = -6$ , то  $(x'')^4 m = 1296$ ;

•  $m = 2$ , а  $x'' = -5$ , то  $(x'')^4 m = 1250$

и т. д. (см. приложение VII).

Используя формулу 70, получаем:

$$\begin{aligned} M_4^* &= 1 + 4 \cdot 0,1 + 6 \cdot 5,38889 + 4(-8,5) + 97,52222 = \\ &= 1 + 0,4 + 32,33333 - 34 + 97,52222 = 97,25555. \end{aligned}$$

Исчисляя  $M_4^*$  непосредственно, получаем

$$M_4^* = \frac{8753}{90} = 97,25555.$$

Результаты вычисления  $M_4^*$  по двум формулам совпадают, что свидетельствует о правильности расчета первых четырех начальных моментов.

Находим начальные моменты первых четырех порядков относительно выбранного начала отсчета 44,5 по формуле 68.

Получаем:

$$M_0' = 1 \cdot 3^0 = 1; \quad M_3' = -8,5 \cdot 3^3 = -229,5;$$

$$M_1' = 0,1 \cdot 3 = 0,3; \quad M_4' = 97,52222 \cdot 3^4 = 7899,29982.$$

$$M_2' = 5,38889 \cdot 3^2 = 48,5.$$

Находим центральные моменты, используя формулы 71а, б, в, г, д. Получаем:

$$\mu_0 = 1;$$

$$\mu_1 = 0;$$

$$\mu_2 = 48,5 - 0,09 = 48,41;$$

$$\mu_3 = -229,5 - 3 \cdot 48,5(0,3) + 2(0,3)^3 = -229,5 - 43,65 + 0,054 = -273,096;$$

$$\mu_4 = 7899,29982 - 4(-229,5) \cdot (0,3) + 6(48,5) \cdot (0,3)^3 - 3(0,3)^4 = 7899,29982 + 275,4 + 26,19 - 0,0243 = 8200,86552.$$

§ 32. Вторым центральным моментом равен дисперсии, т. е.  $\mu_2 = \alpha^2$ . Если среднее квадратическое отклонение ( $\alpha$ ), т. е. корень из дисперсии, иначе говоря — корень из второго центрального момента ( $\sqrt{\mu_2}$ ), принять за стандарт, то отношение центрального момента  $k$ -го порядка к стандарту в  $k$ -й степени будет называться нормированным моментом и обозначаться  $r_k$ :

$$r_k = \frac{\mu_k}{(\sqrt{\mu_2})^k} = \frac{\mu_k}{\alpha^k}; \quad (74)$$

$$r_1 = 0; \quad (74a)$$

$$r_2 = 1; \quad (74б)$$

$$r_3 = \frac{\mu_3}{\alpha^3}; \quad (74в)$$

$$r_4 = \frac{\mu_4}{\alpha^4}. \quad (74г)$$

Пример 46. По найденным в примере 45 центральным моментам найти нормированные моменты первых четырех порядков.

Из примера 45 имеем:

$$\mu_0 = 1;$$

$$\mu_1 = 0;$$

$$\alpha^2 = \mu_2 = 48,41;$$

$$\mu_3 = -273,096;$$

$$\mu_4 = 8200,8655.$$

Находим сначала стандарт:

$$\sqrt{\mu_2} = \alpha = \sqrt{48,41} = 6,958,$$

а затем нормированные моменты:

$$r_1 = 0;$$

$$r_2 = 1;$$

$$r_3 = \frac{-27,096}{6,958^3} = -\frac{27,096}{336,837} \approx -0,08;$$

$$r_4 = \frac{8200,8655}{6,958^4} = \frac{8200,8655}{48,41^2} \approx \frac{8200,8655}{2343,528} = 3,5.$$

§ 33. Нормированные моменты используются при изучении вариационных рядов. Третий нормированный момент  $r_3 = \frac{\mu_3}{\sigma^3}$  называется мерой асимметрии или косости вариационного ряда. Знак перед  $r_3$  указывает на направление асимметрии ряда. Если  $r_3 > 0$ , то вариационный ряд будет с левосторонней скошенностью, а если  $r_3 < 0$  — с правосторонней скошенностью. В симметричном ряде  $r_3 = 0$ .

Четвертый нормированный момент  $r_4 = \frac{\mu_4}{\sigma^4}$  называется мерой крутости.

Если  $r_4 > 3$  — распределение высоковершинное, если  $r_4 < 3$  — распределение низкововершинное, если  $r_4 = 3$  — распределение, близкое к нормальному (см. раздел IV).

По результатам вычисления нормированных моментов в примере 46 видно, что  $r_3$  отрицателен ( $-0,08$ ), т. е. распределение с незначительной правосторонней скошенностью, а  $r_4$  больше 3. Это указывает на высоковершинность данного распределения. В целом данное распределение не очень сильно отличается от нормального.

§ 34. В качестве характеристик биномиального распределения признака (бернуллиево распределение) (см. раздел II) используют начальные, начальные относительно  $x_0$  и центральные моменты:

$$M_k = \Sigma m^k \frac{n!}{m!(n-m)!} p^m q^{n-m}, \quad (75)$$

где  $m$  — число появлений события;

$p$  — вероятность события;  $q = 1 - p$ ;

$n$  — число испытаний.

*Начальные моменты:*

$$M_1 = np; \quad (75a)$$

$$M_2 = np + n(n-1) \cdot p^2. \quad (75b)$$

*Центральные моменты* получаются из начальных по общему правилу (формула 71).

Например,

$$\begin{aligned} \mu_2 &= M_2 - (M_1)^2 = np + n(n-1)p^2 - n^2p^2 = \\ &= np + n^2p^2 - np^2 - n^2p^2 = np(1-p) = npq. \end{aligned}$$

§ 35. Биномиальный ряд может быть представлен следующими характеристиками:

$$\bar{x} = np; \quad (76)$$

$$\alpha = \sqrt{npq}; \quad (77)$$

$$r_3 = \frac{q-p}{\alpha}; \quad (78)$$

Нормированный момент  
3-го порядка

$$r_4 = 3 - \frac{6pq-1}{\alpha^2} \quad (79)$$

Нормированный момент  
4-го порядка

Пример 47. Дан вариационный ряд, в котором числа появлений события будут рассматриваться как варианты ( $x$ ), а вероятности данного исхода при  $n=5$ ;  $p = \frac{3}{4}$  и  $q = \frac{1}{4}$ , рассчитанные по формуле разложения вероятностей в биномиальную сторону, как частоты ( $m$ ) (см. раздел II). Вычислить характеристики данного ряда по формулам 76, 77, 78 и 79.

Таблица 35

$x$	$m$
1	2
0	$\frac{1}{1024}$
1	$\frac{15}{1024}$
2	$\frac{90}{1024}$
3	$\frac{270}{1024}$
4	$\frac{405}{1024}$
5	$\frac{243}{1024}$
Итого . . .	1,00

$$\bar{x} = np = 5 \cdot \frac{3}{4} = \frac{15}{4} = 3,75;$$

$$\alpha^2 = npq = 5 \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{15}{16} = 0,9375;$$

$$\alpha = \sqrt{0,9375} = 0,9682;$$

$$r_3 = \frac{q-p}{\alpha} = \frac{\frac{1}{4} - \frac{3}{4}}{0,9682} = -\frac{1}{1,9364} \approx -0,516;$$

$$r_4 = 3 - \frac{6pq-1}{\alpha^2} = 3 - \frac{6 \cdot \frac{3}{4} \cdot \frac{1}{4} - 1}{0,9375} = 3 - \frac{1}{7,5} = 2,877.$$

§ 36. В качестве показателя отклонения вариационного ряда от симметрии применяется чрезвычайно простой эмпирический коэффициент асимметрии ( $K_a$ ):

$$K_a = \frac{\bar{x} - M_0}{\alpha}. \quad (80)$$

Если  $K_a > 0$ , то скошенность левосторонняя;  
 если  $K_a < 0$ , то скошенность правосторонняя;  
 если  $K_a = 0$ , то вариационный ряд симметричен.

---

## РАЗДЕЛ II

### ОСНОВНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ

При статистической обработке опытных данных используют определения и правила, установленные теорией вероятностей.

Теория вероятностей изучает закономерности, которым подчинены массовые случайные явления.

§ 1. Измерение вероятностей основывается на подсчете шансов<sup>1</sup>.

Вероятность того, что событие «А» произойдет, измеряется отношением числа шансов, благоприятствующих данному событию «А» ( $M$ ), к общему числу шансов как благоприятствующих, так и неблагоприятствующих ( $N$ ).

Вероятность обозначают:

$$p_A = \frac{M}{N} . \quad (1)$$

**Пример 1.** Из партии в 200 деталей, среди которых 4 бракованных и 196 доброкачественных (соответствующих стандарту), производится выборка одной детали. Как велика вероятность того, что отобранная деталь окажется бракованной? Находим вероятность события «А» в соответствии с приведенным выше определением вероятности. Из возможных 200 случаев исхода данного отбора 4 случая благоприятствуют событию «А», вероятность которого мы ищем. Поэтому:

$$p_A = \frac{M}{N} = \frac{4}{200} = 0,02.$$

§ 2. Вероятность того, что «А» не произойдет, исчисляется в соответствии с определением вероятности, как отношение числа шансов, не благоприятствующих событию «А», к общему числу шансов.

---

<sup>1</sup> Шансами называются равновозможные, друг друга исключающие случаи, один из которых должен непременно осуществиться.

Если  $N$  — общее число шансов и событию «А» благоприятствует  $M$ , то не благоприятствует  $N - M$  шансов, — тогда:

$$q_A = \frac{N - M}{N}. \quad (2)$$

**Пример 2.** По данным примера 1 найти вероятность того, что отобранная деталь не окажется бракованной, т. е. окажется доброкачественной.

Число случаев, благоприятствующих тому, что отобранная деталь не окажется бракованной, равно  $200 - 4 = 196$ .

$$q_A = \frac{N - M}{N} = \frac{196}{200} = 0,98.$$

**§ 3.** Сумма вероятностей двух противоположных событий, т. е. таких, из которых может произойти либо одно, либо другое, равна единице:

$$p_A + q_A = 1. \quad (3)$$

Отсюда вытекает, что, зная вероятность события «А», можно вычитанием ее из единицы определить вероятность того, что «А» не произойдет:

$$q_A = 1 - p_A. \quad (3a)$$

**Пример 3.** Используя данные примера 1, по вероятности события «А» найти вероятность противоположного события.

Имеем:

$$p_A = 0,02;$$

тогда:

$$q_A = 1 - 0,02 = 0,98.$$

**§ 4.** Если все  $N$  случаев благоприятствуют событию «А» ( $M = N$ ), то вероятность события «А» равна единице. Такое событие называется достоверным.

$$p_A = \frac{N}{N} = 1. \quad (4)$$

**Пример 4.** Если в партии из 200 деталей все 200 доброкачественны, то вероятность того, что отобранная наудачу деталь окажется доброкачественной, есть событие достоверное.

$$p_A = \frac{200}{200} = 1.$$



§ 5. Если число случаев, благоприятствующих событию «А», равно нулю ( $M=0$ ), то вероятность события «А» равна нулю. Такое событие называется невозможным.

$$p_A = \frac{0}{N} = 0. \quad (5)$$

Пример 5. Если в партии из 200 деталей нет ни одной бракованной, то вероятность того, что отобранная деталь окажется бракованной, равна нулю. Это событие невозможно.

$$p_A = \frac{0}{200} = 0.$$

§ 6. Вероятности событий находятся в пределах от нуля до единицы:

$$0 \leq p \leq 1. \quad (6)$$

§ 7. При решении различных задач по исчислению вероятностей событий используются основные теоремы теории вероятностей.

*Если события «А» и «Б» несовместимы<sup>1</sup>, то вероятность того, что произойдет — или событие «А», или событие «Б» ( $P_{A \text{ или } B}$ ), равна сумме вероятностей наступления каждого события.*

$$P_{(A \text{ или } B)} = p_A + p_B, \quad (7)$$

где  $p_A$  — вероятность события „А“;

$p_B$  — вероятность события „Б“.

Пример 6. Партия из 200 деталей состоит из:

150 деталей I сорта

30 „ II „

16 „ III „

4 „ брака

Как велика вероятность того, что отобранная наудачу деталь будет либо I сорта, либо II сорта?

Так как при данных условиях событие «А» (I сорт) и событие «Б» (II сорт) несовместимы (одно из них исключает возможность другого), то можно применить теорему сложения вероятностей.

<sup>1</sup> Несовместимые — такие события, которые исключают друг друга.

Имеем:

$$p_A = \frac{150}{200} = 0,75;$$

$$p_B = \frac{30}{200} = 0,15.$$

Следовательно:

$$P_{(A \text{ или } B)} = p_A + p_B = 0,75 + 0,15 = 0,9$$

или

$$P_{(\text{I сорт или II сорт})} = p_{\text{I сорт}} + p_{\text{II сорт}} = 0,9.$$

**Следствие теоремы сложения вероятностей.** Сумма вероятностей единственно возможных и несовместимых событий равна единице.

$$p_{A_1} + p_{A_2} + p_{A_3} + \dots + p_{A_k} = 1. \quad (8)$$

Такие несовместимые события, сумма вероятностей которых равна единице, составляют полную группу событий. Если полная группа состоит из двух событий, то эти события называются противоположными.

**Пример 7.** По данным примера 6 найти вероятность того, что отобранная наудачу деталь окажется либо I сорта, либо II сорта, либо III сорта, либо бракованной.

Имеем:

$$p_{A_1} = 0,75;$$

$$p_{A_2} = 0,15;$$

$$p_{A_3} = 0,08;$$

$$p_{A_4} = 0,02,$$

отсюда:

$$p_{A_1} + p_{A_2} + p_{A_3} + p_{A_4} = 0,75 + 0,15 + 0,08 + 0,02 = 1.$$

**Первая теорема умножения вероятностей.** Если имеется сложное событие, состоящее из совпадения двух независимых<sup>1</sup> друг от друга событий («А», «В»), то вероятность того, что произойдет сложное событие (включающее в себя и событие «А» и событие «В»), равна произведению вероятностей каждого события:

$$P_{(A \text{ и } B)} = p_A \cdot p_B. \quad (9)$$

<sup>1</sup> Независимыми называют такие события, вероятность каждого из которых не зависит от того, произошло или не произошло другое событие.

Указанная теорема может быть распространена и на сложные события, состоящие из нескольких (более двух) независимых событий.

Пример 8. Имеются две партии деталей:

1-я партия		2-я партия	
Таблица 1		Таблица 2	
Сорт	Количество деталей	Сорт	Количество деталей
I	8	I	12
II	3	II	2
III	1	III	1
Брак	1	Брак	2
Итого	13	Итого	17

Из каждой партии наудачу отбирается по одной детали. Как велика вероятность того, что отобранные две детали окажутся I сорта?

Данное событие является сложным и состоит из двух простых: отбор детали I сорта из 1-й партии (событие «А») и отбор детали I сорта из 2-й партии (событие «Б»). Оба простых события «А» и «Б» независимы, так как вероятность отбора детали I сорта из 2-й партии не зависит от того, окажется ли отобранная деталь I сорта из 1-й партии. Следовательно, можно применить 1-ю теорему умножения вероятностей.

Имеем:

$$p_A = \frac{8}{13}; \quad p_B = \frac{12}{17}.$$

$$P_{(A \text{ и } B)} = p_A \cdot p_B = \frac{8}{13} \cdot \frac{12}{17} = \frac{96}{221} \approx 0,4344.$$

Пример 9. Сохраняя условие примера 8, произведем наудачу выборку двух деталей только из 1-й партии с возвратом 1-й отобранной детали после ее извлечения. Как велика вероятность того, что 1-я деталь окажется I сорта, а 2-я — II сорта?

В данном случае все условия применения 1-й теоремы умножения вероятностей налицо.

Имеем:

$$p_{\text{I сорт}} = \frac{8}{13}; \quad p_{\text{II сорт}} = \frac{3}{13}; \quad P_{(\text{I и II сорт})} = \frac{8}{13} \cdot \frac{3}{13} = \frac{24}{169} = 0,142.$$

Следствие 1-й теоремы умножения вероятностей. Вероятность повторения события «А»  $k$  раз при  $k$  независимых испытаниях, в которых вероятность его остается одинаковой, равна вероятности события «А» при каждом испытании, возведенной в степень  $k$ :

$$P_k = p_A^k. \quad (10)$$

Пример 10. По данным примера 8 находим вероятность того, что при отборе наудачу трех деталей из 1-й партии (по одной с возвратом) все три детали окажутся I сорта. Вероятность детали I сорта при одном испытании

$$p_A = \frac{8}{13}.$$

Находим вероятность того, что все три детали окажутся I сорта:

$$P_3 = \left(\frac{8}{13}\right)^3 = \frac{512}{2197} = 0,2326.$$

2-я теорема умножения вероятностей. Вероятность сложного события, состоящего из двух простых событий, равна произведению вероятностей одного события на условную вероятность другого:

$$P_{(A \text{ и } B)} = p_A \cdot P_{B(A)}. \quad (11)$$

Условная вероятность события «Б» исчисляется в предположении, что событие «А» состоялось.

Пример 11. По данным примера 6 определить вероятность того, что при отборе двух деталей (без возврата) первая отобранная деталь окажется I сорта, а вторая — II сорта.

Вероятность того, что первая деталь окажется I сорта:

$$p_A = \frac{150}{200} = 0,75.$$

Предполагая, что данное событие осуществилось, вероятность второго события:

$$P_{B(A)} = \frac{30}{199} = 0,1507.$$

Вычисляем вероятность сложного события, состоящего из двух простых зависимых событий:

$$P_{(A \text{ и } B)} = p_A \cdot P_{B(A)} = 0,75 \cdot 0,1507 \approx 0,113.$$

## § 8. Полная вероятность события «А».

Если некоторое событие «А» может произойти только тогда, когда имеет место какая-нибудь из нескольких гипотез, то *полная* вероятность события «А» равна сумме произведений вероятностей гипотез на вероятности события «А», при условии осуществления каждой данной гипотезы:

$$P_A = p_{z(1)} \cdot p_{A(1)} + p_{z(2)} \cdot p_{A(2)} + \dots + p_{z(k)} \cdot p_{A(k)} = \sum p_z p_A, \quad (12)$$

где  $P_A$  — полная вероятность события „А“;

$p_{z(1)}, p_{z(2)}, \dots, p_{z(k)}$  — вероятность гипотезы 1-й, 2-й и т. д. до  $k$ -й;

$p_{A(1)}, p_{A(2)}, \dots, p_{A(k)}$  — вероятность события „А“ по первой гипотезе, по второй гипотезе и т. д. до  $k$ -й.

При подсчете полной вероятности мы пользуемся двумя теоремами: сложения и умножения вероятностей.

**Пример 12.** Имеется 6 ящиков, одинаковых по внешнему виду и по числу содержащихся в них деталей, но с разным количеством деталей I сорта.

Т а б л и ц а 3

Номер ящика	Количество деталей	
	Всего	I сорта
1	10	8
2	10	8
3	10	8
4	10	6
5	10	6
6	10	5

Определить полную вероятность того, что при выборке наудачу одной детали она окажется первосортной.

Вероятность предположения, что отобранная деталь окажется взятой из ящиков, содержащих 8 деталей I сорта (ящики № 1, 2 и 3 — гипотеза первая), может быть исчислена как отношение числа ящиков, благоприятствующих этому исходу, к общему числу ящиков:

$$p_{z(1)} = \frac{3}{6}.$$

Вероятность второй гипотезы, т. е. того, что деталь окажется взятой из ящиков, содержащих 6 деталей I сорта (ящики № 4 и 5):

$$p_{z(2)} = \frac{2}{6}.$$

Вероятность третьей гипотезы (ящик № 6):

$$p_{z(3)} = \frac{1}{6}.$$

Определим вероятность событий «А», т. е. отбора детали I сорта по гипотезам:

по первой гипотезе вероятность отбора детали I сорта:

$$p_{A(1)} = \frac{8}{10};$$

по второй гипотезе:

$$p_{A(2)} = \frac{6}{10};$$

по третьей гипотезе:

$$p_{A(3)} = \frac{5}{10}.$$

Определяем полную вероятность события „А“:

$$\begin{aligned} P_A = \sum p_z p_{A} &= \frac{3}{6} \cdot \frac{8}{10} + \frac{2}{6} \cdot \frac{6}{10} + \frac{1}{6} \cdot \frac{5}{10} = \\ &= \frac{24}{60} + \frac{12}{60} + \frac{5}{60} = \frac{41}{60}. \end{aligned}$$

Здесь применена сначала теорема сложения вероятностей, затем теорема умножения вероятностей и, наконец, еще раз теорема сложения вероятностей. Когда мы определяли вероятность гипотезы, то считали, что вероятность первой гипотезы равна  $\frac{3}{6}$ , но ведь эта первая гипотеза осуществляется, если деталь окажется взятой из ящика № 1 ( $p = \frac{1}{6}$ ) или № 2 ( $p = \frac{1}{6}$ ), или же № 3 ( $p = \frac{1}{6}$ ). Следовательно, вероятность первых двух гипотез мы и получили, складывая вероятности каждого из событий, составляющих эту гипотезу:

$$p_{z(1)} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6};$$

$$p_{z(2)} = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}.$$

Вероятность того, что произойдет первое сложное событие, состоящее в том, что отобранная деталь окажется взятой по первой гипотезе из ящиков, содержащих 8 деталей I сорта, и будет первосортной, мы определяли по теореме умножения вероятностей:

$$p_{2(1)} \cdot p_{A(1)} = \frac{3}{6} \cdot \frac{8}{10} = \frac{24}{60}.$$

Вероятность второго сложного события, состоящего в том, что отобранная деталь окажется взятой по второй гипотезе из ящиков, содержащих 6 деталей I сорта, и будет первосортной:

$$p_{2(2)} \cdot p_{A(2)} = \frac{2}{6} \cdot \frac{6}{10} = \frac{12}{60}.$$

Также получена и вероятность третьего сложного события:

$$p_{2(3)} \cdot p_{A(3)} = \frac{1}{6} \cdot \frac{5}{10} = \frac{5}{60}.$$

Для того чтобы отобранная деталь оказалась I сорта, должно иметь место либо первое сложное событие, либо второе, либо третье. Поэтому исчисленные вероятности этих событий мы складывали, применяя снова теорему сложения вероятностей:

$$p_A = \frac{24}{60} + \frac{12}{60} + \frac{5}{60} = \frac{41}{60}.$$

### § 9. Теорема деления вероятностей.

Из теоремы умножения вероятностей видно, что условная вероятность события «Б» в предположении, что событие «А» имело место, может быть определена как отношение вероятности совместного совершения обоих событий «А» и «Б» к вероятности события «А»:

$$p_{B(A)} = \frac{p_{(A \text{ и } B)}}{p_A}. \quad (13)$$

§ 10. Теорема Байеса дает возможность судить о величине вероятности какого-либо предположения после опыта, давшего определенный результат, и формулируется следующим образом: вероятность гипотезы  $i$  после испытания, приведшего к осуществлению события «А», равна произведению вероятности этой гипотезы на условную вероятность события «А» при этой гипотезе, деленную на сумму произведений вероятностей всех гипотез на условные вероятности события «А» при этих гипотезах.

тезы до испытания на вероятность события по этой гипотезе, деленному на полную вероятность события «А», т. е. на сумму таких произведений для всех гипотез:

$$P_z(i) = \frac{P_z(i) \cdot P_A(i)}{\sum P_z(i) \cdot P_A(i)}, \quad (14)$$

где  $P_z(i)$  — вероятность  $i$  гипотезы после испытания;

$P_z(i)$  — вероятность  $i$  гипотезы до испытания;

$P_A(i)$  — вероятность события „А“ в предположении, что данная гипотеза осуществилась.

Применяя теорему Байеса, удобно пользоваться следующей схемой:

Т а б л и ц а 4

№ п/п	Гипотезы $z(i)$	Вероят- ность гипотезы до испытания $P_z(i)$	Вероят- ность собы- тия „А“ по каждой гипотезе $P_A(i)$	Произ- ведения $P_z(i)P_A(i)$	Вероятность гипотезы после испытания $P_z(i)$
1	2	3	4	5 (3 · 4)	6
1	$z(1)$	$P_z(1)$	$P_A(1)$	$P_z(1)P_A(1)$	$P_z(1)$
2	$z(2)$	$P_z(2)$	$P_A(2)$	$P_z(2)P_A(2)$	$P_z(2)$
3	$z(3)$	$P_z(3)$	$P_A(3)$	$P_z(3)P_A(3)$	$P_z(3)$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$\kappa$	$z(\kappa)$	$P_z(\kappa)$	$P_A(\kappa)$	$P_z(\kappa)P_A(\kappa)$	$P_z(\kappa)$
—	—	$\sum P_z(i) = 1$	—	$\sum P_z(i)P_A(i)$	$\sum P_z(i) = 1$

Пример 13. По данным примера 12 рассуждаем: пусть мы выбрали наудачу один ящик и из него отобрали одну деталь, которая оказалась I сорта. Требуется определить вероятность того, что: 1) отобранный ящик содержал 8 деталей I сорта (т. е. оказался ящиком № 1, 2 или 3); 2) отобранный ящик содержал 6 деталей I сорта (т. е. оказался ящиком № 4 и 5) и 3) отобранный ящик содержал 5 деталей I сорта (т. е. оказался ящиком № 6).

Для исчисления искомых вероятностей расположим все исходные и расчетные показатели в таблицу.



Таблица 5

№ п п	Гипотезы $z(i)$	Вероятность гипотезы до испытания $P_{z(i)}$	Вероятность пер-восортной детали по гипотезам $P_{A(i)}$	$P_{z(i)} \cdot P_{A(i)}$	Вероятность гипотезы после испытания (т. е. выборки детали 1 сорта $P_{z(i)}$ )
1	2	3	4	5	6
1	Деталь I сорта взята из ящика с 8 деталями I сорта (ящики № 1, 2 или 3) . . . .	$P_{z(1)} = \frac{3}{6}$	$P_{A(1)} = \frac{8}{10}$	$\frac{3}{6} \cdot \frac{8}{10} = \frac{24}{60}$	$\frac{24}{60} : \frac{41}{60} = \frac{24}{41}$
2	Деталь I сорта взята с 6 деталями I сорта (ящики № 4 или 5) . .	$P_{z(2)} = \frac{2}{6}$	$P_{A(2)} = \frac{6}{10}$	$\frac{2}{6} \cdot \frac{6}{10} = \frac{12}{60}$	$\frac{12}{60} : \frac{41}{60} = \frac{12}{41}$
3	Деталь I сорта взята из ящика с 5 деталями I сорта (ящик № 6)	$P_{z(3)} = \frac{1}{6}$	$P_{A(3)} = \frac{5}{10}$	$\frac{1}{6} \cdot \frac{5}{10} = \frac{5}{60}$	$\frac{5}{60} : \frac{41}{60} = \frac{5}{41}$
	—	$\Sigma P_{z(i)} = 1$	—	$\Sigma P_{z(i)} \cdot P_{A(i)} = \frac{41}{60}$	$\Sigma P_{z(i)} = 1$

Окончательный расчет для ответа на все три вопроса дан в колонке 6. По формуле (15) получаем:

$$P_{z(1)} = \frac{P_{z(1)} \cdot P_{A(1)}}{\Sigma P_{z(i)} \cdot P_{A(i)}} = \frac{24}{60} : \frac{41}{60} = \frac{24}{41};$$

$$P_{z(2)} = \frac{P_{z(2)} \cdot P_{A(2)}}{\Sigma P_{z(i)} \cdot P_{A(i)}} = \frac{12}{60} : \frac{41}{60} = \frac{12}{41};$$

$$P_{z(3)} = \frac{P_{z(3)} \cdot P_{A(3)}}{\Sigma P_{z(i)} \cdot P_{A(i)}} = \frac{5}{60} : \frac{41}{60} = \frac{5}{41}.$$

Сумма вероятностей гипотез до испытания и после испытания равна 1 (что видно из сумм колонок 3 и 6).

§ 11. Биномиальный закон вероятностей при  $n$  повторных независимых испытаниях очень часто называют бернуллиевым распределением вероятностей.

При повторных испытаниях, в каждом из которых может осуществиться некоторое событие «А» (с одной и той же вероятностью  $p$ ), вероятности любого числа его появлений соответ-

ствуют членам разложения бинома Ньютона в степени, равной числу испытаний.

$$(p+q)^n = p^n + np^{n-1} \cdot q + C_n^{n-2} p^{n-2} q^2 + C_n^{n-3} p^{n-3} q^3 + \dots + C_n^m p^m q^{n-m} + \dots + q^n, \quad (15)$$

где  $p$  — вероятность события „А“;  
 $q = 1 - p$  — вероятность того, что событие „А“ не произойдет;  
 $n$  — число испытаний;  
 $m$  — число осуществлений события „А“ или частота события „А“:

$$\left. \begin{aligned} C_n^{n-2} &= C_n^2 \\ C_n^{n-3} &= C_n^3 \\ C_n^m &= C_n^{n-m} \end{aligned} \right\} \text{— число сочетаний из } n \text{ элементов, по } n-2; n-3; n-m \text{ элементов;}$$

$p^n$  — первый член биномиальной строки дает значение вероятности такого исхода (комбинации), при котором событие „А“ осуществляется  $n$  раз (см. следствие первой теоремы умножения вероятностей—формула (10);

$np^{n-1}q$  — второй член биномиальной строки дает вероятность исхода, при котором событие „А“ осуществляется  $n-1$  раз, а не осуществляется один раз;

и т. д. до  $q^n$  — последнего члена строки, дающего вероятность такого исхода, при котором событие „А“ ни разу не осуществилось.

Таким образом, вероятность осуществления события «А»  $m$  раз в  $n$  независимых испытаниях с одинаковой вероятностью  $p$  можно рассчитать по формуле общего члена разложения бинома Ньютона:

$$P_m = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m}, \quad (16)$$

где  $P_m$  — вероятность того, что при  $n$  испытаниях событие „А“ осуществится  $m$  раз;

$n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$  (произведение натурального ряда чисел от 1 до  $n$ ) — читается „ $n$  факториал“;

$m! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot m$  читается „ $m$  факториал“;

$(n-m)! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-m)$ .

Примечание.  $0!$  считается равным единице (см. гамму функцию от  $n$  в разделе III, § 20).

Пример 14. По данным примера 6 найти вероятность того, что при отборе наудачу 5 деталей:

- 1) все 5 деталей окажутся I сорта;
- 2) 4 детали — I сорта и 1 деталь — не I сорта;
- 3) 3 детали — I сорта и 2 детали — не I сорта;
- 4) 2 детали — I сорта и 3 детали — не I сорта;
- 5) 1 деталь — I сорта и 4 детали — не I сорта;
- 6) все 5 деталей окажутся не I сорта.

Вероятности отбора детали I сорта при единичном испытании

$$p = \frac{150}{200} = \frac{3}{4}.$$

Вероятность отбора детали не I сорта при единичном испытании

$$q = \frac{50}{200} = \frac{1}{4}.$$

Рассчитываем по формуле (16):

1) вероятность того, что все 5 деталей окажутся I сорта.

Имеем:

$$p = \frac{3}{4}; n = 5;$$

$$q = \frac{1}{4}; m = 5.$$

$$\begin{aligned} P_5 &= \frac{5!}{5!(5-5)!} \left(\frac{3}{4}\right)^5 \cdot \left(\frac{1}{4}\right)^{5-5} = \left(\frac{3}{4}\right)^5 \cdot \left(\frac{1}{4}\right)^0 = \\ &= \frac{243}{1024} = 0,23730; \end{aligned}$$

2) вероятность того, что 4 детали окажутся I сорта, а одна деталь не I сорта.

Имеем:

$$p = \frac{3}{4}; n = 5;$$

$$q = \frac{1}{4}; m = 4;$$

$$\begin{aligned} P_4 &= \frac{5!}{4!(5-4)!} \cdot \left(\frac{3}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^1 = 5 \cdot \left(\frac{3}{4}\right)^4 \cdot \frac{1}{4} = \\ &= \frac{405}{1024} = 0,39551; \end{aligned}$$

3) вероятность того, что 3 детали окажутся I сорта, а 2 детали не I сорта:

$$p = \frac{3}{4}; n = 5;$$

$$q = \frac{1}{4}; m = 3.$$

$$\begin{aligned} P_3 &= \frac{5!}{3!(5-3)!} \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^{5-3} = 10 \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^2 = \\ &= \frac{270}{1024} = 0,26367; \end{aligned}$$

$$4) p = \frac{3}{4}; n = 5;$$

$$q = \frac{1}{4}; m = 2.$$

$$5) p = \frac{3}{4}; n = 5;$$

$$q = \frac{1}{4}; m = 1.$$

$$P_1 = \frac{5!}{1!(5-1)!} \cdot \left(\frac{3}{4}\right)^1 \cdot \left(\frac{1}{4}\right)^{5-1} = 5 \cdot \frac{3}{4} \left(\frac{1}{4}\right)^4 = \frac{15}{1024} = 0,01465;$$

$$6) p = \frac{3}{4}; n = 5;$$

$$q = \frac{1}{4}; m = 0;$$

$$P_0 = \frac{5!}{0!(5-0)!} \cdot \left(\frac{3}{4}\right)^0 \cdot \left(\frac{1}{4}\right)^5 = \left(\frac{1}{4}\right)^5 = \frac{1}{1024} = 0,00098.$$

Сумма всех найденных вероятностей равна:

$$\left(\frac{3}{4} + \frac{1}{4}\right)^5 = 1^5 = 1.$$

§ 12. Для определения коэффициентов при разложении в биномиальную строку используют треугольник Паскаля, в котором каждый коэффициент строки образуется сложением двух стоящих над ним (справа и слева) коэффициентов предыдущей строки.

$n$	Коэффициенты биномиальной строки																									
1						1		1																		
2					1		2		1																	
3				1		3		3		1																
4			1		4		6		4		1															
5		1		5		10		10		5		1														
6		1		6		15		20		15		6		1												
7		1		7		21		35		35		21		7		1										
8		1		8		28		56		70		56		28		8		1								
9		1		9		36		84		126		126		84		36		9		1						
10		1		10		45		120		210		252		210		120		45		10		1				
11		1		11		55		165		330		462		462		330		165		55		11		1		
12		1		12		66		220		495		792		924		792		495		220		66		12		1

$$C_n^0 \quad C_n^1 \quad C_n^2 \quad C_n^3$$

и т. д.

$$C_n^{n-1} \quad C_n^n$$

Пример 15. В примере 14 нам необходимы были коэффициенты разложения при  $n=5$ , которые можно взять из треугольника Паскаля (см. 5-ю строку).

При  $n=5$  имеем коэффициенты: 1; 5; 10; 10; 5; 1.

Следовательно, искомые вероятности соответственно равны:

$$\left(\frac{3}{4}\right)^5; 5\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right); 10\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^2; 10\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^3; 5\frac{3}{4}\left(\frac{1}{4}\right)^4; \left(\frac{1}{4}\right)^5.$$

§ 13. Каждый возможный исход событий обладает определенной вероятностью. Если на оси абсцисс наносить возможные исходы событий, а по оси ординат — вероятности этих исходов,

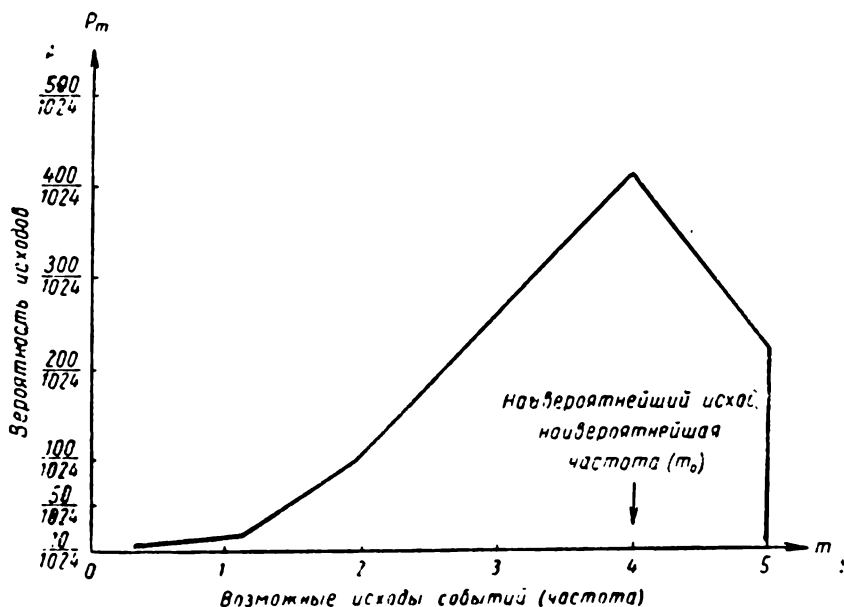


График 7. Полигон распределения вероятностей при  $p = \frac{3}{4}$ ;  $n = 5$ .

то ломаная линия, характеризующая изменение вероятностей различных исходов событий при повторных испытаниях, называется полигоном (многоугольником) распределения вероятностей (см. раздел I, § 8).

Пример 16. По результатам разложения в биномиальную строку в примере 14 построить полигон распределения вероятностей.

В системе координат строим полигон (см. график 7).

Обычно при построении полигона распределения вероятностей используют формулу 16. Придавая  $m$  различные значения от 0 до  $n$ , получают вероятности  $P_0; P_1; P_2 \dots, P_n$ , которые наносятся на график.

**Пример 17.** Дано  $p = \frac{1}{2}$ ;  $n = 11$ ; построить многоугольник распределения вероятностей.

Находим вероятности различных исходов (частот) по формуле 16:

$$P_0 = C_{11}^0 \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^{11} = \left(\frac{1}{2}\right)^{11} = 0,00049;$$

$$P_1 = C_{11}^1 \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^{10} = \frac{11}{2048} = 0,00537;$$

$$P_2 = C_{11}^2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^9 = \frac{55}{2048} = 0,02685;$$

$$P_3 = C_{11}^3 \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^8 = \frac{165}{2048} = 0,08057;$$

$$P_4 = C_{11}^4 \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^7 = \frac{330}{2048} = 0,16113;$$

$$P_5 = C_{11}^5 \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^6 = \frac{462}{2048} = 0,22559;$$

$$P_6 = C_{11}^6 \cdot \left(\frac{1}{2}\right)^6 \cdot \left(\frac{1}{2}\right)^5 = \frac{462}{2048} = 0,22559;$$

$$P_7 = C_{11}^7 \cdot \left(\frac{1}{2}\right)^7 \cdot \left(\frac{1}{2}\right)^4 = \frac{330}{2048} = 0,16113;$$

$$P_8 = C_{11}^8 \cdot \left(\frac{1}{2}\right)^8 \cdot \left(\frac{1}{2}\right)^3 = \frac{165}{2048} = 0,08057;$$

$$P_9 = C_{11}^9 \cdot \left(\frac{1}{2}\right)^9 \cdot \left(\frac{1}{2}\right)^2 = \frac{55}{2048} = 0,02685;$$

$$P_{10} = C_{11}^{10} \cdot \left(\frac{1}{2}\right)^{10} \cdot \left(\frac{1}{2}\right)^1 = \frac{11}{2048} = 0,00537;$$

$$P_{11} = C_{11}^{11} \cdot \left(\frac{1}{2}\right)^{11} \cdot \left(\frac{1}{2}\right)^0 = \frac{1}{2048} = 0,00049.$$

Наносим полученные данные на график (см. график 8).

§ 14. По полигонам распределения вероятностей наглядно видно, что частоты обладают различными вероятностями. Одни

частоты менее вероятны, другие более вероятны. Имеется частота, обладающая наибольшей вероятностью (иногда две частоты с одинаковыми наибольшими вероятностями). Такая частота на-

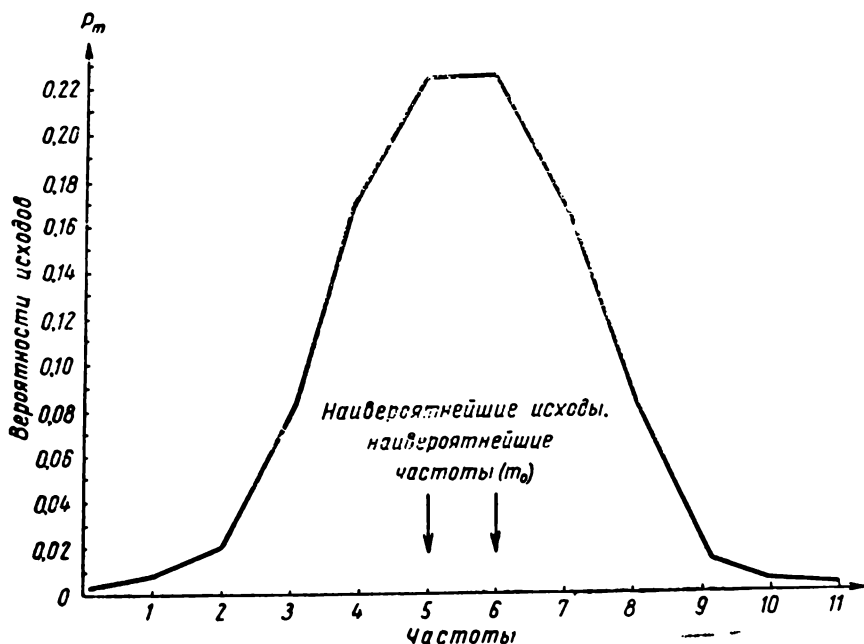


График 8. Полигон распределения вероятностей при  $p = \frac{1}{2}$ ;  $n = 11$ .

зывается наивероятнейшей частотой, или наивероятнейшим исходом, обозначается  $m_0$  и определяется из следующего неравенства:

$$np - q \leq m_0 \leq np + p, \quad (17)$$

где  $n$  — число испытаний;

$p$  — вероятность события „А“ при одном испытании;

$q = 1 - p$ .

Примечания: 1.  $m_0$  выражается в целых числах.

2. Если границы неравенства (левая и правая) — целые числа, то, следовательно, имеются две наивероятнейшие частоты.

3. Если границы неравенства (левая и правая) — дроби, то между ними выбирают целое число, которое и будет наивероятнейшей частотой. Таким образом, заключаем, что  $m_0 \approx np$  есть ближайшее к  $np$  целое число.

Пример 18. По данным примеров 14 и 6 при отборе 5 деталей определить наивероятнейший исход.

Имеем:

$$p = \frac{3}{4}; n = 5.$$

Следовательно:

$$m_0 \approx np \approx \frac{3}{4} \cdot 5 = \frac{15}{4} = 3 \frac{3}{4}.$$

по формуле 17 получаем:

$$\frac{15}{4} - \frac{1}{4} \leq m_0 \leq \frac{15}{4} + \frac{3}{4}; 3,5 < m < 4,5.$$

Границы (левая и правая) есть дроби 3,5 и 4,5.

Целым числом между ними является 4. Следовательно,  $m_0 = 4$ .

По графику 7, где частота ( $m$ ), равная 4, обладает наибольшей вероятностью, убеждаемся в совпадении результатов.

**Пример 19.** По данным примера 17 определить наимвероятнейшую частоту.

Имеем:

$$p = \frac{1}{2}; n = 11; m_0 \approx \frac{1}{2} \cdot 11 \approx \frac{11}{2} \approx 5,5.$$

по формуле 17 получаем:

$$5,5 - 0,5 \leq m_0 \leq 5,5 + 0,5; 5 < m_0 < 6.$$

Границы (левая и правая) — целые числа. Следовательно, мы имеем две наимвероятнейшие частоты: 5 и 6. По графику 8, где частоты 5 и 6 обладают равными наибольшими вероятностями, убеждаемся в совпадении результатов.

§ 15. Для расчета вероятностей всех исходов вместо разложения в биномиальную строку можно воспользоваться одной из следующих формул:

$$P_{m+1} = P_m \cdot \frac{n-m}{m+1} \cdot \frac{p}{q}; \quad (18)$$

$$P_{m-1} = P_m \cdot \frac{m}{n-m+1} \cdot \frac{q}{p}. \quad (19)$$

**Пример 20.** По условию примера 17 рассчитать вероятности всех частот по формулам 18 и 19.

Для использования формул 18 и 19 нужно найти вероятность какой-нибудь одной частоты. Пусть мы нашли вероятность частоты 5 по формуле 16:

$$P_5 = \frac{462}{2048}.$$

Используя формулу 18, находим вероятности всех частот, больших 5:

$$P_6 = \frac{462}{2048} \cdot \frac{11-5}{5+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{462}{2048} \cdot \frac{6}{6} = \frac{462}{2048};$$



$$P_7 = \frac{462}{2048} \cdot \frac{11-6}{6+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{462}{2048} \cdot \frac{5}{7} = \frac{330}{2048};$$

$$P_8 = \frac{330}{2048} \cdot \frac{11-7}{7+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{330}{2048} \cdot \frac{4}{8} = \frac{165}{2048};$$

$$P_9 = \frac{165}{2048} \cdot \frac{11-8}{8+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{165}{2048} \cdot \frac{3}{9} = \frac{55}{2048};$$

$$P_{10} = \frac{55}{2048} \cdot \frac{11-9}{9+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{55}{2048} \cdot \frac{2}{10} = \frac{11}{2048};$$

$$P_{11} = \frac{11}{2048} \cdot \frac{11-10}{10+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{11}{2048} \cdot \frac{1}{11} = \frac{1}{2048}.$$

Используем формулу (19):

$$P_4 = \frac{462}{2048} \cdot \frac{5}{11-5+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{462}{2048} \cdot \frac{5}{7} = \frac{330}{2048};$$

$$P_3 = \frac{330}{2048} \cdot \frac{4}{11-4+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{330}{2048} \cdot \frac{4}{8} = \frac{165}{2048};$$

$$P_2 = \frac{165}{2048} \cdot \frac{3}{11-3+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{165}{2048} \cdot \frac{3}{9} = \frac{55}{2048};$$

$$P_1 = \frac{55}{2048} \cdot \frac{2}{11-2+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{55}{2048} \cdot \frac{2}{10} = \frac{11}{2048};$$

$$P_0 = \frac{11}{2048} \cdot \frac{1}{11-1+1} \cdot \frac{\frac{1}{2}}{\frac{1}{2}} = \frac{11}{2048} \cdot \frac{1}{11} = \frac{1}{2048}.$$

Как видим, вероятности всех частот совпадают с результатами, полученными в примере 17.

§ 16. Для расчета вероятности наивероятнейшей частоты используется формула, дающая приближенный результат. Точность формулы зависит от числа испытаний ( $n$ ) и по мере увеличения числа испытаний возрастает.

$$P_{m_0} \approx \frac{1}{\sqrt{2\pi npq}}, \quad (20)$$

где  $m_0 \approx np$  — есть наивероятнейшая частота;  
 $\pi$  — число, равное 3,14159.

Для расчетов используют эту формулу в несколько измененном виде:

$$P_{m_0} = \frac{1}{\sqrt{2\pi}} : \sqrt{npq} = \frac{0,3989}{\sqrt{npq}}. \quad (21)$$

Пример 21. В примерах 18 и 19 найдены наивероятнейшие частоты. Вычислить приближенно их вероятности.

По примеру 18 имеем:  $n = 5$ ;  $p = \frac{3}{4}$ ;  $m_0 = 4$ .

$$\begin{aligned} P_{m_0} = P_4 &\approx \frac{0,3989}{\sqrt{5 \cdot \frac{3}{4} \cdot \frac{1}{4}}} = \frac{0,3989}{\sqrt{\frac{15}{16}}} = \frac{0,3989}{\sqrt{0,9375}} = \\ &= \frac{0,3989}{0,96825} = 0,41198. \end{aligned}$$

По примеру 19 имеем:  $n = 11$ ;  $p = \frac{1}{2}$ ;  $m_0 = 5$  и 6.

$$\begin{aligned} P_{m_0} = P_{5 \text{ и } 6} &\approx \frac{0,3989}{\sqrt{11 \cdot \frac{1}{2} \cdot \frac{1}{2}}} = \frac{0,3989}{\sqrt{\frac{11}{4}}} = \\ &= \frac{0,3989}{\sqrt{2,75}} = \frac{0,3989}{1,65831} = 0,24055. \end{aligned}$$

Сопоставляя приближенные результаты вероятностей наивероятнейших исходов, полученные в данном примере, с точными, полученными в примерах 14 и 17, видим их несомненную близость.

В примере 14:  $P_4=0,39551$ , в данном примере  $P_4=0,41198$ ;  
в примере 17:  $P_{5 и 6}=0,22559$ , в данном примере  $P_{5 и 6}=0,24055$ .  
Погрешности не превышают 5%.

§ 17. Для приближенного вычисления вероятностей биномиального распределения используется формула Лапласа:

$$P_x \approx \frac{1}{\sqrt{2\pi npq}} \cdot e^{-\frac{x^2}{2npq}}, \quad (22)$$

где  $x$  — отклонение частоты от  $np$ , т. е.  $x = m - np$ ;

$P_x$  — вероятность такого события, при котором частота отклоняется от  $np$  на величину  $x$ ;

$e$  — неперово число, основание натуральных логарифмов, приближенно равное 2,71828;

$\frac{1}{\sqrt{2\pi npq}} = P_{m_0}$  — вероятность наивероятнейшей частоты (см. формулу 20).

**Пример 22.** Из партии, в которой доля деталей I сорта равна  $\frac{2}{5}$ , отобрано 50 единиц (с возвратом). Определить вероятность того, что среди отобранных единиц 25 окажутся I сорта. .

Дано:  $n = 50$ ;  $m_0 = np = 50 \cdot \frac{2}{5} = 20$ .

$$p = \frac{2}{5};$$

$$q = \frac{3}{5};$$

$$m = 25.$$

Мы должны найти вероятность частоты 25 ( $P_{25}$ ). Находим отклонение искомой частоты от наивероятнейшей:  $x = 25 - 20 = 5$ .

По формуле 21 определяем вероятность наивероятнейшей частоты:

$$\begin{aligned} P_{m_0=20} &= \frac{1}{\sqrt{2\pi npq}} = \frac{0,3989}{\sqrt{npq}} = \\ &= \frac{0,3989}{\sqrt{50 \cdot \frac{2}{5} \cdot \frac{3}{5}}} = \frac{0,3989}{\sqrt{12}} = \frac{0,3989}{3,4641} = 0,11515, \end{aligned}$$

а по формуле (22) получаем вероятность того, что частота отклонится от наивероятнейшей частоты на 5.

$$P_{x=5} = 0,11515 \cdot e^{-\frac{25}{2 \cdot 50 \cdot \frac{2}{5} \cdot \frac{3}{5}}} = 0,11515 \cdot e^{-1,042}.$$

Для нахождения второго множителя обратимся к таблице (см. приложение X), из которой приближенно найдем:

$$e^{-1,042} \approx 0,35345,$$

тогда получаем:

$$P_{x=5} = 0,11515 \cdot 0,35345 = 0,0407.$$

Характерно, что и вероятность частоты 15 будет такой же. В формулу Лапласа  $x$  входит возведенным в квадрат, а поэтому  $+x$  и  $-x$  приводят по этой формуле к одинаковой вероятности:

$$P_x = P_{-x}. \quad (23)$$

Так, если мы ищем вероятность частоты 15, то

$x = 15 - 20 = -5$ , а поэтому  $P_{x=5} = P_{x=-5}$ .

§ 18. Для вычисления по формуле 22 в ней производят замену:

$$\frac{x}{\sqrt{npq}} = \frac{m - m_0}{\sqrt{npq}} = t. \quad (24)$$

$$P_t \approx \frac{1}{\sqrt{2\pi npq}} \cdot e^{-\frac{t^2}{2}}. \quad (25)$$

Имеются таблицы значений:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

(см. приложение II).

Следовательно:

$$P_t = \frac{\varphi(t)}{\sqrt{npq}}. \quad (26)$$

Придавая  $m$  различные значения, определяем сначала  $t$  по формуле 24, затем  $\varphi(t)$  по приложению II, найденные  $\varphi(t)$  делим на  $\sqrt{npq}$ .

Пример 23. По условию примера 22 определить искомую вероятность, пользуясь формулой 26.

Дано:

$$n = 50; \quad p = \frac{2}{5}; \quad q = \frac{3}{5}; \quad m = 25;$$

$$m_0 = 20; \quad x = 5; \quad \sqrt{npq} = 3,4641.$$

Находим по формуле (24)  $t$  при  $m = 25$ .

$$t = \frac{5}{3,4641} = 1,443, \text{ берем приближенно } t = 1,44.$$

По приложению II определяем:  $\varphi(1,44) = 0,1415$ .

По формуле 26 получаем:

$$P_{t=1.44} = \frac{0,1415}{3,4641} = 0,0408.$$

Расхождение в 0,0001 получилось за счет неточно взятого  $t$ .

§ 19. Из формулы 25 можно получить формулу, выражающую стандартизованное распределение вероятностей.

Мы знаем, что множитель  $\frac{1}{\sqrt{2\pi npq}}$  в формуле 25 есть вероятность наивероятнейшей частоты —  $P_{m_0}$ . Тогда, заменяя, получим:

$$P_t = P_{m_0} \cdot e^{-\frac{t^2}{2}}. \quad (27)$$

Используя приложение III, в котором даны значения  $e$  в различных степенях, можно по формуле 27 построить таблицу стандартизованного распределения вероятностей.

Таблица 6

$t$	$x$	$P_t = P_{m_0} \cdot e^{-\frac{t^2}{2}}$
1	2	3
0	0	$P_{m_0}$
$\pm \frac{1}{4}$	$\pm \frac{1}{4} \sqrt{npq}$	$0,96948 \cdot P_{m_0}$
$\pm \frac{1}{2}$	$\pm \frac{1}{2} \sqrt{npq}$	$0,88251 \cdot P_{m_0}$
$\pm 1$	$\pm \sqrt{npq}$	$0,60653 \cdot P_{m_0}$
$\pm \frac{3}{2}$	$\pm \frac{3}{2} \sqrt{npq}$	$0,32466 \cdot P_{m_0}$
$\pm 2$	$\pm 2 \sqrt{npq}$	$0,13534 \cdot P_{m_0}$
$\pm \frac{5}{2}$	$\pm \frac{5}{2} \sqrt{npq}$	$0,04394 \cdot P_{m_0}$
$\pm 3$	$\pm 3 \sqrt{npq}$	$0,01111 \cdot P_{m_0}$
$\pm 4$	$\pm 4 \sqrt{npq}$	$0,000336 \cdot P_{m_0}$
$\pm 5$	$\pm 5 \sqrt{npq}$	$0,00000373 \cdot P_{m_0}$
$\pm 6$	$\pm 6 \sqrt{npq}$	$0,000000015 \cdot P_{m_0}$

По полученным в таблице показателям строят график стандартизованного распределения вероятностей.

На оси абсцисс наносят значения  $t$  и  $x$ . На оси ординат откладывают найденные вероятности ( $P_t$ ).

Графическое изображение стандартизованного распределения вероятностей называется нормальной кривой, или кривой Гаусса-Лапласа (способы построения см. раздел IV).

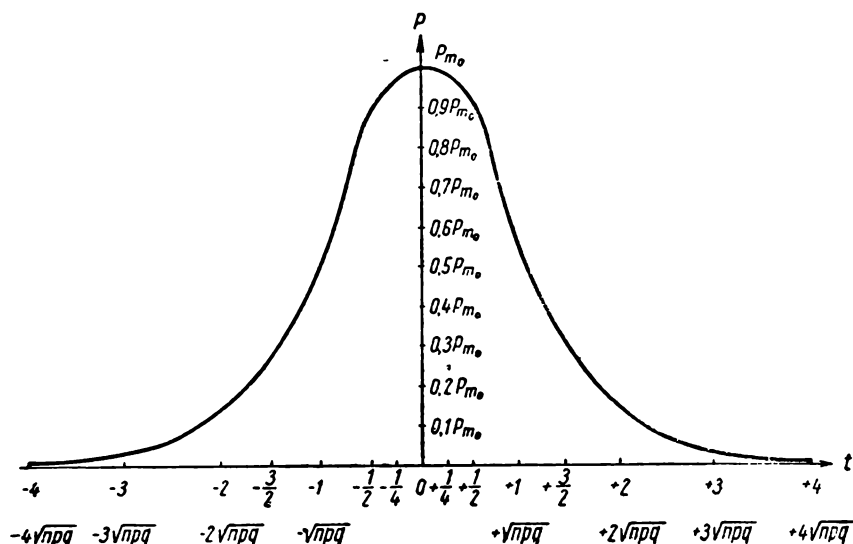


График 9. Стандартизованное распределение вероятностей.

Особенностями данной кривой являются:

Максимальная ордината равна вероятности наивероятнейшей частоты.

Кривая симметрична относительно оси ординат.

По мере удаления  $t$  от 0 в обе стороны вероятности уменьшаются. Это означает, что более вероятными являются частоты, примыкающие к наивероятной частоте.

§ 20. Для приближенного исчисления вероятностей по схеме «невозвращенного шара» (бесповторный отбор, при котором отобранная единица в совокупность не возвращается) можно воспользоваться несколько иной формулой:

$$P_x = \frac{1}{\sqrt{2\pi npq \left(1 - \frac{n}{N}\right)}} \cdot e^{-\frac{x^2}{2npq \left(1 - \frac{n}{N}\right)}} \quad (28)$$

где  $N$  — численность совокупности, из которой производится отбор.

Если  $n$  по сравнению с  $N$  — величина незначительная, то формула 28 сводится к формуле 22.

§ 21. Если необходимо рассчитать вероятность того, что при  $n$  испытаниях число осуществлений события «А» будет находиться

в заданных границах  $a$  и  $b$ , используют интегральную формулу Лапласа:

$$P_{(a < m < b)} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} \cdot dt = \frac{1}{2} [F(t=\beta) - F(t=\alpha)], \quad (29)$$

где  $a$  и  $b$  — заданные границы, числа осуществлений события „А“;

$\alpha$  и  $\beta$  — пределы интегрирования (см. ниже);

$P_{(a < m < b)}$  — вероятность того, что при  $n$  неограниченно возрастающем числе испытаний, число появлений события „А“ будет заключено между числами  $a$  и  $b$ .

С геометрической точки зрения интеграл

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$$

представляет площадь, заключенную между нормальной кривой вероятностей и соответствующими ординатами.

В приложении III даны значения  $F(t)$ , по которым можно определять искомое значение вероятности.

Числа  $a$  и  $b$  связаны с границами интегрирования  $\alpha$  и  $\beta$  определенным соотношением:

$$\alpha = \frac{a - m_0}{\sqrt{npq}};$$

$$\beta = \frac{b - m_0}{\sqrt{npq}}.$$

(Практическое применение смотреть в примере 24 — второй вопрос.)

§ 22. Если заданные границы числа осуществлений события «А» ( $a$  и  $b$ ) отличаются от  $np$  на одинаковую величину  $t$ , то интегральная формула Лапласа примет вид:

$$P_{(a < m < b)} = \frac{1}{\sqrt{2\pi}} \cdot \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} \cdot dt = F(t). \quad (30)$$

В этом случае

$$t = \frac{m_0 - a}{\sqrt{npq}} = \frac{b - m_0}{\sqrt{npq}}.$$

**Пример 24.** Из партии в 500 деталей, в которой первосортных 300, отбирается наудачу 150 деталей с возвратом. Какова вероятность того, что число деталей I сорта будет находиться в пределах:

1) от 78 до 102?    2) от 78 до 108?

Для первого вопроса дано:

$$n = 150; \quad p = \frac{300}{500} = \frac{3}{5};$$

$$q = \frac{2}{5}; \quad a = 78; \quad b = 102.$$

Находим:

$$np = 150 \cdot \frac{3}{5} = 90;$$

$$\sqrt{npq} = \sqrt{150 \cdot \frac{3}{5} \cdot \frac{2}{5}} = 6.$$

Границы  $a$  и  $b$  отличаются от  $m_0$  на одинаковую величину равную 12;

$a$	$m_0$	$b$
78	90	102

Находим  $t$  из соотношений:

$$t = \frac{m_0 - a}{\sqrt{npq}} \quad \text{или} \quad t = \frac{b - m_0}{\sqrt{npq}}.$$

Получаем:

$$t = \frac{90 - 78}{6} = 2 \quad \text{и} \quad t = \frac{102 - 90}{6} = 2.$$

Для использования таблицы берем  $t = 2$ .

Следовательно:

$$P(78 \leq m \leq 102) = F(2).$$

По таблице приложения III находим  $F(2) = 0,9545$ .

Для второго вопроса ищем вероятность того, что частота лежит в границах от 78 до 108.

Имеем:

$$a = 78; \quad b = 108; \quad n = 150; \quad p = \frac{3}{5}; \quad q = \frac{2}{5};$$

$$m_0 = 90; \quad \sqrt{npq} = 6.$$

Находим  $\alpha$  и  $\beta$  из соотношений:

$$\alpha = \frac{a - m_0}{\sqrt{npq}}; \quad \beta = \frac{b - m_0}{\sqrt{npq}}.$$



Получаем:

$$\alpha = \frac{78 - 90}{6} = -2;$$

$$\beta = \frac{108 - 90}{6} = 3.$$

По формуле 29:

$$\begin{aligned} P_{(78 < m < 108)} &= \frac{1}{\sqrt{2\pi}} \int_{-2}^3 e^{-\frac{t^2}{2}} \cdot dt = \frac{1}{2} [F(t=3) - F(t=-2)] = \\ &= \frac{1}{2} [F(t=3) + F(t=2)]. \end{aligned}$$

По приложению III находим:

$$F(t=3) = 0,9973;$$

$$F(t=2) = 0,9545.$$

Тогда:

$$\frac{1}{2} [0,9973 + 0,9545] = 0,9759.$$

Значит, вероятность того, что число деталей I сорта находится в границах от 78 до 108, равна 0,9759.

§ 23. Величина  $x$ , принимающая в зависимости от некоторых случайных обстоятельств одно из значений  $x_1, x_2, x_3 \dots x_n$ , имеющих определенные вероятности  $p_1, p_2, p_3 \dots p_n$ , называется случайной величиной.

Случайные величины бывают дискретными (с дискретным рядом возможных значений) и непрерывными (имеющие сколь угодно близкие возможные значения). Совокупность значений случайных величин и соответствующих вероятностей называют распределением случайной величины.

Пример 25. Если производится отбор 11 деталей из партии, в которой вероятность первосортной детали равна  $\frac{1}{2}$ , то частота ( $m$ ) представляет собой дискретную случайную величину, вероятность значений которой можно рассчитать по соответствующим формулам (например, по формуле 16).

Таблица 7

$m$	0	1	2	3	4
$P_m$	$C_{11}^0 p^0 q^{11}$	$C_{11}^1 p^1 q^{10}$	$C_{11}^2 p^2 q^9$	$C_{11}^3 p^3 q^8$	$C_{11}^4 p^4 q^7$

5	6	7	8	9	10	11
$C_{11}^5 p^5 q^6$	$C_{11}^6 p^6 q^5$	$C_{11}^7 p^7 q^4$	$C_{11}^8 p^8 q^3$	$C_{11}^9 p^9 q^2$	$C_{11}^{10} p^{10} q^1$	$C_{11}^{11} p^{11} q^0$

Здесь совокупность  $m$  и  $P_m$  будет распределением случайной величины  $m$ . В данном примере приводится биномиальное распределение.

§ 24. Математическое ожидание дискретной случайной величины  $x$  равно сумме произведений каждого возможного значения этой величины на его вероятность.

$$E(x) = \sum_{i=0}^{i=n} x_i \cdot p_i = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n, \quad (31)$$

где  $E$  — символ математического ожидания;

$E(x)$  — математическое ожидание случайной величины  $x$ , которое очень часто называют центром распределения или центром рассеивания.

Пример 26. По результатам примера 17 ищем математическое ожидание случайной величины.

Дано:

$$n = 11; \quad p = \frac{1}{2}; \quad q = \frac{1}{2}.$$

Располагаем полученные данные в таблицу и вычисляем произведения частот на их вероятности:

Таблица 8

$m$	$P_m$	$mP_m$
1	2	3
0	0,00049	0
1	0,00537	0,00537
2	0,02685	0,05370
3	0,08057	0,24171
4	0,16113	0,64452
5	0,22559	1,12795
6	0,22559	1,35354
7	0,16113	1,12791
8	0,08057	0,64456
9	0,02685	0,24165
10	0,00537	0,05370
11	0,00049	0,00539
Сумма . . .	1,00000	5,50000

Получаем по формуле (31):

$$E(m) = \sum_{l=0}^{l=n} m_l P_l = 5,5.$$

§ 25. Математическое ожидание непрерывной случайной величины:

$$E(x) = \int_{-\infty}^{+\infty} x \cdot \varphi(x) \cdot dx, \quad (32)$$

где  $\varphi(x)$  — функция плотности вероятностей.

§ 26. Математическое ожидание случайной величины равно средней ее значений, взвешенной по вероятностям:

$$E(x) = \sum_{l=0}^{l=n} x_l p_l = \bar{x}. \quad (33)$$

§ 27. Свойства математического ожидания:

а. Математическое ожидание постоянной величины равно этой постоянной:

$$E(a) = a. \quad (34)$$

б. Математическое ожидание алгебраической суммы случайных величин равно алгебраической сумме их математических ожиданий:

$$E(x \pm y \pm z \pm \dots \pm w) = E(x) \pm E(y) \pm \dots \pm E(w). \quad (35)$$

в. Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий:

$$E(xy) = E(x) \cdot E(y), \quad (36)$$

если  $x$  и  $y$  независимы<sup>1</sup>.

г. Постоянный множитель выносится за знак математического ожидания:

$$E(ax) = a \cdot E(x). \quad (37)$$

д. Математическое ожидание случайной величины, принимающей только два значения: 1 — при осуществлении некоторого события и 0 — при его неосуществлении, равно вероятности события:

$$E(x) = p. \quad (38)$$

<sup>1</sup> Случайные величины независимы, если вероятности значений одной не зависят от того, какое значение приняла другая.

**Пример 27.** В партии 500 деталей, из них I сорта 300. Производится отбор одной детали. Найти математическое ожидание числа деталей I сорта.

Имеем два варианта: 1 и 0.

Найдем вероятности вариантов: .

$$p_1 = \frac{300}{500} = 0,6 \text{ (вероятность варианта 1);}$$

$$p_2 = 1 - 0,6 = 0,4 \text{ (вероятность варианта 0).}$$

Получаем:

$$E(x) = \sum x_i p_i = 1 \cdot 0,6 + 0 \cdot 0,4 = 0,6$$

и видим, что математическое ожидание данной случайной величины равно вероятности первого варианта.

е. Математическое ожидание случайной величины всегда заключено между наименьшим и наибольшим ее значением:

$$x_{\min} < E(x) < x_{\max}. \quad (39)$$

ж. Математическое ожидание частоты при биномиальном распределении равно произведению числа испытаний на вероятность события:

$$E(m) = np. \quad (40)$$

**§ 28.** Математическое ожидание квадрата отклонения случайной величины от ее математического ожидания называется ее дисперсией, обозначается  $\sigma^2$  и вычисляется по формуле:

$$\begin{aligned} \sigma^2 = E[x - E(x)]^2 &= [x_1 - E(x)]^2 \cdot p_1 + [x_2 - E(x)]^2 \cdot p_2 + \\ &+ \dots + [x_n - E(x)]^2 \cdot p_n = \sum_{i=1}^n [x_i - E(x)]^2 \cdot p_i. \end{aligned} \quad (41)$$

**§ 29.** Среднее квадратическое отклонение ( $\sigma$ ) равно квадратному корню из дисперсии:

$$\sigma = \sqrt{E[x - E(x)]^2}. \quad (42)$$

**§ 30. Лемма Маркова.** Если имеется  $k$  вариантов случайной величины, могущей принять одно из положительных значений  $x_1, x_2, x_3, \dots, x_{n+1}, x_{n+2}, \dots, x_k$  с соответствующими вероятностями  $p_1, p_2, p_3, \dots, p_{n+1}, p_{n+2}, \dots, p_k$  и  $E(x) = \bar{x}$ , то вероятность того, что  $x \leq \bar{x} \cdot t^2$  будет больше, чем  $1 - \frac{1}{t^2}$ , где число  $t^2$  больше единицы:

$$P[x \leq \bar{x} \cdot t^2] > 1 - \frac{1}{t^2}. \quad (43)$$

Можно лемму Маркова записать и так:

$$P[x \leq a] > 1 - \frac{E(x)}{a}, \quad (44)$$

если принять  $\bar{x} \cdot t^2 = a$  и учитывая, что  $E(x) = \bar{x}$ .

Следствие леммы Маркова.

Вероятность того, что  $x > \bar{x}t^2$  не больше  $\frac{1}{t^2}$ :

$$Q[x > a] \leq \frac{E(x)}{a}. \quad (45)$$

**§ 31. Первое неравенство Чебышева.** Из леммы Маркова вытекает неравенство, называемое первым неравенством Чебышева. Если случайная величина « $u$ » принимает положительные и отрицательные значения и если « $a$ » — произвольная величина, то вероятность того, что случайная величина « $u$ » будет находиться в границах между  $-a$  и  $+a$ , т. е. вероятность неравенства  $-a \leq u \leq +a$  [или, что то же самое, неравенства  $u^2 \leq a^2$ ] будет больше  $1 - \frac{E(u^2)}{a^2}$ :

$$P[-a \leq u \leq +a] > 1 - \frac{E(u^2)}{a^2}. \quad (46)$$

**§ 32. Следствие первого неравенства Чебышева.** Возьмем вместо случайной величины « $u$ » другую случайную величину  $x - E(x)$ , тогда первое неравенство Чебышева примет вид:

$$P[-a \leq x - E(x) \leq +a] > 1 - \frac{E[x - E(x)]^2}{a^2}$$

и, заменяя  $E[x - E(x)]^2$  дисперсией ( $\sigma^2$ ), получим:

$$P[-a \leq x - E(x) \leq +a] > 1 - \frac{\sigma^2}{a^2}. \quad (47)$$

**§ 33. Второе неравенство Чебышева.** Если вместо случайной величины « $u$ » из первого неравенства Чебышева взять другую случайную величину  $\{(x + y + \dots + w - [E(x) + E(y) + \dots + E(w)])\}$ , т. е. отклонение суммы случайных величин от суммы их математических ожиданий, и учесть следствие первого неравенства Чебышева, то получим второе неравенство Чебышева:

$$P\{-a \leq (x + y + \dots + w) - [E(x) + E(y) + \dots + E(w)] \leq +a\} > 1 - \frac{\sigma_{(x+y+\dots+w)}^2}{a^2}. \quad (48)$$

Если при этом величины  $x, y, \dots w$  взаимно независимы, то

$$P\{-a \leq (x + y + \dots + w) - [E(x) + E(y) + \dots + E(w)] \leq +a\} > 1 - \frac{\sigma_x^2 + \sigma_y^2 + \dots + \sigma_w^2}{a^2}. \quad (48a)$$

§ 34. Теорема Чебышева. Второе неравенство Чебышева может быть записано иначе.

Если положить

$$a^2 = t^2 \sigma_{(x+y+\dots+w)}^2;$$

$$a = t \sigma_{(x+y+\dots+w)},$$

то в правой части 48a получим  $1 - \frac{1}{t^2}$ . С другой стороны, слева тройное неравенство в скобках разделим на  $n$ . Тогда получим

$$P\left\{-\frac{a}{n} \leq \frac{x+y+\dots+w}{n} - \frac{E(x)+E(y)+\dots+E(w)}{n} \leq \frac{a}{n}\right\} > 1 - \frac{1}{t^2}.$$

Но если все  $\sigma_x^2, \sigma_y^2, \dots, \sigma_w^2$  меньше некоторого числа  $L$ , то

$$a^2 = t^2 (\sigma_x^2 + \sigma_y^2 + \dots + \sigma_w^2) < t^2 nL$$

и

$$a < t \sqrt{nL}$$

и, наконец,  $\frac{a}{n} < t \sqrt{\frac{L}{n}}$ , что при достаточном  $n$  может быть

сделано сколь угодно малой величиной  $\epsilon$ . Иначе говоря:

$$P\left\{-\epsilon \leq \frac{x+y+\dots+w}{n} - \frac{E(x)+E(y)+\dots+E(w)}{n} \leq \epsilon\right\} > 1 - \frac{1}{t^2}, \quad (49)$$

где  $\epsilon$  — сколь угодно малая величина.

Эта формула 49 выражает знаменитую теорему Чебышева и читается так: с вероятностью, сколь угодно близкой к единице (доверности), можно утверждать, что если  $x, y, \dots w$  суть независимые случайные величины, имеющие определенные математические ожидания и ограниченные дисперсии, то при достаточно большом числе случайных величин их средняя арифметическая

будет как угодно мало отличаться от средней арифметической их математических ожиданий.

§ 35. **Следствие теоремы Чебышева.** Если  $x', x'', x''', \dots, x^{(n)}$  — случайные величины с одинаковыми возможными значениями:  $x_1, x_2, x_3, \dots, x_n$  и одинаковыми их вероятностями:  $p_1, p_2, p_3, \dots, p_m$ , то при достаточно большом числе испытаний средняя арифметическая этих величин будет как угодно мало отличаться от их математического ожидания  $E(x)$

$$P \left\{ -\frac{ts_x}{\sqrt{n}} \leq \left[ \frac{x' + x'' + \dots + x^{(n)}}{n} - E(x) \right] \leq \frac{ts_x}{\sqrt{n}} \right\} > 1 - \frac{1}{t^2}. \quad (50)$$

Чтобы убедиться в этом, достаточно учесть, что в этом случае

$$a = t \sqrt{n\sigma_x^2} = ts_x \sqrt{n} \quad \text{и} \quad \frac{a}{n} = \frac{ts_x}{\sqrt{n}}.$$


---

### РАЗДЕЛ III

## ВЫБОРОЧНЫЙ МЕТОД

Теоретические основы выборочного метода содержатся в теоремах Чебышева и Ляпунова.

Основной предпосылкой применения выборочного метода является возможность судить о характеристиках генеральной (общей) совокупности по отобранной, так называемой выборочной совокупности. Наиболее важным принципом в применении выборочного метода является обеспечение равной возможности всем единицам, входящим в состав генеральной совокупности, быть избранными. При таком объективном подходе к отбору единиц, при котором ни одна единица не обладает преимуществом попасть в отбираемую совокупность, по сравнению с другими единицами, характеристики выборочной совокупности при увеличении объема выборки стремятся к характеристикам генеральной совокупности.

§ 1. *Теорема Чебышева* в отношении к выборочному методу может быть записана в следующем виде:

$$P \left[ (\tilde{x} - \bar{x}) < \frac{t\alpha}{\sqrt{n}} \right] > 1 - \frac{1}{t^2}, \quad (1)$$

где  $\tilde{x}$  — средняя по совокупности выбранных единиц;

$\bar{x}$  — средняя по генеральной совокупности;

$\alpha$  — среднее квадратическое отклонение в генеральной совокупности.

Теорема формулируется так: с вероятностью, сколь угодно близкой к единице (достоверности), можно утверждать, что при достаточно большом объеме выборки, при ограниченной дисперсии генеральной совокупности, разность между выборочной средней ( $\tilde{x}$ ) и генеральной средней ( $\bar{x}$ ) будет сколь угодно мала.

Примечания: 1. Выражение  $\frac{\alpha}{\sqrt{n}}$  часто обозначают  $\mu$ .

2. При практическом использовании теоремы Чебышева генеральную дисперсию  $\alpha^2$ , которая неизвестна, заменяют выборочной дисперсией  $\sigma^2$ .



§ 2. *Теорема Ляпунова.* С вероятностью, равной  $F(t)$ , можно утверждать, что при достаточно большом объеме выборки и ограниченной дисперсии генеральной совокупности разность между выборочной средней и генеральной средней будет в пределах  $\pm t\mu = \pm t \frac{\sigma}{\sqrt{n}}$ .

Формулировка Ляпунова придает теореме Чебышева полную определенность и записывается:

$$P\left[(\tilde{x} - \bar{x}) < t \cdot \frac{\sigma}{\sqrt{n}}\right] = F(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} \cdot dt. \quad (2)$$

Замечание о практическом использовании то же, что и для формулы (1).

§ 3. *Теорема Бернулли* представляет собой частный случай теоремы Чебышева и может быть из нее получена

$$P\left[(w - p) < t \sqrt{\frac{pq}{n}}\right] > 1 - \frac{1}{t^2}, \quad (3)$$

где  $w$  — доля признака среди отобранных единиц (частость);  
 $p$  — доля признака в генеральной совокупности.

Теорема Бернулли охватывает случай, когда из генеральной совокупности производится отбор единиц и доля признака не меняется от испытания к испытанию. Формулировка теоремы Бернулли: с вероятностью, сколь угодно близкой к единице, можно утверждать, что разность между частостью и долей при достаточно большом объеме выборки будет сколь угодно мала. При практическом использовании данной теоремы расчет  $\sqrt{\frac{pq}{n}}$  производится заменой  $p$  на  $w$  и  $q$  на  $(1-w)$ .

§ 4. *Теорема Пуассона* также является частным случаем теоремы Чебышева для случая, когда доля признака в генеральной совокупности ( $p$ ) с ходом выборки все время меняется. В этом случае  $\sigma^2 = pq$ .

Тогда:

$$P\left[(w - p) < t \sqrt{\frac{pq}{n}}\right] > 1 - \frac{1}{t^2}. \quad (4)$$

§ 5. Ошибка репрезентативности (представительства) ( $\epsilon$ ) представляет собой разность между характеристиками выборочной и генеральной совокупности. Генеральная средняя ( $\bar{x}$ ) вычитается из выборочной средней ( $\tilde{x}$ ) или доля признака в генеральной совокупности ( $p$ ) вычитается из доли признака в выборочной совокупности, т. е. частости ( $w$ ):

$$d = \tilde{x} - \bar{x}; \quad d = w - p.$$

Если  $\Delta$  представляет собой предел, которого не превосходит абсолютная величина  $d$ , то

$$\tilde{x} - \Delta \leq \bar{x} \leq \tilde{x} + \Delta \quad (5)$$

$$w - \Delta \leq p \leq w + \Delta \quad (6)$$

§ 6. Производство отбора единиц из генеральной совокупности может мыслиться двумя способами.

При первом способе однажды отобранная и зарегистрированная единица может быть отобрана второй и третий раз. Такой способ отбора называется повторным.

При втором способе ни одна единица два раза не попадает в выборку. Такой способ отбора называется бесповторным и является более точным.

§ 7. В формулах выборочного метода фигурирует дисперсия генеральной совокупности ( $\sigma^2$ ). Но при производстве выборки характеристики генеральной совокупности неизвестны. Однако опыт применения выборочного метода показывает, что без большой погрешности можно заменить дисперсию генеральной совокупности дисперсией выборочной совокупности ( $\sigma^2$ ), которая вычисляется по формулам:

$$\left. \begin{aligned} \sigma^2 &= \frac{\sum (x - \bar{x})^2}{n}; & \text{Дисперсия выборочной} \\ & & \text{совокупности для величины} \\ & & \text{признака (невзвешенная)} \\ \sigma^2 &= \frac{\sum (x - \bar{x})^2 \cdot m}{\sum m}; & \text{Дисперсия выборочной} \\ & & \text{совокупности для величины} \\ & & \text{признака (взвешенная)} \\ \sigma^2 &= w(1 - w). & \text{Дисперсия выборочной со-} \\ & & \text{вокупности для частоты} \end{aligned} \right\} \quad (7)$$

§ 8. Теория устанавливает соотношение между пределом ошибки выборки ( $\Delta$ ), гарантируемым с некоторой вероятностью ( $P$ ), величиной  $t$ , связанной с этой вероятностью (см. приложение III),  $F(t)$ , и так называемой средней ошибкой выборки ( $\mu$ ):

$$\Delta = t \cdot \mu$$

или

$$t = \frac{\Delta}{\mu} \quad (8)$$

Предельная ошибка выборки равна  $t$ -кратному числу средних ошибок выборки.

§ 9. По способу организации выборки различают:

- 1) собственно случайный отбор;
- 2) типический отбор;

- 3) механический отбор;
- 4) серийный отбор;
- 5) комбинированный отбор.

Собственно случайный отбор ориентирован на выборку единиц из генеральной совокупности без всякого расчленения ее на части или группы. При этом теоретически возможно применение собственно случайного повторного отбора и собственно случайного бесповторного отбора.

Формулы средней ошибки выборки при собственно случайном методе отбора:

Т а б л и ц а 1

Средняя ошибка выборки ( $\mu$ )	П р и о т б о р е	
	повторном	бесповторном
Для средней	$\sqrt{\frac{\sigma^2}{n}}$	$\sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
Для доли	$\sqrt{\frac{w(1-w)}{n}}$	$\sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

(9)

Более точно, вместо множителя  $\left(1 - \frac{n}{N}\right)$  следует брать  $\frac{N-n}{N-1}$ , но при большой численности  $N$  различие между этими выражениями практически значения не имеет.

**Пример 1.** Из совокупности 10 000 деталей отобрано собственно случайным бесповторным методом 1000 деталей, для которых средний вес детали оказался равным 50 г, дисперсия 49. Бракованных деталей было обнаружено 20 штук. Вычислить средние ошибки выборки для средней и доли.

Дано:

$$N = 10\,000; \quad n = 1000; \quad \bar{x} = 50; \quad \sigma^2 = 49;$$

$$w = \frac{20}{1000} = 0,02.$$

По формулам (9) из табл. 1 для среднего веса детали при бесповторном отборе получаем:

$$\mu = \sqrt{\frac{49}{1000} \cdot \left(1 - \frac{1000}{10\,000}\right)} = 0,21$$

и для доли брака:

$$\mu = \sqrt{\frac{0,02 \cdot 0,98}{1000} \left(1 - \frac{1000}{10\,000}\right)} = 0,0042, \text{ или } 0,42\%.$$

§ 10. Для случая, когда частость даже приблизительно неизвестна, можно произвести «грубый» расчет средней ошибки выборки для доли, вводя в расчет максимальную величину произведения  $w(1-w)$ , равную 0,25. Тогда для повторного отбора получим:

$$\mu = \sqrt{\frac{w(1-w)}{n}} < \sqrt{\frac{0,25}{n}} = \frac{0,5}{\sqrt{n}} = \frac{1}{2\sqrt{n}}$$

и для бесповторного отбора:

$$\mu = \sqrt{\frac{w(1-w)}{n} \cdot \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,25}{n} \left(1 - \frac{n}{N}\right)} = \frac{1}{2\sqrt{n}} \cdot \sqrt{1 - \frac{n}{N}} \quad (10)$$

Пример 2. Из совокупности, численностью в 900 деталей, взята на выборку 81 деталь. Никаких данных, даже предположительных, об удельном весе деталей I сорта в генеральной совокупности нет.

Определяем среднюю ошибку выборки для доли продукции I сорта.

Дано  $N=900$ ;  $n=81$ ; допускаем, что  $w(1-w)=0,25$ , тогда по формуле (10):

$$\begin{aligned} \mu &= \frac{1}{2\sqrt{81}} \cdot \sqrt{1 - \frac{81}{900}} = \frac{1}{18} \cdot \sqrt{\frac{819}{900}} = \\ &= \frac{1}{18 \cdot 30} \sqrt{819} = \frac{28,62}{540} = 0,053, \text{ или } 5,3\%. \end{aligned}$$

§ 11. Как было показано в § 8,  $\Delta = t\mu$ , из приложения III возьмем три значения  $F(t)$  и  $t$ :

$$\text{при } t=1 \quad F(t)=0,683;$$

$$, \quad t=2 \quad F(t)=0,954;$$

$$, \quad t=3 \quad F(t)=0,997.$$

Это показывает, что 0,683 измеряет вероятность того, что ошибка выборки не превысит предела, равного одной средней ошибке. Значительно больше вероятность того, что ошибка не превысит двойной средней ошибки и т. д.

Вероятность 0,997 практически принимают за достоверность, т. е. считают, что предельная ошибка выборки равна трехкратной средней ошибке.

Иногда для определения размеров предельной ошибки связывают величину  $t$  с объемом выборки, применяя эмпирическую формулу:

$$t = 3 + \frac{6}{n-4}, \quad (11)$$

тогда

$$\Delta = \left(3 + \frac{6}{n-4}\right) \cdot \mu.$$

Чем больше объем выборки, тем ближе предельная ошибка к трем средним ошибкам.

§ 12. Проектировка выборочного наблюдения предполагает заранее заданными величину допустимой ошибки выборки и вероятность ответа. Незвестным, следовательно, остается тот минимальный объем выборки, который должен обеспечить требуемую точность. Из формул 8 устанавливаем необходимую численность выборки.

Формулы для определения численности выборки ( $n$ ) при собственно случайном способе отбора:

Таблица 2

Предполагаемый способ отбора	Ф о р м у л ы		
	для средней	для доли	для доли, если даже приблизительно она неизвестна
Повторный	$\frac{t^2 \sigma^2}{\Delta^2}$	$\frac{t^2 w (1-w)}{\Delta^2}$	$\frac{0,25 t^2}{\Delta^2}$
Бесповторный	$\frac{t^2 \sigma^2 N}{N \Delta^2 + t^2 \sigma^2}$	$\frac{t^2 N w (1-w)}{N \Delta^2 + t^2 w (1-w)}$	$\frac{0,25 t^2 N}{N \Delta^2 + 0,25 t^2}$

(12)

Примечание. При проектировании объема необходимой выборки величины  $\sigma^2$  и  $w$  неизвестны, поэтому вместо точного их значения берут приближенные, установленные на основании уже проведенного другого наблюдения, или нескольких пробных наблюдений, избирая из результатов наибольшие значения  $\sigma^2$  и  $w (1-w)$ .

Пример 3. Проектируется выборочное наблюдение, целью которого является установление среднего размера в совокупности 10 000 деталей. Требуемая точность 1 см. Произведенные пробные выборки дали наибольшую дисперсию, равную 49. Нужно определить необходимую численность случайной бесповтор-

ной выборки, обеспечивающей с вероятностью 0,95 заданную точность.

Дано:  $N = 10\,000$ ;

$\Delta = 1$ ;

$F(t) = 0,95$ ;

$\sigma^2 = 49$ .

По приложению III находим по  $F(t)$  значение  $t = 1,96$  и по формуле 12 для бесповторной выборки, взятой из табл. 2, получаем:

$$n = \frac{t^2 \sigma^2 N}{N \Delta^2 + t^2 \sigma^2} = \frac{1,96^2 \cdot 49 \cdot 10\,000}{10\,000 \cdot 1^2 + 1,96^2 \cdot 49} = 184,7 \approx 185.$$

§ 13. Типический отбор дает более точные результаты. Генеральная совокупность делится по некоторому признаку на типические группы. Количество отбираемых единиц из каждой типической группы устанавливается в следующих размерах:

При отборе, не пропорциональном объему типических групп, общее число отбираемых единиц делится на число типических групп и полученная величина дает численность отбора из каждой типической группы.

При отборе, пропорциональном объему типических групп, число наблюдений по каждой группе определяется по формуле:

$$n_i = n \cdot \frac{N_i}{N}, \quad (13)$$

где  $n_i$  — объем выборки из  $i$ -й типической группы;

$n$  — общий объем выборки;

$N_i$  — объем  $i$ -й типической группы;

$N$  — объем генеральной совокупности.

При отборе с учетом колеблемости признака, дающем наименьшую величину ошибки выборки, процент выборки из каждой типической группы должен быть пропорционален среднему квадратическому отклонению в этой группе ( $\sigma_i$ ). Расчет численности ( $n_i$ ) производится по формулам:

$$\left. \begin{aligned} n_i &= \frac{n N_i \sigma_i}{\sum N_i \sigma_i} \\ n_i &= \frac{n N_i \sqrt{w_i (1 - w_i)}}{\sum N_i \sqrt{w_i (1 - w_i)}} \end{aligned} \right\} \quad (14)$$

Формулы средних ошибок выборки ( $\mu$ ) при типическом методе отбора

Способ отбора	Повторный		Бесповторный	
	для средней	для доли	для средней	для доли
Не пропорциональный объем	$\frac{1}{N} \sqrt{\sum \frac{\sigma_i^2}{n_i} \cdot N_i^2}$	$\frac{1}{N} \sqrt{\sum \frac{w_i(1-w_i)}{n_i} N_i^2}$	$\frac{1}{N} \sqrt{\sum \frac{\sigma_i^2}{n_i} N_i^2 \left(1 - \frac{n_i}{N_i}\right)}$	$\frac{1}{N} \sqrt{\sum \frac{w_i(1-w_i)}{n_i} \cdot N_i^2 \left(1 - \frac{n_i}{N_i}\right)}$
Пропорциональный объем	$\sqrt{\frac{\sigma^2}{n}}$	$\sqrt{\frac{w(1-w)}{n}}$	$\sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$	$\sqrt{\frac{w(1-w)}{n} \cdot \left(1 - \frac{n}{N}\right)}$
Пропорциональный коэффициент	$\frac{1}{N} \cdot \frac{\sum \sigma_i N_i}{\sqrt{n_i}}$	$\frac{1}{N} \cdot \frac{\sum w_i(1-w_i) N_i}{\sqrt{n_i}}$	$\frac{1}{N} \cdot \frac{\sum \sigma_i N_i}{\sqrt{n_i}} \cdot \sqrt{1 - \frac{n}{N}}$	$\frac{1}{N} \cdot \frac{\sum w_i(1-w_i) \cdot N_i}{\sqrt{n_i}} \cdot \sqrt{1 - \frac{n}{N}}$

где  $\sigma^2$  — средняя из выборочных дисперсий типических групп;

$w(1-w)$  — средняя из произведения частот на дополнение их до единицы;

$\sigma_i^2$  — выборочная дисперсия  $i$ -й типической группы;

$\sigma_i$  — среднее квадратическое отклонение в выборке из  $i$ -й типической группы.

<sup>1</sup> Наиболее удобные пропорции отбора.

Пример 4<sup>1</sup>. Из совокупности 10 000 единиц производится выборка типическим методом для определения средней. Вся совокупность делится на 5 типических групп. Отбор единиц внутри типических групп производится случайным бесповторным методом пропорционально объему каждой группы. Отбирается 2000 единиц. При отборе получены следующие результаты:

Таблица 3

Типические группы (i)	Численность групп (N <sub>i</sub> )	Численность выборки (n <sub>i</sub> )	Выборочная средняя ( $\tilde{x}_i$ )	Выборочная дисперсия ( $\sigma_i^2$ )
1	2	3	4	5
1	500	100	10	4
2	2 000	400	12	6
3	3 000	600	15	10
4	1 500	300	18	20
5	3 000	600	20	16
Всего . . .	10 000	2 000	16,1	22,89

Вычислить: а) среднюю ошибку для каждой группы и для всей выборочной совокупности (при собственно случайном и типическом способах отбора); б) границы, в которых с вероятностью 0,997 находится генеральная средняя по группам и по всей совокупности (при собственно случайном и типическом методах отбора).

Прежде всего рассчитывают численность отбираемых единиц из каждой типической группы пропорционально ее объему (см. колонку 3 табл. 3). Так, для первой типической группы имеем по формуле 13 при заданном объеме всей выборки, равном 2000 единиц:

$$n_1 = 2000 \cdot \frac{500}{10\,000} = 100 \text{ единиц};$$

для второй типической группы:

$$n_2 = 2000 \cdot \frac{2000}{10\,000} = 400 \text{ единиц}$$

и т. д.

Для определения средней ошибки выборки по группам и общей средней ошибки выборки при собственно случайном способе

<sup>1</sup> Примеры 4 и 5 взяты из неопубликованного задачника по выборочному методу П. П. Шушерины.



отбора (бесповторном) используем формулы 9. Получаем среднюю ошибку выборки для первой типической группы:

$$\mu_1 = \sqrt{\frac{4}{100} \cdot \left(1 - \frac{100}{500}\right)} \approx 0,18;$$

для второй типической группы:

$$\mu_2 = \sqrt{\frac{6}{400} \left(1 - \frac{400}{2000}\right)} \approx 0,11$$

и т. д. по всем группам (см. колонку 2 табл. 4).

Для удобства располагаем все получаемые результаты в таблице.

Таблица 4

Типические группы		Средняя ошибка выборки ( $\mu$ )	Предельная ошибка вы- борки с ве- роятностью 0,997 ( $\Delta$ )	Границы, в которых с вероятностью 0,997 на- ходится генеральная средняя	
				от	до
1		2	3	4	5
1		0,18	0,54	9,46	10,54
2		0,11	0,33	11,77	12,33
3		0,12	0,36	14,64	15,36
4		0,23	0,69	17,31	18,69
5		0,15	0,45	19,55	20,45
Итого	При собст- венно слу- чайном методе	0,096	0,29	15,81	16,39
	При типи- ческом методе	0,07	0,21	15,89	16,31

Для расчета средней ошибки выборки всей совокупности при собственно случайном методе отбора и границ генеральной средней при этом же методе отбора нужно знать общую выборочную среднюю и общую дисперсию выборочной совокупности. Производим расчет общей выборочной средней из групповых выборочных средних путем взвешивания последних по численности отобранных групп.

$$\tilde{x} = \frac{\sum \tilde{x}_i \cdot n_i}{\sum n_i} = \frac{10 \cdot 100 + 12 \cdot 400 + 15 \cdot 600 + 18 \cdot 300 + 20 \cdot 600}{2000} = 16,1$$

(см. итог колонки 4 табл. 3).

Для определения общей выборочной дисперсии используют теорему сложения вариации.

Находим сначала среднюю взвешенную из выборочных дисперсий:

$$\bar{\sigma}_i^2 = \frac{\sum \sigma_i^2 \cdot n_i}{\sum n_i} = \frac{4 \cdot 100 + 6 \cdot 400 + 10 \cdot 600 + 20 \cdot 300 + 16 \cdot 600}{2000} = 12,2,$$

а затем межгрупповую дисперсию

$$\begin{aligned} \bar{\delta}^2 &= \frac{\sum (\bar{x}_i - \bar{x})^2 n_i}{\sum n_i} = \\ &= \frac{(-6,1)^2 \cdot 100 + (-4,1)^2 \cdot 400 + (-1,1)^2 \cdot 600 + (1,9)^2 \cdot 300 + (3,9)^2 \cdot 600}{2000} = \\ &= 10,69. \end{aligned}$$

Получаем общую дисперсию выборочной совокупности

$$\sigma^2 = \bar{\sigma}_i^2 + \bar{\delta}^2 = 12,2 + 10,69 = 22,89$$

(см. итог по колонке 5 табл. 3).

Находим среднюю ошибку выборки всей совокупности при собственно случайном методе отбора.

$$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{22,89}{2000} \left(1 - \frac{2000}{10\,000}\right)} \approx 0,096$$

(см. первую строку итога колонки 2 табл. 4).

Предельная ошибка собственно случайной выборки:

$$\Delta = t\mu \approx 3 \cdot 0,096 = 0,29$$

(см. первую строку итога колонки 3 табл. 4).

Соответственно находим границы генеральной средней при собственно случайном методе отбора:

$$\bar{x} = \bar{x} \pm \Delta = 16,1 \pm 0,29;$$

$$15,81 < \bar{x} < 16,39$$

(см. первую строку итога колонок 4 и 5 табл. 4).

Рассчитываем среднюю ошибку типической выборки, пропорциональной объему типических групп, по формуле 15.

Получим:

$$\mu = \sqrt{\frac{\sigma_i^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{12,2}{2000} \left(1 - \frac{2000}{10\,000}\right)} \approx 0,070$$

(см. вторую строку итога колонки 2 табл. 4).

Далее определяем ошибку типической выборки  $\Delta = 3 \cdot 0,07 = 0,21$  и границы генеральной средней  $\bar{x} = 16,1 \pm 0,21$ , т. е.

$$15,89 < \bar{x} < 16,31$$

(см. вторую строку итога колонок 4 и 5 табл. 4).

**Пример 5.** Для определения доли признака производится типическая выборка 400 единиц из совокупности 10 500 единиц, разбитых на 3 типические группы, численностью в 5000, 2500 и 3000 единиц. Имеются основания (прошрое обследование) считать, что искомая доля по типическим группам составляет около 10%; 20% и 50%.

В каком объеме произвести выборку из типических групп, чтобы пропорции отбора были наивыгоднейшими?

Определяем численность первой типической группы по формуле 14, при объеме всей выборки, равной 400 единицам:

для

$$\begin{aligned} n_1 &= \frac{400 \cdot 5000 \cdot \sqrt{0,1 \cdot 0,9}}{5000 \cdot \sqrt{0,1 \cdot 0,9} + 2500 \cdot \sqrt{0,2 \cdot 0,8} + 3000 \cdot \sqrt{0,5 \cdot 0,5}} = \\ &= \frac{600\,000}{1500 + 1000 + 1500} = \frac{600\,000}{4000} = 150; \end{aligned}$$

для

$$n_2 = \frac{400\,000}{4000} = 100;$$

для

$$n_3 = \frac{600\,000}{4000} = 150.$$

§ 14. При механической выборке совокупность механически делится на столько групп, сколько единиц должно войти в выборку, и из каждой группы отбирается одна единица.

Средняя ошибка выборки подсчитывается по формулам 9 собственно случайной выборки.

§ 15. При серийном отборе с равновеликими сериями генеральную совокупность делят на одинаковые по объему группы — серии и производят выборку не единиц совокупности, а серий. Попавшие в выборку серии обследуются сплошь. Серии могут отбираться повторным и бесповторным методом.

Средние ошибки выборки при таком отборе рассчитывают по формулам:

Способ отбора серий	Ф о р м у л ы	
	для средней	для доли
Повторный	$\sqrt{\frac{\delta_x^2}{r}}$	$\sqrt{\frac{\delta_p^2}{r}}$
Бесповторный	$\sqrt{\frac{\delta_x^2}{r} \left( \frac{R-r}{R-1} \right)} \approx$ $\approx \sqrt{\frac{\delta_x^2}{r} \left( 1 - \frac{r}{R} \right)}$	$\sqrt{\frac{\delta_p^2}{r} \left( \frac{R-r}{R-1} \right)} \approx$ $\approx \sqrt{\frac{\delta_p^2}{r} \left( 1 - \frac{r}{R} \right)}$

(16)

где  $R$  — число серий в генеральной совокупности;

$r$  — число отобранных серий;

$\delta_x^2$  — межсерийная (межгрупповая) дисперсия средних;

$\delta_p^2$  — межсерийная (межгрупповая) дисперсия доли.

**Пример 6.** Генеральная совокупность состоит из 500 единиц, разбитых на 50 равных по величине серий (по 100 единиц). Бесповторным методом отобрано 10 серий. Результаты выборки представлены в следующей таблице:

Таблица 5

Номера серий	1	2	3	4	5	6	7	8	9	10
Средний размер признака ( $\tilde{x}_i$ )	51	54	60	61	62	70	73	77	79	92
Частная дисперсия ( $\sigma_i^2$ )	25,0	23,2	16,1	35,9	45,8	91,2	88,3	77,8	110,5	145,1

Исчислить среднюю ошибку серийной бесповторной выборки.

Вычисляем: а) общую среднюю всей выборочной совокупности по серийным средним:

$$\tilde{x} = \frac{\sum \tilde{x}_i}{n} = \frac{679}{10} = 67,9;$$

б) межсерийную (межгрупповую) дисперсию средних:

$$\begin{aligned} \bar{\delta}_x^2 &= \frac{\sum (\tilde{x}_i - \tilde{x})^2}{n} = \\ &= \frac{(51 - 67,9)^2 + (54 - 67,9)^2 + \dots + (92 - 67,9)^2}{10} = \frac{1440,9}{10} = 144,09; \end{aligned}$$

в) среднюю ошибку серийной выборки:

$$\mu = \sqrt{\frac{\bar{\delta}_x^2}{r} \left( \frac{R-r}{R-1} \right)} = \sqrt{\frac{144,09}{10} \cdot \left( \frac{50-10}{50-1} \right)} = \\ = \sqrt{11,7624} \approx 3,43.$$

§ 16. Необходимая численность отбираемых серий при серийном отборе получается использованием формул 12, в которых вместо  $N$ ,  $n$  и  $\bar{\delta}^2$  подставляют  $R$ ,  $r$  и  $\bar{\delta}^2$ .

Пример 7. Совокупность разбита на 50 серий. Имеются основания предполагать, что межсерийная дисперсия равна 16. Сколько серий нужно отобрать бесповторным методом, чтобы с вероятностью 0,954 утверждать, что ошибка выборочной средней не превысит 2,3.

Дано:

$$\bar{\delta}^2 = 16; \quad R = 50; \quad t = 2; \quad \Delta = 2,3.$$

Находим необходимое число серий, отбор которых обеспечит требуемую точность:

$$r = \frac{t^2 \bar{\delta}^2 R}{R \Delta^2 + t^2 \bar{\delta}^2} = \frac{4 \cdot 16 \cdot 50}{50 \cdot 5,29 + 4 \cdot 16} = \frac{3200}{328,5} \approx 10 \text{ серий}$$

§ 17. Комбинированная выборка (равновеликие серии) предполагает комбинацию серийного отбора с индивидуальным отбором.

Генеральная совокупность разбивается на одинаковые по объему серии. Сначала отбираются серии, а затем из отобранных серий производится индивидуальная выборка единиц.

Квадрат средних ошибок выборки ( $\mu^2$ ) рассчитывают по формулам:

Способ отбора	Ф о р м у л ы	
	для средней	для доли
Повторный	$\frac{\bar{\sigma}_i^2}{n} + \frac{\bar{\delta}_x^2}{r}$	$\frac{\bar{w}_i(1-\bar{w}_i)}{n} + \frac{\bar{\delta}_p^2}{r}$
Бесповторный	$\frac{\bar{\sigma}_i^2}{n} \left( 1 - \frac{n}{n_r} \right) + \frac{\bar{\delta}_x^2}{r} \left( 1 - \frac{r}{R} \right)$	$\frac{\bar{w}_i(1-\bar{w}_i)}{n} \left( 1 - \frac{n}{n_r} \right) + \frac{\bar{\delta}_p^2}{r} \left( 1 - \frac{r}{R} \right)$

(17)

где  $n_r$  — общее число единиц, попавших в выборку при отборе серий; определяется по формуле:

$$n_r = r \frac{N}{R}, \quad (18)$$

$n$  — число единиц, попавших в выборку из серий.

**Пример 8.** Генеральная совокупность состоит из 100 000 единиц, разбитых на 200 равных по объему серий. Произведена бесповторная выборка 50% серий и из каждой серии по 20% единиц. Средняя из серийных дисперсий оказалась равной 12, а межсерийная дисперсия — 5. Определить среднюю ошибку выборки.

Дано:

$N = 100\,000$ ;  $R = 200$ ;  $r = 100$ ;  $\bar{s}_i^2 = 12$ ;  $\delta^2 = 5$ ;  $n = 10\,000$ ;  
получаем по формуле 18

$$n_r = 100 \cdot \frac{100\,000}{200} = 50\,000.$$

$$\mu = \sqrt{\frac{12}{10\,000} \left(1 - \frac{10\,000}{50\,000}\right) + \frac{5}{100} \left(1 - \frac{100}{200}\right)} = 0,161.$$

[По формуле (17) для бесповторного отбора].

Мы получили среднюю ошибку комбинированной выборки при отборе из генеральной совокупности 10 000 единиц. Можно было бы произвести выборку такого же объема, но отобрав 20% серий и по 50% единиц внутри серий.

При тех же значениях — средней из серийных дисперсий и межсерийной дисперсии — средняя ошибка выборки получилась бы

$$\mu = \sqrt{\frac{12}{10\,000} \left(1 - \frac{10\,000}{20\,000}\right) + \frac{5}{40} \left(1 - \frac{40}{200}\right)} = 0,317.$$

Таким образом, величина ошибки увеличилась бы больше чем в два раза.

В иных случаях большая точность достигается большим числом наблюдений в пределах отобранных серий за счет сокращения числа последних.

§ 18. Выборочная средняя отличается от генеральной средней на  $t$  кратное число средних ошибок ( $\mu$ ). Если в результате выборов получены две выборочные средние ( $x_1$  и  $x_2$ ), для каждой из которых найдена средняя ошибка выборки ( $\mu_1$  и  $\mu_2$ ), то среднюю ошибку разности этих двух выборочных средних ( $\mu_{разн}$ ) можно определить по средним ошибкам этих выборочных средних.

$$\mu_{разн} = \sqrt{\mu_1^2 + \mu_2^2 - 2\mu_1 \cdot \mu_2 \cdot R}, \quad (19)$$

где  $R$  — коэффициент корреляции между вариантами двух выборочных совокупностей (см. раздел V).

В случае некоррелированности признаков, т. е. равенства коэффициента корреляции нулю, формула примет вид:

$$\mu_{разн} = \sqrt{\mu_1^2 + \mu_2^2}. \quad (19a)$$

**Пример 9.** Из генеральной совокупности произведены две выборки. При этом средние ошибки выборочных средних оказались равными 0,48 и 0,43. Признаки некоррелированы. Найти среднюю ошибку разности двух выборочных средних:

$$\mu_{разн} = \sqrt{0,48^2 + 0,43^2} = \sqrt{0,4153} \approx 0,64.$$

§ 19. В качестве критерия оценки существенности расхождения между двумя выборочными средними принимается неравенство:

$$\frac{\tilde{x}_1 - \tilde{x}_2}{\mu_{разн}} > 3. \quad (20)$$

**Пример 10.** Выделено 5 участков лесонасаждений, и с каждого участка взяты пробные площадки. В среднем на 1 га по пяти участкам получилось следующее распределение деревьев по толщине:

**Таблица 6**

Диаметр на высоте груди в см (x)	Число деревьев по участкам на 1 га				
	$m^I$	$m^{II}$	$m^{III}$	$m^{IV}$	$m^V$
1	2	3	4	5	6
12	88	107	150	232	381
16	73	84	125	193	222
20	62	72	98	136	106
24	58	61	72	87	41
28	47	46	44	42	14
32	36	32	27	18	2
36	26	22	16	8	1
40	18	13	8	3	—
44	11	7	4	1	—
48	6	4	1	—	—
52	4	1	1	—	—
56	1	1	—	—	—
60	1	1	—	—	—
Итого . .	431	451	546	720	767

Определить существенность расхождения средних диаметров деревьев по участкам.

а. Находим средние диаметры деревьев по участкам:

$$\text{I. } \tilde{x}^{\text{I}} = \frac{\Sigma x m^{\text{I}}}{\Sigma m^{\text{I}}} = 23,4;$$

$$\text{II. } \tilde{x}^{\text{II}} = \frac{\Sigma x m^{\text{II}}}{\Sigma m^{\text{II}}} = 21,8;$$

$$\text{III. } \tilde{x}^{\text{III}} = \frac{\Sigma x m^{\text{III}}}{\Sigma m^{\text{III}}} = 19,7;$$

$$\text{IV. } \tilde{x}^{\text{IV}} = \frac{\Sigma x m^{\text{IV}}}{\Sigma m^{\text{IV}}} = 17,9;$$

$$\text{V. } \tilde{x}^{\text{V}} = \frac{\Sigma x m^{\text{V}}}{\Sigma m^{\text{V}}} = 15,3.$$

б. Вычисляем средние квадратические отклонения по участкам:

$$\text{I. } \sigma^{\text{I}} = \sqrt{\frac{\Sigma (x - \tilde{x}^{\text{I}})^2 m^{\text{I}}}{\Sigma m^{\text{I}}}} = 9,9;$$

$$\text{II. } \sigma^{\text{II}} = \sqrt{\frac{\Sigma (x - \tilde{x}^{\text{II}})^2 m^{\text{II}}}{\Sigma m^{\text{II}}}} = 9,1;$$

$$\text{III. } \sigma^{\text{III}} = \sqrt{\frac{\Sigma (x - \tilde{x}^{\text{III}})^2 m^{\text{III}}}{\Sigma m^{\text{III}}}} = 7,5;$$

$$\text{IV. } \sigma^{\text{IV}} = \sqrt{\frac{\Sigma (x - \tilde{x}^{\text{IV}})^2 m^{\text{IV}}}{\Sigma m^{\text{IV}}}} = 5,9;$$

$$\text{V. } \sigma^{\text{V}} = \sqrt{\frac{\Sigma (x - \tilde{x}^{\text{V}})^2 m^{\text{V}}}{\Sigma m^{\text{V}}}} = 4,1.$$

в. Вычисляем средние ошибки выборочных средних

$$\text{I. } \mu^{\text{I}} = \frac{\sigma^{\text{I}}}{\sqrt{n^{\text{I}}}} = 0,48;$$

$$\text{II. } \mu^{\text{II}} = \frac{\sigma^{\text{II}}}{\sqrt{n^{\text{II}}}} = 0,43;$$



$$\text{III. } \mu^{\text{III}} = \frac{\sigma^{\text{III}}}{\sqrt{n^{\text{III}}}} = 0,33;$$

$$\text{IV. } \mu^{\text{IV}} = \frac{\sigma^{\text{IV}}}{\sqrt{n^{\text{IV}}}} = 0,22;$$

$$\text{V. } \mu^{\text{V}} = \frac{\sigma^{\text{V}}}{\sqrt{n^{\text{V}}}} = 0,14.$$

г. Находим, например, следующие разности выборочных средних по участкам:

$$\text{I. } \tilde{x}^{\text{II}} - \tilde{x}^{\text{I}} = 1,6;$$

$$\text{II. } \tilde{x}^{\text{III}} - \tilde{x}^{\text{II}} = 2,1;$$

$$\text{III. } \tilde{x}^{\text{IV}} - \tilde{x}^{\text{III}} = 1,8;$$

$$\text{IV. } \tilde{x}^{\text{V}} - \tilde{x}^{\text{IV}} = 2,6.$$

д. Находим средние ошибки разности соответствующих пар выборочных средних:

$$\text{I. } \mu_{\text{разн}}^{\text{I}} = \sqrt{(\mu^{\text{I}})^2 + (\mu^{\text{II}})^2} = 0,64;$$

$$\text{II. } \mu_{\text{разн}}^{\text{II}} = \sqrt{(\mu^{\text{II}})^2 + (\mu^{\text{III}})^2} = 0,54;$$

$$\text{III. } \mu_{\text{разн}}^{\text{III}} = \sqrt{(\mu^{\text{III}})^2 + (\mu^{\text{IV}})^2} = 0,4;$$

$$\text{IV. } \mu_{\text{разн}}^{\text{IV}} = \sqrt{(\mu^{\text{IV}})^2 + (\mu^{\text{V}})^2} = 0,26.$$

е. Находим критерий оценки существенности расхождения соответствующих выборочных средних:

$$\text{I. } \frac{\tilde{x}^{\text{II}} - \tilde{x}^{\text{I}}}{\mu_{\text{разн}}^{\text{I}}} = \frac{1,6}{0,64} = 2,5;$$

$$\text{II. } \frac{\tilde{x}^{\text{III}} - \tilde{x}^{\text{II}}}{\mu_{\text{разн}}^{\text{II}}} = \frac{2,1}{0,54} = 3,9;$$

$$\text{III. } \frac{\tilde{x}^{\text{IV}} - \tilde{x}^{\text{III}}}{\mu_{\text{разн}}^{\text{III}}} = \frac{1,8}{0,4} = 4,5;$$

$$\text{IV. } \frac{\tilde{x}^{\text{V}} - \tilde{x}^{\text{IV}}}{\mu_{\text{разн}}^{\text{IV}}} = \frac{2,6}{0,26} = 10,0.$$

ж. **Вывод.** Из критериев оценки существенности заключаем, что выделения II, III, IV и V участков произведены правильно, так как критерии оценки существенности больше трех. И, следовательно, мы имеем разные насаждения.

При сравнении I и II участков вопрос остается открытым.

§ 20. *Малая выборка.* При необходимости оценки генеральной совокупности по результатам малого числа наблюдений, т. е. при  $n$  меньше 20, формулы для обычной (большой) выборки, основанные на нормальном распределении вероятностей, дают значительные неточности.

Оценка результатов малой выборки производится путем «исправления» выборочной дисперсии и использования распределения вероятностей по Стюденту.

*Выборочная дисперсия* малой выборки исчисляется по формуле:

$$\sigma_{м. в} = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}, \quad (21)$$

т. е. она отличается от выборочной дисперсии  $\sigma$  тем, что сумму квадратов отклонений от выборочной средней делят не на  $n$ , а на  $n-1$ . Зная выборочную дисперсию  $\sigma$ , можно путем ее «исправления» вычислить выборочную дисперсию малой выборки  $\sigma_{м. в}$  по формуле:

$$\sigma_{м. в} = \sigma \sqrt{\frac{n}{n-1}}. \quad (21a)$$

**Пример 11.** Произведена выборка 16 единиц. Выборочная дисперсия ( $\sigma$ ) оказалась равной 100. Вычислить выборочную дисперсию малой выборки ( $\sigma_{м. в}$ ):

$$\sigma_{м. в} = 100 \cdot \sqrt{\frac{16}{15}} = 103,28.$$

*Средняя ошибка* малой выборки исчисляется по формуле:

$$\mu_{м. в} = \frac{\sigma_{м. в}}{\sqrt{n}} = \sqrt{\frac{\sum (x - \bar{x})^2}{n(n-1)}}. \quad (22)$$

**Пример 12.** Продолжая пример 11, можно вычислить среднюю ошибку малой выборки:

$$\mu_{м. в} = \frac{103,28}{\sqrt{16}} \approx 25,82.$$

Среднюю ошибку малой выборки можно получить и использованием «неисправленной» выборочной дисперсии:

$$\mu_{м. в} = \frac{\sigma}{\sqrt{n-1}}; \quad (22a)$$

Среднюю ошибку разности двух выборочных средних исчисля-  
ют по формуле:

$$\mu_{м.с. разн} = \sqrt{\frac{[\Sigma (x - \tilde{x}^I)^2 + \Sigma (x - \tilde{x}^{II})^2] (n_1 + n_2)}{(n_1 + n_2 - 2) n_1 n_2}}. \quad (23)$$

Нормированное отклонение или стандартизованная разность  
малой выборки ( $t$ ) получается аналогично тому, как это полу-  
чалось в обычной выборке:

$$t = \frac{\tilde{x} - \bar{x}}{\mu_{м.с.}} = \frac{\Delta_{м.с.}}{\mu_{м.с.}}. \quad (24)$$

Предельная ошибка малой выборки:

$$\Delta_{м.с.} = t \cdot \mu_{м.с.} \quad (25)$$

Опираясь на предположение о нормальном распределении  
признака в генеральной совокупности, Стюдент нашел закон рас-  
пределения  $t$ , который называется *распределением Стюдента*:

$$P_{(t)} = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)} \cdot \Gamma\left(\frac{n-1}{2}\right)} \cdot \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}} = S(t), \quad (26)$$

где  $P_{(t)} = S_{(t)}$  — вероятности, что стандартизованная разность между  
выборочной и генеральной средней имеет величину  $t$ ;

$\Gamma\left(\frac{n}{2}\right)$  и  $\Gamma\left(\frac{n-1}{2}\right)$  — гамма функции.

Для любого положительного числа  $n$  гамма функция опреде-  
ляется следующим равенством:

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx = (n-1)!$$

Частные случаи:

$\Gamma(1) = 1$ ;  $\Gamma(2) = 1$ ;  $\Gamma(3) = 2! = 2$ ;  $\Gamma(4) = 3! = 6$  и т. д.

Свойство гаммы функции:

$$\Gamma(n+1) = n \cdot \Gamma(n).$$

Первый частный случай гаммы функции и указанное ее свой-  
ство дают:

$$\Gamma(1) = 0! = 1.$$

Особенностью распределения Стюдента является то, что вероятность того или иного значения  $t$  зависит только от двух величин: объема выборки ( $n$ ) и величины  $t$ . При возрастании объема выборки распределение Стюдента приближается к нормальному

$$P(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Если сделать определенные допущения о величине генеральной средней, то можно вычислить *фактическое нормированное отношение* ( $t_\phi$ ) при помощи интеграла Стюдента:

$$t_\phi = \frac{\bar{x} - \tilde{x}}{\mu_{\text{м.с}}}, \quad (27)$$

тогда:

$$P[t < t_\phi] = C \cdot \int_{-t}^{+t} \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}} \cdot dt = S(t_\phi),$$

$$\text{где } C = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)} \cdot \Gamma\left(\frac{n-1}{2}\right)};$$

$S(t_\phi) = P[t < t_\phi]$  — вероятность того, что стандартизованная разность ( $t$ ) между действительной генеральной средней и выборочной средней будет меньше стандартизованной разности, вычисленной по результатам малой выборки ( $t_\phi$ );

$S(t_\phi)$  — определяется из приложения IV. При этом значение  $n$  определяется вычитанием единицы из числа наблюдений.

Интеграл Стюдента используют для решения ряда обычных задач малой выборки как для случаев, когда генеральная совокупность обладает нормальным распределением, так и для случаев, когда распределение признака в генеральной совокупности не совсем совпадает с нормальным.

Функция  $S(t_\phi)$  используется для определения также и вероятностей того, что: 1)  $t > t_\phi$ ; 2)  $|t| > t_\phi$  и 3)  $|t| < t_\phi$ :

$$P[t > t_\phi] = 1 - S(t_\phi), \quad (27a)$$

где  $P[t > t_\phi]$  — вероятность значений  $t$ , больших, чем  $t_\phi$ ;

$$P[|t| > t_\phi] = 2[1 - S(t_\phi)], \quad (27б)$$

где  $P[|t| > t_\phi]$  — вероятность значений  $t$  по абсолютной величине, больших, чем  $t_\phi$ ;

$$P[|t| < t_\phi] = 2S(t_\phi) - 1, \quad (27в)$$

где  $P[|t| < t_\phi]$  — вероятность значений  $t$  по абсолютной величине, меньших, чем  $t_\phi$ .

**Пример 13.** Первая типовая задача малой выборки. Оценка выборочной средней.

Произведена малая выборка урожая пшеницы. Срок уборки урожая своевременный. На выборку собственно случайным повторным методом взято 8 участков. Результаты выборки по отдельным участкам следующие:

**Т а б л и ц а 7**

Участки	Урожай пшеницы в ц с га ( $x'$ )	( $x'$ ) <sup>2</sup>
1	16,5	272,25
2	16,2	262,44
3	18,9	357,21
4	20,1	404,01
5	19,3	372,49
6	10,1	102,01
7	12,8	163,84
8	15,0	225,00
Итого . . . . .	128,9	2159,25

Определить вероятность того, что разность между выборочным и генеральным средним урожаем не больше 0,5 ц с га.

Дано:

$$n = 8;$$

$$\tilde{x}' = \frac{128,9}{8} = 16,113;$$

$$\tilde{x}' - \bar{x} = |0,5|.$$

Находим  $\sigma^2$  по формуле (54) (см. раздел I):

$$\begin{aligned} \sigma^2 &= \overline{(x')^2} - (\bar{x}')^2 = \frac{2159,25}{8} - (16,113)^2 = 269,906 - \\ &\quad - 259,629 = 10,277. \end{aligned}$$

Определяем  $\sigma = \sqrt{10,277} = 3,2057$ .

„Исправляем“  $\sigma$  и получаем:

$$\sigma_{\text{исп.}} = \sigma \sqrt{\frac{n}{n-1}} = 3,2057 \cdot \sqrt{\frac{8}{7}} = 3,2057 \cdot 1,069 = 3,427.$$

Вычисляем среднюю ошибку малой выборки ( $\mu_{м.с}$ ).

$$\mu_{м.с} = \frac{\sigma_{м.с}}{\sqrt{n}} = \frac{3,427}{\sqrt{8}} = 1,212.$$

Определяем величину нормированного отклонения по выборочным данным и предполагаемым границам генеральной средней ( $t_{\phi}$ ):

$$t_{\phi} = \frac{\tilde{x}' - \bar{x}}{\mu_{м.с}} = \frac{0,5}{1,212} = 0,412.$$

По формуле 27б находим:

$$P[|t| > 0,412] = 2[1 - S(0,412)].$$

Так как число наблюдений равно 8, то берем  $n=7$ ;

тогда по приложению IV находим:  $S(0,4) = 0,649$ .

Следовательно:

$$P[|t| > 0,412] = 2(1 - 0,649) = 2 \cdot 0,351 = 0,702 \approx 0,7.$$

Таким образом, видно, что вероятность нормированных отклонений, по абсолютной величине превышающих 0,412, или, иными словами, вероятность отклонений генеральной средней от выборочной средней на абсолютную величину, превышающую 0,5  $\mu$  с га, не мала (0,7). Поэтому разность между генеральной и выборочной средними легко могла превысить 0,5  $\mu$  с га.

Можно было воспользоваться формулой (27в) и определить вероятность нормированных отклонений по своей абсолютной величине, меньших, чем 0,412:

$$\begin{aligned} P[|t| < 0,412] &= 2 \cdot S(0,412) - 1 = 2 \cdot 0,649 = \\ &= 1,298 - 1 \approx 0,3 \end{aligned}$$

и прийти к тому же заключению.

Вероятность того, что генеральная средняя находится в определенных границах, определяется по формуле:

$$P[\tilde{x} - t_{\phi} \mu_{м.с} < \bar{x} < \tilde{x} + t_{\phi} \mu_{м.с}] = 2 \cdot S(t_{\phi}) - 1. \quad (28)$$

Пример 14. Вторая типовая задача малой выборки. Нахождение границ интервала, в котором находится генеральная средняя.

Из данных предыдущего примера 13 найти с вероятностью 0,954 границы интервала, в котором содержится генеральная средняя урожая.

Дано:

$$P = 0,954;$$

$$\mu_{\text{м. в.}} = 1,212.$$

Находим  $t_\phi$  по формуле (28):

$$2 \cdot S(t_\phi) - 1 = 0,954;$$

$$2 \cdot S(t_\phi) = 1,954;$$

$$S(t_\phi) = 0,977.$$

По приложению IV находим  $t_\phi$ , равное 2,5.

Следовательно, границы генеральной средней  $\bar{x} = \tilde{x} \pm t_\phi \cdot \mu_{\text{м. в.}} = 16,113 \pm 2,5 \cdot 1,212 = 16,113 \pm 2,05$ .

Окончательно: с вероятностью 0,954 можно утверждать, что  $13,06 \leq \bar{x} \leq 19,16$ .

Теория малой выборки дает возможность оценить существенность разности между двумя выборочными средними. Вероятность значений разностей между двумя выборочными средними по абсолютной величине, не меньших, чем разность, полученная в результате опыта, т. е. фактическая, определяется по формуле:

$$P[|\tilde{x}' - \tilde{x}''| > \delta_\phi] = 2[1 - S(t_\phi)], \quad (29)$$

где  $\tilde{x}'$  и  $\tilde{x}''$  — выборочные средние;

$\delta_\phi = \tilde{x}' - \tilde{x}''$  — фактическая разность между двумя выборочными средними;  
а величина  $t_\phi$  определяется по формуле:

$$t_\phi = \frac{\delta_\phi}{\mu_{\text{м. в. разн}}} = \frac{\delta_\phi}{\sqrt{\frac{[\Sigma(x' - \tilde{x}')^2 + \Sigma(x'' - \tilde{x}'')^2] \cdot (n_1 + n_2)}{(n_1 + n_2 - 2) \cdot n_1 \cdot n_2}}}. \quad (30)$$

Примечания: 1. При определении вероятности, равной  $2[1 - S(t_\phi)]$ , по приложению IV в качестве  $n$  следует брать  $n_1 + n_2 - 1$ .

2. Если вероятность ( $P$ ) получается большей, то это свидетельствует о том, что следовало ожидать разностей, превышающих ту, которую мы получили фактически. И, следовательно, фактическая разность, будучи меньше тех, которых следовало ожидать с большой вероятностью, не дает основания считать, что различия между средними существенны.

При полученной малой вероятности ( $P$ ) различие между средними не случайно, а существенно.

3. При вычислении  $\Sigma(x - \bar{x})^2$  можно использовать равенство  $\Sigma(x - \bar{x})^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$ .

Пример 15. Третья типовая задача малой выборки.

Оценка разности двух выборочных средних.

Произведена малая выборка 9 участков, аналогично тому, как это сделано в примере 13. Уборка урожая произведена с большим опозданием.

Результат сбора урожая по участкам представлен в табл. 8 (в колонках 1 и 2).

Т а б л и ц а 8

Участки	Урожай пшеницы в ц с га ( $x''$ )	( $x''$ ) <sup>2</sup>
1	10,7	114,49
2	9,0	81,00
3	13,9	193,21
4	9,4	88,36
5	11,9	141,61
6	11,3	127,69
7	10,5	110,25
8	9,9	98,01
9	7,4	54,76
И т о г о . . . . .	94,0	1009,38

Оценить расхождение между средним урожаем, полученным при своевременной уборке урожая (пример 13) и уборке урожая с большим опозданием.

Дано из примера 13:

$$n_1 = 8; \bar{x}' = 16,113; \Sigma x' = 128,9; \Sigma (x')^2 = 2159,25.$$

Из данного примера:

$$n_2 = 9; \Sigma x'' = 94; \bar{x}'' = \frac{94}{9} = 10,444; \Sigma (x'')^2 = 1009,38.$$

Вычисляем:

$$\Sigma (x' - \bar{x}')^2 = \Sigma (x')^2 - \frac{(\Sigma x')^2}{n_1} = 2159,25 - \frac{(128,9)^2}{8} = 82,349.$$

$$\Sigma (x'' - \bar{x}'')^2 = \Sigma (x'')^2 - \frac{(\Sigma x'')^2}{n_2} = 1009,38 - \frac{(94)^2}{9} = 25,38;$$

$$\delta_{\phi} = \bar{x}' - \bar{x}'' = 16,113 - 10,444 = 5,669.$$



По формуле (30) получаем:

$$t_{\phi} = \frac{5,669}{\sqrt{\frac{(82,349 + 25,38) \cdot (8 + 9)}{(8 + 9 - 2) \cdot 8 \cdot 9}}} = \frac{5,669}{\sqrt{\frac{107,729 \cdot 17}{15 \cdot 72}}} = 4,3.$$

Из приложения IV для  $n = 8 + 9 - 1 = 16$ , находим:

$$S(4,3) = 0,999,$$

тогда:

$$P[|\tilde{x}' - \tilde{x}''| > 5,669] = 2[1 - S(4,3)] = 2(1 - 0,999) = 2 \cdot 0,001 = 0,002.$$

Так как вероятность ( $P$ ) очень мала, то следует считать, что средние урожаи существенно отличаются друг от друга, т. е. что опоздание в сроках уборки существенно снижает урожай.

§ 21. При оценке расхождения между двумя выборочными средними часто применяют правило трех сигм:

$$\left. \begin{array}{l} t_{\phi} > 3\sigma_{t(\phi)} \\ t_{\phi} < 3\sigma_{t(\phi)} \end{array} \right\} \quad (31)$$

где  $\sigma_{t(\phi)}$  — среднее квадратическое отклонение, вычисляемое по формуле:

$$\sigma_{t(\phi)} = \sqrt{\frac{n_1 + n_2 - 2}{n_1 + n_2 - 4}}. \quad (32)$$

В первом случае, т. е. если  $t_{\phi}$  больше 3 сигм, расхождение между средними двух выборок полагают не случайным.

Пример 16. По данным примеров 13 и 15 оценить расхождение между двумя выборочными средними по формулам (31) и (32).

Дано:

$$\tilde{x}' = 16,113; n_1 = 8 \quad \tilde{x}'' = 10,444; n_2 = 9; t_{\phi} = 4,3.$$

Находим:

$$\sigma_{t(\phi)} = \sqrt{\frac{8 + 9 - 2}{8 + 9 - 4}} = \sqrt{\frac{15}{13}} = \sqrt{1,1538} \approx 1,07.$$

Получаем:

$$3\sigma_{t(\phi)} \approx 3 \cdot 1,07 = 3,21$$

и, следовательно,

$$4,3 > 3,21, \text{ т. е. } t_{\phi} > 3\sigma_{t(\phi)}.$$

Поэтому расхождение между двумя выборочными средними следует считать существенным, что согласуется с выводом примера 15.

## РАЗДЕЛ IV

### КРИВЫЕ РАСПРЕДЕЛЕНИЯ

**§ 1. Эмпирическое и теоретическое распределение.** Как отмечалось в разделе I, экспериментальные статистические данные для удобства их обработки сводятся в ряды распределения, обязательными реквизитами которых является наличие вариантов данного ряда с указанием частот каждого варианта. Графически такой ряд распределения может быть изображен в виде полигона распределения, или эмпирической кривой распределения.

Математическая статистика, имея ряд распределения некоторого эмпирического материала, наблюдая однотипность этих рядов, стремится выражать распределения аналитически в виде функции. Такая функция называется функцией распределения. Функция распределения выражает теоретические частоты распределения как функцию вариантов.

Графическое изображение функции распределения образует теоретическую кривую распределения, или просто кривую распределения.

**§ 2. Биномиальное распределение.** Как уже указывалось в разделе II, распределение вероятностей событий при повторении испытаний подчинено закону биномиального распределения:

$$P_m = C_n^m p^m q^{n-m} = \frac{n!}{m! (n-m)!} p^m q^{n-m},$$

где  $C_n^m$  представляют собою *биномиальные коэффициенты* и могут быть найдены из треугольника Паскаля.

Исчисление вероятностей по формуле биномиального распределения и составление ряда распределения вероятностей можно иллюстрировать на следующем примере.

**Пример 1.** Производится отбор 10 деталей подряд из большой партии, в которой доля брака составляет 0,1. Найти вероятность появления 0, 1, 2, ..., 10 бракованных деталей.

Ввиду наличия большого числа деталей в партии, мы можем полагать, что вероятность появления бракованных деталей в ходе

отбора не изменяется. Тогда вероятность появления 0 бракованных деталей

$$P_0 = \frac{10!}{0!10!} (0,1)^0 \cdot (0,9)^{10} = 0,34867.$$

Так как в условиях этого примера  $n = 10$ ;  $m = 0$ ;  $p = 0,1$ ;  $q = 0,9$ :

$$P_1 = \frac{10!}{1!9!} \cdot (0,1)^1 \cdot (0,9)^9 = 0,38742;$$

$$P_2 = \frac{10!}{2!8!} \cdot (0,1)^2 \cdot (0,9)^8 = 0,13376;$$

$$P_3 = \frac{10!}{3!7!} \cdot (0,1)^3 \cdot (0,9)^7 = 0,05739;$$

$$P_4 = \frac{10!}{4!6!} \cdot (0,1)^4 \cdot (0,9)^6 = 0,01116;$$

$$P_5 = \frac{10!}{5!5!} \cdot (0,1)^5 \cdot (0,9)^5 = 0,00149;$$

$$P_6 = \frac{10!}{6!4!} \cdot (0,1)^6 \cdot (0,9)^4 = 0,00014;$$

$$P_7 = \frac{10!}{7!3!} \cdot (0,1)^7 \cdot (0,9)^3 = 0,00001;$$

$$P_8 = \frac{10!}{8!2!} \cdot (0,1)^8 \cdot (0,9)^2 = 0,00000;$$

$$P_9 = \frac{10!}{9!1!} \cdot (0,1)^9 \cdot (0,9)^1 = 0,00000;$$

$$P_{10} = \frac{10!}{10!0!} \cdot (0,1)^{10} \cdot (0,9)^0 = 0,00000.$$

Табличная запись вероятностей значений  $m$  составит ряд распределения.

Таблица 1

$m$	0	1	2	3	4
Вероятность . . . . .	0,34867	0,38742	0,13376	0,05739	0,01116

$m$	5	6	7	8	9	10
Вероятность . . .	0,00149	0,00014	0,00001	0,00000	0,00000	0,00000

Графическое изображение этих вероятностей дает кривую распределения.

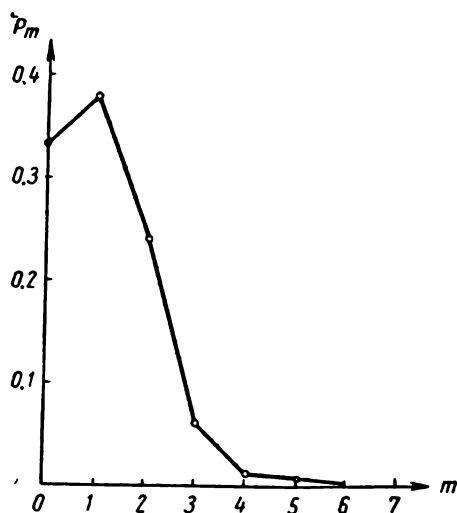


График 10. Распределение вероятностей брака при 10 испытаниях.

§ 3. Нормальное распределение. Функция нормального распределения имеет вид:

$$y_x = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, \quad (1)$$

где  $y_x$  — теоретическая ордината (частота) распределения;  
 $x$  — текущая абсцисса данного эмпирического распределения;  
 $\bar{x}$  — средняя арифметическая этого распределения;  
 $\sigma$  — стандарт, или среднее квадратическое отклонение;  
 $\pi$  и  $e$  — известные в математике величины.

При замене  $\frac{x-\bar{x}}{\sigma} = t$  получаем другой вид этой функции:

$$y_t = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

где величина  $t = \frac{x-\bar{x}}{\sigma}$  носит название стандартизованного отклонения (см. раздел II).

При допущении, что среднее квадратическое отклонение  $\sigma = 1$ , получим функцию стандартизированной кривой нормального распределения:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}. \quad (2)$$

Графическое изображение этой функции дает кривую нормального распределения.

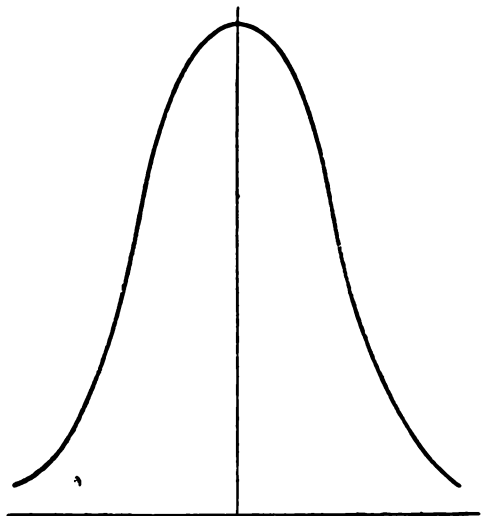


График 11. Кривая нормального распределения.

§ 4. Исследование функции нормального распределения приводит к следующим выводам:

а) при всех значениях  $t$  (как при положительных, так и при отрицательных) функция положительна, т. е. теоретически вычисленные частоты всегда являются величинами положительными, а график этой кривой располагается над осью  $ot$ ;

б) функция симметрична относительно оси  $oy$ , т. е. при любых противоположных по знаку, но равных по абсолютной величине значениях стандартизованного отклонения кривая имеет одинаковые ординаты;

в) при  $t = \pm \infty \lim y_t = 0$ , т. е. кривая нормального распределения асимптотически приближается к оси  $ot$  при больших значениях  $t$ .

При увеличении  $t$  хотя бы до 6 величина функции становится крайне малой:

$$y_{t=6} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} = 0,000000015;$$

г) при  $t=0$   $y_t$  имеет максимум, т. е.

$$y_{\max} = y_0 = \frac{1}{\sqrt{2\pi}};$$

д) при  $t = \pm 1$  кривая имеет точки перегиба;

е) площадь, заключенная между кривой и осью  $ot$ , равна единице, т. е.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1.$$

**§ 5. Построение кривой нормального распределения.** Нормальную кривую распределения по эмпирическим данным можно построить тремя способами:

- 1) применением функции стандартизованного распределения;
- 2) применением стандартизованного отклонения;
- 3) интегральным способом.

**1-й способ** основан на применении функции стандартизованного нормального распределения, в котором, как уже указывалось,  $y_0=1$  — величина наибольшей ординаты принимается за единицу.

За начало отсчета признака при этом способе построения принимается его средняя арифметическая. Ей соответствует наибольшая ордината.

Вычисление ординат производится по формуле:

$$y = \frac{Nb}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \quad (3)$$

где  $N$  — число наблюдений;

$b$  — величина интервала эмпирического распределения.

Так как значение наибольшей ординаты получается при  $t=0$ , когда  $e^{-\frac{t^2}{2}}=1$ , то величина наибольшей ординаты будет

$$y = \frac{Nb}{\sigma\sqrt{2\pi}}. \quad (4)$$

Придавая  $t$  последовательно значения 0,5; 1,0; 1,5; 2,0, т. е. сначала меньшие, а потом увеличивающиеся, находим в таблице стандартизованного нормального распределения для данных  $t$  соответствующие  $y_t$  и, умножив полученную величину на значение наибольшей ординаты, будем иметь соответственно ординаты для этих значений  $t$ .

Например, при  $t=0,5$  величина стандартизованного нормаль-

ного распределения  $y_t = 0,8825$ . Так как величина наибольшей ординаты  $y_0 = \frac{Nb}{\sigma}$ , то величина ординаты в точке  $t = 0,5$  будет:

$$y_t = 0,5 = \frac{Nb}{\sigma\sqrt{2\pi}} \cdot 0,8825 \approx 15.$$

Пример 2. Взяты результаты измерения 100 отклонений шага резьбы ( $x$ ) по всей длине резьбы. Получен следующий ряд распределения, для которого по общим правилам производится расчет средней и дисперсии.

Т а б л и ц а 2

$x$	$m$	$xm$	$x^2m$
-6 -5	3	-16,5	91,0
-5 -4	6	-27,0	122,0
-4 -3	9	-31,5	110,0
-3 -2	14	-35,0	87,5
-2 -1	16	-24,0	36,0
-1 0	16	- 8,0	4,0
0 +1	12	+ 6,0	3,0
+1 +2	10	+15,0	22,5
+2 +3	6	+15,0	37,5
+3 +4	5	+17,5	61,5
+4 +5	3	+13,5	61,0
	100	-75	636,0

Отсюда

$$\bar{x} = -\frac{75}{100} = -0,75$$

$$\sigma^2 = 6,36 - 0,56 = 5,8$$

$$\sigma = \pm 2,4.$$

Производим расчет наибольшей ординаты:

$$y_0 = \frac{Nb}{\sigma\sqrt{2\pi}} = \frac{100 \cdot 1}{2,4\sqrt{2\pi}};$$

так как величина  $\frac{1}{\sqrt{2\pi}} \approx 0,4$ , то имеем:

$$y_0 = \frac{100 \cdot 0,4}{2,4} \approx \frac{400}{24} \approx \frac{100}{6} \approx 17.$$

Беря значение  $t=0,5$  по таблице стандартизованного нормального распределения, находим  $\varphi(t)$  при  $t=0,5$  равно 0,88251. Это и есть коэффициент, который при умножении на значение наибольшей ординаты дает величину ординаты в этой точке. Потом аналогично находим ординаты для  $t=\pm 1$  и т. д.

Для данного примера будем иметь:

$t$	$\varphi(t)$	$y_t$
$\pm 0,5$	0,8825	15
$\pm 1,0$	0,6065	10
$\pm 1,5$	0,3246	5
$\pm 2,0$	0,1353	2
$\pm 2,5$	0,0439	0,6
$\pm 3,0$	0,0111	0,2

Полученный результат наносим на график, а для сравнения наносим на график и результаты непосредственных измерений отклонений (график 12).

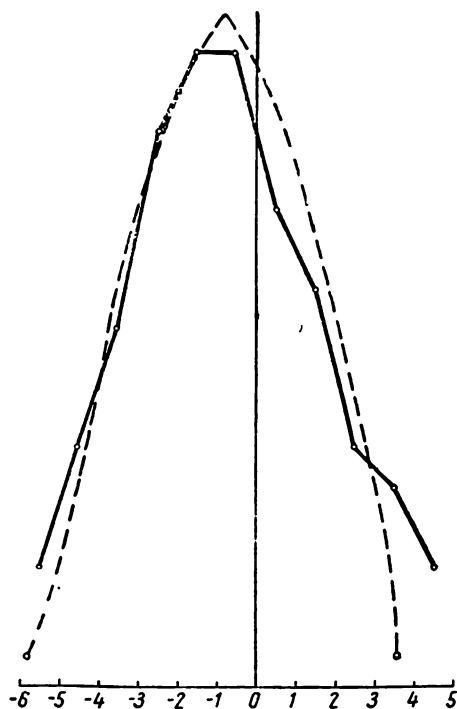


График 12. Распределение 100 отклонений шага резьбы ко всей длине резьбы.

Как видно из графика, теоретическая кривая довольно близко воспроизводит полигон эмпирического распределения.



Пример 3. Даны результаты измерений отклонений шага резьбы ( $x$ ) в  $\mu$  на 1 витке, которые приводятся с соответствующими расчетами:

Т а б л и ц а 3

$x$	$m$	$xm$	$x^2m$	$x$	$m$	$xm$	$x^2m$
-3 -2	8	-20	50	+1 +2	22	33	49
-2 -1	11	-16,5	25	+2 +3	4	10	25
-1 0	25	-12,5	6	+3 +4	3	10,5	37
0 +1	27	+13,5	7				
					100	+18	199

$$\bar{x} = +0,18 \mu; \sigma^2 = 1,96; \sigma = \pm 1,40 \mu.$$

Расчет теоретических частот (ординат) производится, как и в предыдущем случае. Сначала находится величина наибольшей частоты:

$$y_0 = \frac{100 \cdot 0,4}{1,39} \approx 29,$$

затем другие частоты:

$t$	$\varphi(t)$	$y_t$
$\pm 0,5$	0,88	25
$\pm 1,0$	0,60	17
$\pm 1,5$	0,32	9
$\pm 2,0$	0,13	4
$\pm 2,5$	0,04	2
$\pm 3,0$	0,01	0,3

Эмпирические и теоретические частоты наносим на график (график 13) и убеждаемся, что эмпирическое распределение довольно близко воспроизводится теоретически.

2-й способ основан на применении стандартизованного отклонения  $\frac{x-\bar{x}}{\sigma}$ . При этом сначала находят значения  $\bar{x}$  и  $\sigma$  для данного эмпирического распределения. Затем находят отклонения  $x-\bar{x}$  и стандартизованные отклонения  $\frac{x-\bar{x}}{\sigma}$  для данного ва-

рианта. Значения же частоты для нее исчисляются по известной уже формуле:

$$y = \frac{Nb}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

где сохраняются прежние обозначения.

Так как величина  $\frac{Nb}{\sigma}$  остается одной и той же для всего распределения, то достаточно ее найти один раз и умножить

на величину  $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$  при данном  $t$ ; получим искомую теоретическую частоту (ординату).

Значения  $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$  для любых  $t$  даются в приложении II.

**Пример 4.** В следующей таблице дается распределение 500 спиралей сорта 220 V×150 W одной выработки по весу (в мг) и расчет нормальной кривой для этого распределения.

Т а б л и ц а 4

$x$	$m$	$xm$	$x^2$	$x^2m$	$x-\bar{x}$	$t$	$\varphi(t)$	$y_t$
1	2	3	4	5	6	7	8	9
33,0	2	76,0	1 444	2 888	—3	—3	0,004	1
38,5	3	115,5	1 482	4 446	—2,5	—2,5	0,017	4
39,0	10	390,0	1 521	15 210	—2,0	—2,0	0,054	13
39,5	31	1 224,5	1 560	48 330	—1,5	—1,5	0,130	32
40,0	72	2 880,0	1 600	115 200	—1,0	—1,0	0,242	61
40,5	85	3 442,5	1 640	139 400	—0,5	—0,5	0,352	89
41,0	94	3 854,0	1 681	158 014	0	0	0,399	98
41,5	88	3 652,0	1 722	151 536	0,5	0,5	0,352	89
42,0	62	2 604,0	1 764	109 368	1,0	1,0	0,242	61
42,5	37	1 572,5	1 803	66 822	1,5	1,5	0,130	32
43,0	12	516,0	1 849	22 188	2,0	2,0	0,054	13
43,5	3	130,5	1 892	5 676	2,5	2,5	0,017	4
44,0	1	44,0	1 936	1 936	3,0	3,0	0,004	1
	500	20501,5		841 044				

$\bar{x} = 41,0$  мг;  $\bar{x}^2 = 1681$ ;  $\sigma^2 = 1,0$ ;  $\sigma = \pm 1$  мг.

Колонки 1, 2, 3, 4, 5 данной таблицы нужны, как обычно, для вычислений параметров  $x$  и  $\sigma$ , колонки же 6, 7, 8, 9 содержат

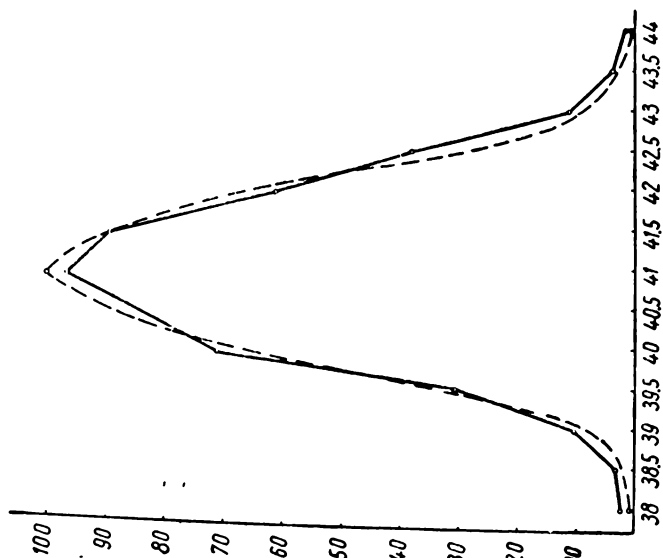


График 13. Распределение отклонений шага резьбы на 1 витке.

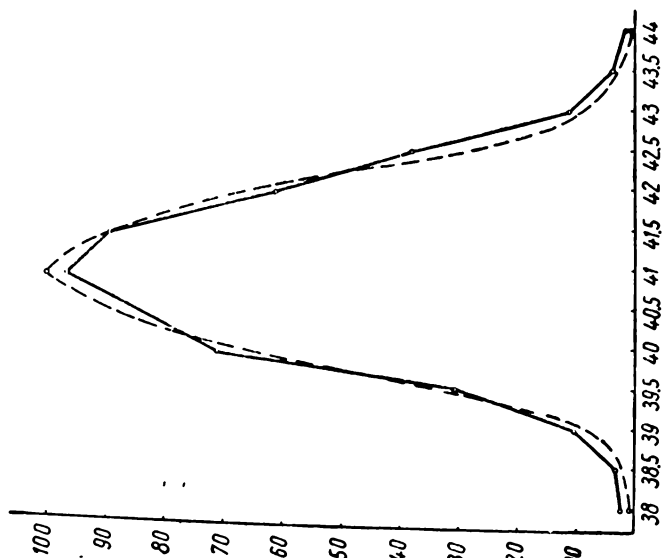


График 14. Распределение веса 500 спиралей.

необходимые расчеты теоретических частот, при этом колонка 6 и колонка 7 нужны для расчета стандартизованного отклонения.

Колонка 8 содержит найденные значения  $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} = \varphi(t)$  взятые из приложения II.

Наносим эмпирические и теоретические частоты на график (см. график 14), из сравнения которых видим достаточную близость теоретически вычисленных данных к эмпирическим.

Пример 5. В следующей таблице дается эмпирическое распределение 110 замеров межцентрового расстояния ( $x$ ) при шевинговании зубцов шестерни динамо-машины 110412.

Единицей измерения являются сотые доли миллиметра. В этой же таблице производится вычисление теоретических частот (ординат кривой нормального распределения). Результаты вычислений наносятся на график, как и результаты полученных эмпирических наблюдений (см. график 15).

Таблица 5

$x$	$m$	$xm$	$x^2$	$x^2m$	$x - \bar{x}$	$t$	$\varphi(t)$	$y_t$	$y_t$
1	2	3	4	5	6	7	8	9	10
3,0	2	6,0	9,00	18,00	-2,6	-2,3	0,03	1,2	1
3,5	7	24,5	12,25	85,75	-2,1	-1,9	0,06	2,6	3
4,0	6	24,0	16,00	96,00	-1,6	-1,2	0,20	8,8	9
4,5	13	58,5	20,25	263,25	-1,1	-0,9	0,27	11,9	12
5,0	16	80,0	25,00	400,00	-0,6	-0,5	0,35	15,4	15
5,5	16	88,0	30,25	484,00	-0,1	-0,01	0,40	17,6	18
6,0	16	96,0	36,00	576,00	+0,1	0,03	0,38	16,7	17
6,5	14	91,0	42,25	591,50	+0,9	0,70	0,31	13,6	14
7,0	11	77,0	49,00	539,00	+1,4	0,2	0,20	8,8	9
7,5	4	30,0	56,25	225,00	+1,9	1,5	0,13	5,7	6
8,0	3	24,0	64,00	192,00	+2,4	2,0	0,15	2,2	2
8,5	1	8,5	72,25	72,25	+2,9	2,4	0,02	0,8	1
9,0	1	9,0	81,00	81,00	+3,4	2,9	0,006	0,2	
	110	616,5		3623,75					

$\bar{x} = 5,6$  мм;  $\bar{x}^2 = 31,36$ ;  $\bar{x}^2 = 32,90$ ;  $s^2 = 32,90 - 31,36 = 1,54$ ;  $\sigma = \pm 1,2$  мм.

**3-й способ** построения кривой нормального распределения (или вычисления теоретических частот) по имеющимся эмпирическим данным основан на применении функции

$$\frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt = F(t), \quad (5)$$

которая дает нам площадь нормальной кривой, заключенной между  $-t$  и  $+t$ .

Вообще говоря, можно находить площадь нормальной кривой, заключенную между любыми точками  $t_0$  и  $t_1$  как

$$\frac{1}{\sqrt{2\pi}} \int_{t_0}^{t_1} e^{-\frac{t^2}{2}} dt,$$

применяя функцию  $F(t)$ . Искомая площадь будет представлять собой  $\frac{1}{2} [F(t_1) - F(t_0)]$ , причем для отрицательных  $t$  надо брать  $F(t)$  со знаком минус.

**Пример 6.** Получены результаты 208 измерений межцентровых расстояний при шевинговании зубцов шестерни динамомашины, по которым вычислим нужные параметры и теоретические частоты и построим графики эмпирического и теоретического распределений.

Здесь, как и раньше, колонки 1, 2, 3, 4 и 5 необходимы для расчетов  $\bar{x}$  и  $\sigma$ ; колонка 6 есть расчет отклонений концов интервалов от средней; колонка 7 — величина стандартизованного отклонения  $\frac{x - \bar{x}}{\sigma} = t_i$  для них же. Колонка 8 содержит значения

$F(t_i)$ , взятые из приложения III, умноженные на  $\frac{1}{2}$  и на 208, т. е.

на 104. В верхней строке приведено и значение  $t$  для конца интервала, предшествующего первому, т. е. для начала этого первого интервала  $\left(\frac{1,75-5,4}{2,01}\right)$ . Чтобы получить теоретическую частоту для

каждого интервала, достаточно из стоящего в его строке в 8-й колонке числа вычесть число, стоящее в предыдущей строке (в той же колонке).

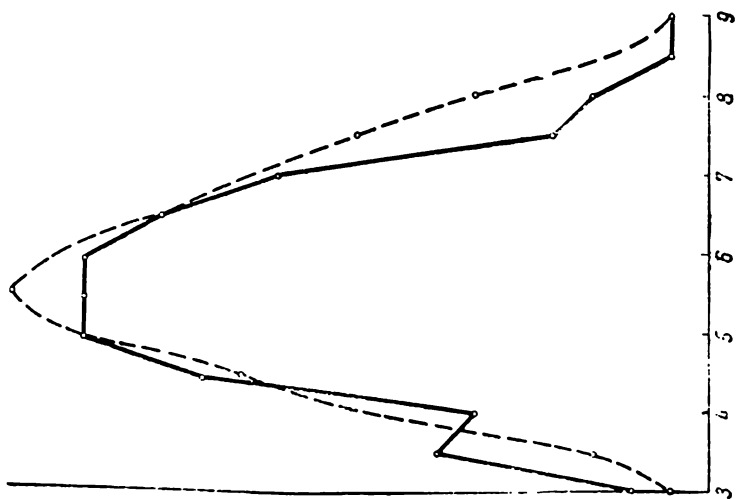


График 15. Распределение 110 замеров межцентрового расстояния.

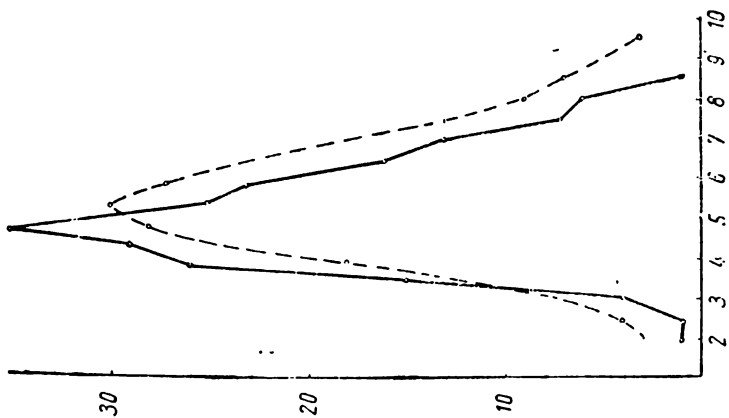


График 16. Распределение межцентровых расстояний при шевинговании зубцов.

Т а б л и ц а 6

$x$	$m$	$xm$	$x^2$	$x^2m$	$x-\bar{x}+0,25$	$t$	$\frac{1}{2} F(t_i) \cdot 210$	$y_t$
2,0	1	2,0	4,00	4,00	-3,65	-2,57	-103,908	
2,5	1	2,5	6,25	6,25	-3,15	-2,11	-101,241	4
3,0	4	12,0	9,00	36,00	-2,65	-1,86	-98,408	4
3,5	15	52,5	12,25	183,75	-2,15	-1,51	-91,245	7
4,0	26	104,0	16,00	416,00	-1,65	-1,16	-79,170	12
4,5	29	130,5	20,25	587,25	-1,15	-1,81	-61,110	18
5,0	35	175,0	25,00	875,00	-0,65	-0,46	-37,212	24
5,5	25	137,5	30,25	756,25	-0,15	-0,15	-10,516	27
6,0	23	138,0	36,00	828,00	0,35	0,24	19,908	30
6,5	16	104,0	42,25	676,00	+0,85	0,60	47,297	27
7,0	13	91,0	49,00	637,00	+1,35	0,95	68,901	22
7,5	7	52,5	56,25	393,75	+1,85	1,30	84,672	16
8,0	6	48,0	64,00	384,00	+2,35	1,58	93,009	8
9,0	7	63,0	81,00	567,00	+2,85	2,00	100,212	7
					+3,85	2,72	104,307	4

208 | 1112,5 | | 6350,25 | | | |

$$\bar{x} = 5,4; \sigma^2 = 2,01; \sigma = \pm 1,42$$

На графике 16 показано, что теоретическое распределение достаточно точно отражает эмпирически полученный материал, только наблюдается некоторое смещение теоретической кривой вправо, что, очевидно, вызвано большим удельным весом правого конца эмпирического распределения.

Пример 7. Дается ряд распределений ударной вязкости в 240 испытаниях. Приведем этот ряд распределения и построим для него теоретическое распределение.

Т а б л и ц а 7

$x$	$m$	$xm$	$x^2$	$x^2m$	$x-\bar{x}$	$t$	$\frac{1}{2} F(t_i) \cdot 240$	$y_t$
3,0—3,4	5	16,0	10,24	51,20	-2,6	-2,28	-117,28	
3,4—3,8	11	39,6	12,96	142,56	-2,2	-1,92	-113,40	4
3,8—4,2	12	48,0	16,00	192,00	-1,8	-1,51	-104,16	9
4,2—4,6	18	79,2	19,36	348,48	-1,4	-1,22	-93,12	11
4,6—5,0	18	79,2	19,36	348,48	-1,0	-0,87	-72,26	21
5,0—5,4	30	144,0	23,04	691,20	-0,6	-0,52	-46,52	26
5,4—5,8	30	156,0	27,04	811,20	-0,2	-0,17	-15,36	31
5,8—6,2	28	156,8	31,36	878,08	+0,2	0,17	15,36	32
6,2—6,6	40	240,0	36,00	1440,00	+0,6	0,52	46,52	31
6,6—7,0	24	153,6	40,96	983,04	+2,0	0,87	72,96	26
7,0—7,4	15	102,0	46,24	693,60	+1,4	1,22	93,12	21
7,4—7,8	10	72,0	51,84	518,40	+1,8	1,51	104,16	11
7,8—8,2	8	60,8	57,76	462,08	+2,2	1,92	113,28	9
8,2—8,6	4	32,0	64,00	256,00	+2,6	2,28	117,12	4
8,6—9,0	2	16,8	70,56	141,12	+3,0	2,60	118,87	1
	3	26,4	77,44	232,32	+3,4	2,96	119,52	0,7
	240	1343,2		7841,28				

$$\bar{x} = 5,6; \sigma^2 = 32,67 - 31,36 = 1,31; \sigma = \pm 1,15$$

§ 6. Критерии «нормальности» распределения. Подчинение многих эмпирических распределений закону нормального распределения различно, или различна степень близости эмпирических распределений к нормальному.

Мерами близости эмпирических распределений к теоретическому распределению вообще, применяемыми часто к нормальному распределению, служат различные «критерии», которые даны Пирсоном, Мизесом, Колмогоровым, Смирновым, Ястремским и др. Рассмотрим некоторые из них.

**Критерий  $\chi^2$  Пирсона.** Для оценки расхождения между теоретическим и эмпирическим распределениями частот К. Пирсон предложил критерий  $\chi^2$ .

Для нахождения  $\chi^2$  вычисляется сумма квадратов разностей эмпирических и теоретических частот, отнесенных к теоретическим частотам, т. е.

$$\chi^2 = \sum \frac{(m - m')^2}{m'}, \quad (6)$$

где  $m$  — эмпирическая частота;

$m'$  — частота, вычисленная теоретически.

Для суждения о близости между эмпирическим и теоретическим распределением надо знать значения  $P(\chi^2)$  — вероятность достижения  $\chi^2$  данного значения.

Значения  $P(\chi^2)$  вычислены для разных  $\chi^2$  и приведены в приложении IX, в котором дается таблица с двумя входами, где одним из аргументов (по строкам) являются значения  $\chi^2$ , а другим (по столбцам) — значения  $k$  — «число степеней свободы» вариации данного ряда распределения. Число степеней свободы вариации в вариационном ряду равно числу групп минус число исчисленных статистических характеристик (средняя, дисперсия, моменты распределения и т. д.), использованных в вычислении теоретического распределения.

Пересечение данного столбца и строки дает искомую вероятность.

При вероятностях  $P(\chi^2)$ , значительно отличающихся от нуля,

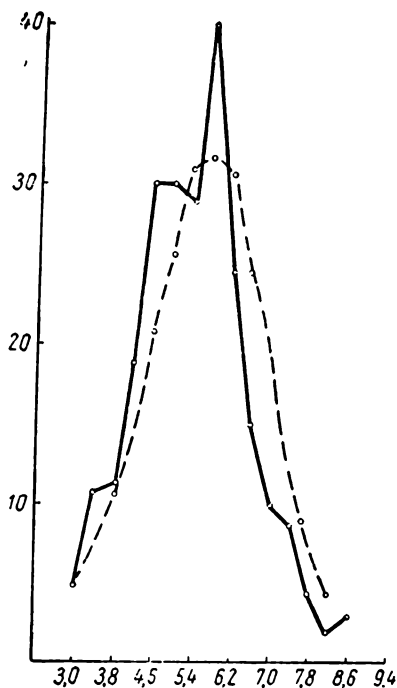


График 17. Распределение ударной вязкости в 240 испытаниях.



расхождение между теоретическими и эмпирическими частотами можно считать случайным.

Для более грубой оценки расхождения между теоретическими и эмпирическими частотами проф. Романовским В. И. предложен более простой метод использования  $\chi^2$ .

Он предлагает вычислять отношение:

$$\frac{\chi^2 - k}{\sqrt{2k}}, \quad (7)$$

где  $k$  — число степеней свободы вариации.

Если указанное отношение дает в результате значение, меньшее трех, то проф. Романовский предлагает расхождение между теоретическими и эмпирическими частотами считать несущественным, если же отношение будет больше трех, то это расхождение существенно. Несущественность расхождения говорит о возможности принимать за закон данного распределения закон нормального распределения.

По примерам вычисления нормальных кривых распределения на эмпирических данных сравним теоретические и эмпирические данные по критерию Пирсона.

Пример 8. Вычисление  $\chi^2$  для распределения межцентрового расстояния в 110 наблюдениях:

Т а б л и ц а 8

$x$	$m$	$m'$	$m - m'$	$(m - m')^2$	$\frac{(m - m')^2}{m'}$
3,0	2	1	1	1	1
3,5	7	3	4	16	5,3
4,0	6	9	-3	9	1
4,5	13	12	1	1	0,08
5,0	16	15	1	1	0,07
5,5	16	18	-2	4	0
6,0	16	17	-1	1	0
6,5	14	14	0	0	0
7,0	11	9	2	4	0,45
7,5	4	6	-2	4	0,66
8,0	3	2	1	1	0,5
8,5	1	1	0	0	0
1,0	1	0,2	0,8	0,64	3,2
					12,31

$$\chi^2 = \sum \frac{(m - m')^2}{m'} = 12,31; k = 12.$$

Входим в таблицу  $P(\chi^2)$  — приложение IX. Для  $\chi^2 = 12$  и  $k = 12$  находим вероятность  $P(\chi) = 0,4457$ ; она достаточно велика,

значит, расхождение между теоретическими и эмпирическими частотами можно считать случайным, а распределение — подчиняющимся закону нормального распределения.

Находим отношение Романовского:

$$\frac{|\chi^2 - k|}{\sqrt{2k}} \approx \frac{0,31}{5} \approx 0,06.$$

Это соотношение значительно меньше трех, поэтому расхождение между теоретическими и эмпирическими частотами можно считать несущественным, и, таким образом, теоретическое распределение достаточно воспроизводит эмпирическое.

Пример 9. Вычисление критерия  $\chi^2$  для распределения веса 500 спиралей:

Т а б л и ц а 9

$x$	$m$	$m'$	$m - m'$	$(m - m')^2$	$\frac{(m - m')^2}{m'}$
38,0	2 } 5	5	0	0	0
38,5					
39,0		13	3	19	0,70
39,5		32	1	1	0,03
40,0	10	61	11	121	2,00
40,5	85	89	4	16	0,05
41,0	94	98	4	16	0,16
41,5	88	89	1	1	0,01
42,0	62	61	1	1	0,01
42,5	37	32	5	25	0,78
43,0	12	13	6	11	0,10
43,5	3 } 4	5	1	1	0,20
44,0					

$$\chi^2 = 4,02; k = 12.$$

$$4,02$$

По таблице находим вероятность  $P(\chi^2) = 0,9834$ , которая близка к достоверности и поэтому расхождение между теоретическим и эмпирическим распределением может быть случайным.

Отношение Романовского

$$\frac{|\chi^2 - k|}{\sqrt{2k}} \approx \frac{7,98}{5} \approx 1,59$$

также значительно меньше трех, поэтому теоретическое воспроизведение эмпирического ряда достаточно удовлетворительное.

**Критерий  $\lambda$  Колмогорова.** Критерий  $\lambda$ , предложенный акад. Колмогоровым, устанавливает близость теоретических и эмпирических распределений, сравнивая их интегральные распределения.  $\lambda$  исчисляется исходя из  $D$  — максимального верхнего предела

абсолютного значения разности их «накопленных частот», отнесенного к квадратному корню из числа наблюдений  $N$

$$\lambda = \frac{D}{\sqrt{N}}, \quad (8)$$

где  $D$  — максимальная граница разности  $M'$  — накопленных теоретических частот и  $M$  — накопленных эмпирических частот.

В приложении VIII приводятся вероятности  $P(\lambda)$  того, что  $\lambda$  достигнет данной величины.

Если найденному значению  $\lambda$  соответствует очень малая вероятность  $P(\lambda)$ , то расхождение между эмпирическим и теоретическим распределением нельзя считать случайным и, таким образом, первое мало отражает второе; наоборот, если  $P(\lambda)$  — величина значительная (больше 0,05), то расхождение между частотами может быть случайным и распределения хорошо отвечают одно другому.

Проследим применение этого критерия на двух примерах.

Пример 10.

Т а б л и ц а 10

$x$	$m$	$m'$	$M$	$M'$	$M - M'$
3,0	2	1	2	1	1
3,5	7	3	9	4	5
4,0	6	9	15	13	2
4,5	13	12	28	25	3
5,0	16	15	44	40	4
5,5	16	18	60	58	2
6,0	16	17	76	75	1
6,5	14	14	90	89	1
7,0	11	9	101	98	3
7,5	4	6	105	104	1
8,0	3	2	108	106	2
8,5	1	1	109	107	2
9,0	1	0,2	110	107	3

$$\lambda = \frac{5}{\sqrt{110}} \approx \frac{5}{10,5} \approx 0,5.$$

В таблице  $P(\lambda)$  находим для  $\lambda = 0,5$ ,  $P(\lambda) = 0,9636$ .

Эта большая вероятность указывает на то, что расхождение между наблюдаемым и теоретическим распределением вполне могло быть случайным.

Пример 11.

Таблица 11

$x$	$m$	$m'$	$M$	$M'$	$M-M'$
38,0	2	1	2	1	1
38,5	3	4	5	5	0
39,0	10	13	15	18	3
39,5	31	32	46	50	4
40,0	72	61	118	111	3
40,5	85	89	203	200	3
41,0	94	98	297	298	1
41,5	88	89	385	387	2
42,0	62	61	447	448	1
42,5	37	32	484	480	4
43,0	12	13	496	493	2
43,5	3	4	499	497	2
44,0	1	1	500	498	2

$$\lambda = \frac{4,0}{500} \cong \frac{4,0}{23,31} \cong 0,17; \quad P(\lambda) = 0,7112.$$

Значительная вероятность  $P(\lambda)$  говорит и здесь о несущественности расхождений между теоретическим и эмпирическим распределением.

**Критерий Ястремского.** В общем виде критерий Ястремского можно записать следующим неравенством:

$$I \leq 3\sqrt{2n+4\theta}, \quad (9)$$

где  $I = |C - n|$ ;

$$C = \sum \frac{(m - m')^2}{m'q},$$

$m$  — эмпирические частоты;

$m'$  — теоретические частоты;

$n$  — число групп.

Для числа групп меньших 20  $\theta = 0,6$ ;  $q = 1 - p$ .

Значение  $I$  меньше  $3\sqrt{2n+4\theta}$  в критерии Ястремского показывает несущественность расхождений между эмпирическими и теоретическими частотами в данном распределении. При значениях  $I$ , больших  $3\sqrt{2n+4\theta}$ , расхождение между теоретическим и эмпирическим распределением существенно.

**Пример 12.** Определим величину  $I$  и оценим эмпирическое распределение 500 спиралей ( $m$ ) по сравнению с соответствующим нормальным ( $m'$ )

Т а б л и ц а 12

$x$	$m$	$m'$	$p$	$q$	$m-m'$	$m'q$	$\frac{(m-m')^2}{m'q}$
38,0	2	1	0,002	0,998	1	0,998	1,00
38,5	3	4	0,005	0,992	1	3,968	0,25
39,0	10	13	0,016	0,974	9	12,662	0,71
39,5	31	32	0,064	0,936	1	22,975	0,04
40,0	72	61	0,122	0,878	121	53,558	2,24
40,5	85	89	0,175	0,822	16	73,158	0,22
41,0	94	98	0,196	0,804	16	78,992	0,20
41,5	88	89	0,178	0,822	1	73,158	0,01
42,0	62	61	0,122	0,878	1	53,558	0,02
42,5	37	32	0,064	0,936	25	22,975	1,09
43,0	12	13	0,026	0,974	1	12,662	0,08
43,5	3	4	0,008	0,992	1	3,968	0,25
44,0	1	1	0,002	0,992	0	0,992	0
	500	500					6,25

$$I < 3 \sqrt{2n + 4\theta};$$

$$C = 6,25;$$

$$|13 - 6,25| < 3 \sqrt{26 + 2,4};$$

$$|13 - 6,25| < 3 \sqrt{28,4};$$

$$6,75 < 3 \cdot 5,34;$$

$$6,75 < 16,02,$$

что говорит о нормальном распределении исследуемой совокупности.

§ 7. Элементарные приемы определения «нормальности» распределения. Для определения элементарными способами близости данного опытного распределения к нормальному прибегают к числам Вестергарда и к сравнению средней арифметической, моды и медианы.

Числами Вестергарда являются: 0,3; 0,7; 1,1; 3. Для пользования ими определяют сначала основные характеристики — среднюю арифметическую ( $\bar{x}$ ) и среднее квадратическое отклонение ( $\sigma$ ).

Для того, чтобы сказать, что данное эмпирическое распределение подчинено закону нормального распределения, необходимо, чтобы распределение удовлетворяло следующим условиям:

1) в промежутке от  $\bar{x} - 0,3\sigma$  до  $\bar{x} + 0,3\sigma$  было расположено  $\frac{1}{4}$  всей совокупности;

2) в промежутке от  $\bar{x} - 0,7\sigma$  до  $\bar{x} + 0,7\sigma$  было расположено  $\frac{1}{2}$  всей совокупности;

3) в промежутке от  $\bar{x} - 1,1\sigma$  до  $\bar{x} + 1,1\sigma$  было расположено  $\frac{3}{4}$  всей совокупности;

4) в промежутках от  $\bar{x} - 3\sigma$  до  $\bar{x} + 3\sigma$  было расположено 0,998 всей совокупности.

Для приводимого распределения 500 спиралей по весу (пример 8) все эти условия соблюдаются, что говорит о подчинении данного распределения закону нормального распределения.

К элементарным приемам определения «нормальности» следует отнести применение графического метода, особенно удобное с помощью сетки Турбина. На сетке накопленные частоты распределения при нормальном их распределении дают прямую линию. Всякое отклонение от прямой свидетельствует об отклонении эмпирического распределения от «нормального».

§ 8. Распределение Пуассона. Вероятности частот событий, редко встречающихся при некотором числе испытаний, находят по формуле:

$$P_m = \frac{(np)^m e^{-np}}{m!} . \quad (10)$$

Это выражение носит название закона распределения Пуассона.

Здесь  $m$  — частота данного события;

$n$  — число испытаний;

$p$  — вероятность события при одном испытании;  $e = 2,71828$ .

Подставив вместо  $np$  среднее число, наблюдавшееся фактически в эмпирическом материале, — теоретические ординаты кривой распределения по закону Пуассона, найдем формулу:

$$m'_x = n \cdot \frac{\bar{x}^x \cdot e^{-\bar{x}}}{x!} , \quad (11)$$

где  $x$  — переменное значение (числа раз);

$\bar{x}$  — среднее число раз в эмпирическом распределении;

$n$  — число наблюдений.

При

$$x = 0, m' = n e^{-\bar{x}} ;$$

при

$$x = 1 \quad m' = n \bar{x} e^{-\bar{x}} ;$$

при

$$x = 2 \quad m' = \frac{n \bar{x}^2 e^{-\bar{x}}}{2} ;$$

при

$$x = 3 \quad m' = \frac{n \bar{x}^3 e^{-\bar{x}}}{6}$$

и т. д.

Пример 13. В. И. Романовский<sup>1</sup> наблюдал распределение семян повилики в 1000 выборках семян клевера. Результаты эксперимента записаны в следующей таблице:

<sup>1</sup> Романовский В. И., Применение математической статистики в опытном деле, М., 1947.

Т а б л и ц а 13

Число семян по- вилики в одной выборке $x$	Число выборок $m$
0	599
1	315
2	74
3	12

Определим по закону Пуассона теоретические частоты разного числа семян повилики. Для этого исчислим предварительно среднее число семян повилики в одной выборке.

$$\bar{x} = \frac{\sum xm}{\sum m} = \frac{499}{1000} = 0,5.$$

Из таблицы находим  $e^{-0,5} = 0,606$ .

Найдем теоретическое число выборок, в которых число семян повилики будет равно 0:

$$m' = ne^{-\bar{x}} = 1000 \cdot 0,606 = 606;$$

то же, для числа семян повилики, равного 1:

$$m' = n\bar{x}e^{-\bar{x}} = 1000 \cdot 0,50 \cdot 0,606 = 303;$$

для числа семян повилики, равного 2:

$$m' = \frac{n\bar{x}^2 e^{-\bar{x}}}{2} = \frac{1000 \cdot 0,25 \cdot 0,606}{2} = 76;$$

для числа семян повилики, равного 3:

$$m' = \frac{n\bar{x}^3 e^{-\bar{x}}}{6} = \frac{1000 \cdot 0,125 \cdot 0,606}{6} = 13;$$

для числа семян повилики более 3:

$$m' = 2.$$

Графическое сопоставление обоих распределений говорит о соответствии между эмпирическим и теоретическим распределениями.

§ 9. Распределение Максвелла. В технике часто встречается распределение по закону Максвелла. Это — распределение существенно положительных величин. Например, в технике исследование

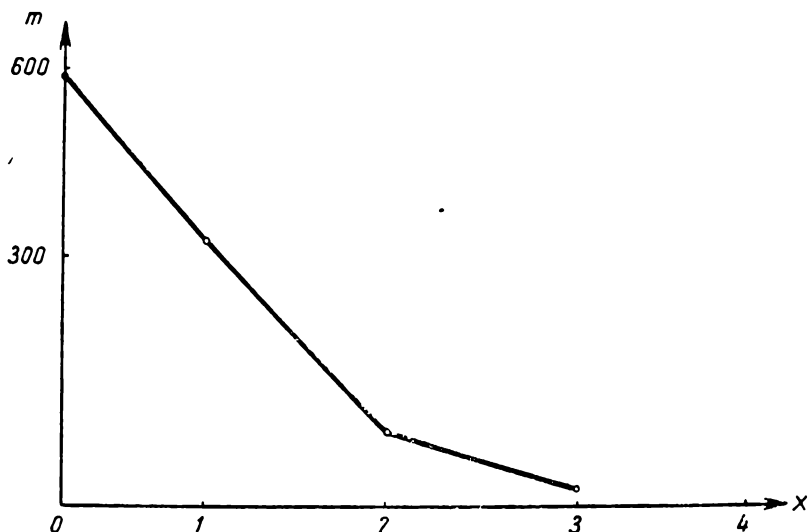


График 18. Распределение семян повилки в клевере.

эксцентриситетов биений приводит к распределению Максвелла.

Дифференциальный закон распределения Максвелла выражается следующей формулой:

$$\psi(t) = \frac{2}{\sqrt{2\pi}} \cdot \frac{t^2}{\alpha} \cdot e^{-\frac{t^2}{2}}, \quad (12)$$

где  $\alpha$  — параметр распределения, равный  $0,6267 \bar{x}$ .

$$\sigma^2 = 3\alpha^2 - \frac{8}{\pi} \alpha^2 = \left(3 - \frac{8}{\pi}\right) \alpha^2 = 0,454\alpha^2;$$

$$\sigma = 0,647\alpha; \quad t = \frac{x}{\alpha}; \quad \frac{dx}{\alpha} = dt.$$

Интегральный закон распределения выразится тогда:

$$\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t t^2 e^{-\frac{t^2}{2}} dt = F(t) - 2t \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

$$\Phi(t) = F(t) - 2t\varphi(t) \quad (13)$$

Пример 14. Возьмем из книги Длин<sup>1</sup> таблицу распределения симметричности гнезд относительно торцов в круглых плашках (в 0,01 мм) и проведем дополнительные расчеты.

<sup>1</sup> Длин А. М., Математическая статистика в технике, М., 1949.



Т а б л и ц а 14

Симметричность гнезд относительно торцов	Число наблюдений	Средины интервалов	Гр. 2 × гр. 3
1	2	3	4
3,5—8,5	9	6	54
8,5—13,5	24	11	264
13,5—18,5	38	16	608
18,5—23,5	33	21	693
23,5—28,5	43	26	1118
28,5—33,5	27	31	837
33,5—38,5	11	36	396
38,5—43,5	6	41	246
43,5—48,5	7	46	322
48,5—53,5	2	51	102
	200		4640

Из этой таблицы легко определим среднюю симметричность:

$$\bar{x} = \frac{4640}{200} = 23,2;$$

параметр рассеяния:

$$\alpha = 0,627\bar{x};$$

$$\alpha = 0,627 \cdot 23,2 = 14,55.$$

Приведенная формула интегрального распределения по закону Максвелла позволяет найти накопленные, а затем и теоретические частоты и частоты по схеме:

Т а б л и ц а 15

$x$	$m$	$t = \frac{x}{\alpha}$	$\varphi(t)$	$2t\varphi(t)$	$F(t)$	Теорети- ческие накоплен- ные частоты	Теорети- ческие частоты	Теорети- ческие частоты
1	2	3	4	5	6	7	8	9
1,5	9	0,584	0,3364	0,3929	0,4408	0,048	0,048	10
13,5	24	0,927	0,2596	0,4813	0,6460	0,165	0,117	23
18,5	38	1,271	0,1779	0,4522	0,7962	0,344	0,179	37
23,5	33	1,616	0,1081	0,3496	0,8939	0,544	0,200	40
28,5	43	1,96	0,0584	0,2289	0,9500	0,721	0,177	35
33,5	27	2,30	0,0283	0,1302	0,9786	0,849	0,128	26
38,5	11	2,64	0,0122	0,0644	0,9917	0,927	0,078	16
43,5	6	2,99	0,0046	0,0274	0,9972	0,970	0,043	9
48,5	7	3,33	0,0016	0,0107	0,9990	0,998	0,018	4
53,5	2	3,68	0,0005	0,0037	1,0000	0,996	0,008	2

Изобразим на графике 19 данные эмпирического и теоретического рядов распределения. Определим близость их по критерию

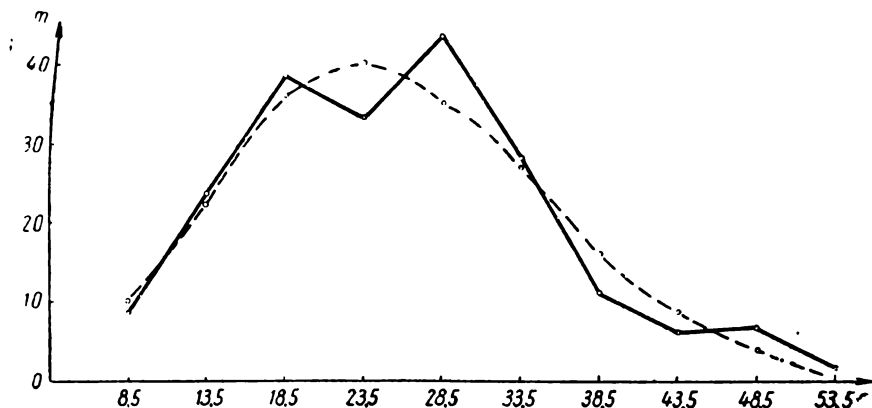


График 19. Распределение симметричности гнезд относительно торцов.

согласия Ястремского, для чего приведем расчет величины:

$$C = \Sigma \frac{(m - m')^2}{m'}$$

Т а б л и ц а 16

$x$	$m$	$m'$	$m - m'$	$(m - m')^2$	$\frac{(m - m')^2}{m'}$
8,5	9	10	-1	1	0,1
13,5	24	23	+1	1	0,05
18,5	38	37	+1	1	0,03
23,5	33	40	-7	49	1,22
28,5	43	35	+8	64	1,83
33,5	27	26	+1	1	0,33
38,5	11	16	-5	25	1,50
43,5	6	9	-3	9	1,00
48,5	7	4	+3	9	2,22
53,5	2	2	0	0	0
					7,98

По критерию  $I = \frac{|C - n|}{\sqrt{2\pi + 4n}}$ ;  $C = 7,98$ ;

$$I = \frac{|7,98 - 10|}{\sqrt{20 + 2,4}} = \frac{2,02}{\sqrt{22,4}} = \frac{2,02}{4,74} \approx 0,42,$$

что явно меньше 3. Следовательно, данное эмпирическое распределение хорошо согласуется с законом распределения Максвелла.

## РАЗДЕЛ V

### КОРРЕЛЯЦИЯ

§ 1. Понятие о корреляционной связи. При изучении зависимости между величинами  $x$  и  $y$  различают зависимости функциональные и статистические. При функциональных зависимостях каждому значению  $x$  соответствует одно совершенно определенное значение  $y$ .

При статистических зависимостях каждому значению  $x$  может соответствовать ряд значений  $y$ . Статистическая связь может быть выражена при помощи перечневой и корреляционной таблицы. Перечневая таблица получится в том случае, если мы выпишем значения  $x$  у всех единиц совокупности и рядом значения  $y$  — для тех же единиц.

Пример 1. Даны 10 равных участков земли с разным количеством удобрений, внесенных на 1 га, и разной урожайностью на этих участках. Обозначим урожайность через  $y$  и количество внесенных удобрений через  $x$ . В результате статистического наблюдения получили следующие результаты:

Т а б л и ц а 1	
Количество внесенных удобрений ( $m$ на 1 га) $x$	Урожайность $y$
1	14
2	14
5	18
3	15
2	15
3	16
3	16
4	16
5	18
4	17

Корреляционная таблица дает возможность устранить повторения значений и представить статистическую связь более ком-

пактно. Для предыдущего примера корреляционная таблица будет иметь следующий вид:

Т а б л и ц а 2

$y$ (урожайность) $x$ (удобрение)	14	15	16	17	18	Всего
1	1	—	—	—	—	1
2	1	1	—	—	—	2
3	—	1	2	—	—	3
4	—	—	1	1	—	2
5	—	—	—	—	2	2
Всего . . .	2	2	3	1	2	10

Цифра, стоящая на пересечении граф и строк, показывает количество участков с данной урожайностью и данным количеством удобрений.

Одним из выражений статистических связей являются связи корреляционные, т. е. зависимость между  $x$  и средним значением  $y$ . Приведем пример корреляционной связи между урожайностью и количеством внесенных удобрений.

Т а б л и ц а 3

$x$ (удобрение)	$y$ (урожайность)
1	14
2	14,5
3	15,7
4	16,5
5	18

Статистическая связь может быть прямой, когда с увеличением  $x$  происходит увеличение  $y$ , и обратной, когда с увеличением  $x$  происходит уменьшение  $y$ .

Корреляционная зависимость становится наглядной при изображении ее в виде многоугольника корреляции. Многоугольник корреляции получается при нанесении данных значений  $x$  и средних  $y$  на поле координат (см. график 20).

§ 2. **Выравнивание.** Если с увеличением  $x$  одновременно происходит более или менее равномерное нарастание (или падение)  $y$ , то говорят, что имеет место линейная зависимость между  $x$  и  $y$ . Математическая статистика при этом находит уравнение прямой линии, отображающей корреляционную зависимость между  $y$  и  $x$ . Нахождение аналитического выражения зависимости между признаками носит название выравнивания.

Линейная корреляционная зависимость выражается уравнением прямой:

$$\bar{y}_x = a_0 + a_1 x, \quad (1)$$

где  $\bar{y}_x$  — среднее значение  $y$  при данном значении  $x$ ;  
 $a_0$  и  $a_1$  — параметры уравнения.

Зная параметры и величину  $x$ , легко отыскать  $\bar{y}_x$ . Отыскание параметров  $a_0$  и  $a_1$  производится с помощью способа наименьших квадратов, который ставит условием, чтобы сумма квадратов

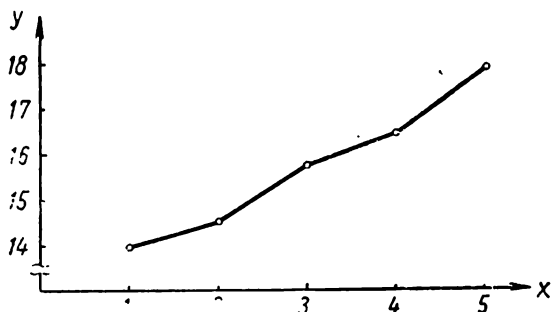


График 20. Зависимость урожайности (y) от количества внесенных удобрений.

отклонений  $\bar{y}_x$  от  $y$  была минимальной. Применение способа наименьших квадратов приводит к системе нормальных уравнений:

$$\left. \begin{aligned} na_0 + a_1 \Sigma x &= \Sigma y \\ a_0 \Sigma x + a_1 \Sigma x^2 &= \Sigma xy \end{aligned} \right\}, \quad (2)$$

решение которых относительно  $a_0$  и  $a_1$  определит их величину.  
 $n$  — число пар значений.

Чтобы составить систему уравнений для примера 1, строим таблицу, позволяющую найти  $\Sigma x$ ,  $\Sigma y$ ,  $\Sigma x^2$ ,  $\Sigma xy$ :

Т а б л и ц а 4

$x$	$y$	$x^2$	$xy$	$\bar{y}_x$
1	14,0	1	14	13,750
2	14,5	4	29	14,745
3	15,7	9	47,1	15,740
4	16,5	16	66,0	16,735
5	18,0	25	90,0	17,730
15	78,7	55	246,1	78,700

Из таблицы получим систему уравнений:

$$\begin{cases} 5a_0 + 15a_1 = 78,7; \\ 15a_0 + 55a_1 = 246,1, \end{cases}$$

откуда

$$a_1 = 0,995; \quad a_0 = 12,755.$$

Следовательно, уравнение корреляционной связи примет вид:

$$\bar{y}_x = 12,755 + 0,995x.$$

Наглядное представление о выравнивании дает график 21. Исчисленные в таблице  $\bar{y}_x$  являются значениями урожайности,

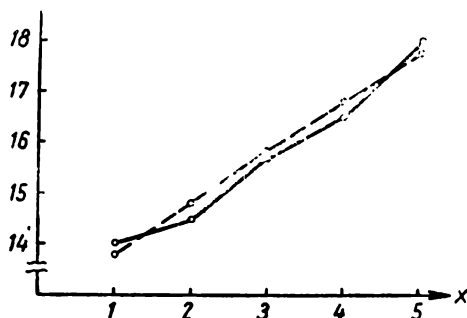


График 21. Зависимость урожайности ( $y$ ) от количества удобрений ( $x$ ) и выравненные значения  $y_x$ .

найденными по уравнению. Итоги по данным значениям ( $y$ ) и вычисленным значениям ( $\bar{y}_x$ ) одинаковы (с округлением).

§ 3. При изменении по параболе 2-го порядка корреляционное уравнение имеет вид:

$$\bar{y}_x = a_0 + a_1x + a_2x^2. \quad (3)$$

Параметры  $a_0$ ,  $a_1$  и  $a_2$  находят, решая систему нормальных уравнений (3 уравнения с тремя неизвестными):

$$\left. \begin{aligned} na_0 + a_1\sum x + a_2\sum x^2 &= \sum y \\ a_0\sum x + a_1\sum x^2 + a_2\sum x^3 &= \sum yx \\ a_0\sum x^2 + a_1\sum x^3 + a_2\sum x^4 &= \sum yx^2 \end{aligned} \right\} \quad (4)$$

**Пример 2.** Корреляционная связь между уровнем производительности труда торфяника и коэффициентом использования агрегата<sup>1</sup> выражается параболой второго порядка.

Т а б л и ц а 5

Месяцы	Коэффициент использования агрегата	Уровень произв. труда (в т на 1 чел.-день)	$x - \bar{x}$	$y - \bar{y}$	$x^2$	$x^1$	$x^1$	$xy$	$x^2y$	$\bar{y}_x$
Апрель . . .	60,6	16,2	-8,2	-6,2	67,24	-551,368	4521,2176	50,84	416,888	16,4
Май . . .	66,1	20,3	-2,7	-2,1	7,29	-19,683	25,1441	5,67	-15,309	19,4
Июнь . . .	74,2	28,3	5,4	5,9	29,16	157,464	850,3056	31,86	172,055	27,5
Июль . . .	72,9	25,1	4,1	2,7	16,81	68,921	282,5761	11,07	45,387	25,8
Август . . .	70,2	22,1	1,4	0,3	1,96	2,744	3,8416	-0,42	-0,588	22,9
	344,0	112,0	0	0	122,46	-341,922	5771,9850	99,02	-215,354	112,0

Параметры  $a_0$ ,  $a_1$ ,  $a_2$  находим, решая систему уравнений:

$$5a_0 + 0a_1 + 122,46a_2 = 0;$$

$$0a_0 + 122,46a_1 - 341,922a_2 = 99,02;$$

$$122,46a_0 + 341,922a_1 + 5771,085a_2 = 215,354,$$

откуда получаем:

$$a_0 = 0,852;$$

$$a_1 = 0,906;$$

$$a_2 = 0,0348.$$

В этой системе уравнений произведена подстановка  $x - \bar{x}$  вместо  $x$  и  $y - \bar{y}$  вместо  $y$ . Само корреляционное уравнение после соответствующих подстановок примет вид:

$$\bar{y}_x = 123,942 - 3,882x + 0,0348x^2.$$

График 22 дает наглядное представление о достигнутом результате.

§ 4. При обратной пропорциональности применяется уравнение гиперболы:

$$\bar{y}_x = a_0 + \frac{a_1}{x}. \quad (5)$$

Параметры  $a_0$  и  $a_1$  находятся решением системы уравнений:

$$a_0 + a_1 \Sigma \frac{1}{x} = \Sigma y;$$

$$a_0 \Sigma \frac{1}{x} + a_1 \Sigma \frac{1}{x^2} = \Sigma \frac{y}{x}. \quad (6)$$

<sup>1</sup> См. Магарилл, Учебное пособие по измерению связей, Иркутск, 1944.

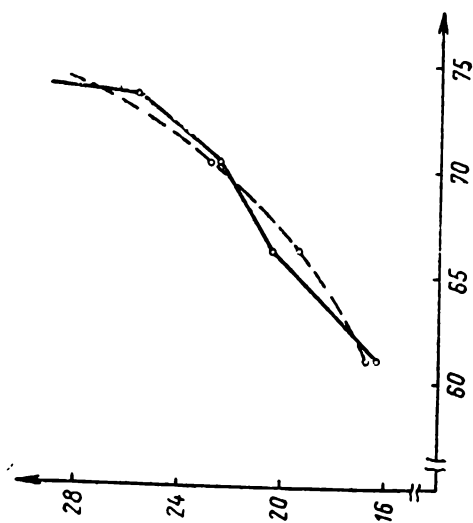


График 22 Зависимость уровня производительности труда от коэффициента использования агрегата.

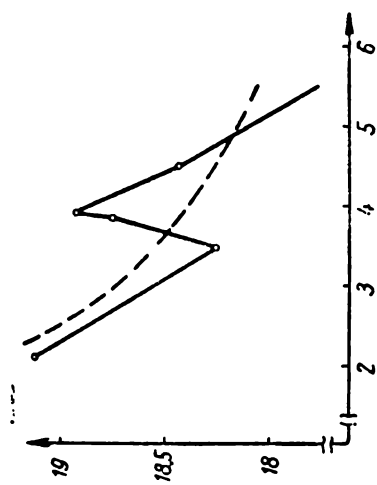


График 23. Зависимость себестоимости продукции от выпуска продукции (тыс. шт.).



**Пример 3.** Зависимость между себестоимостью единицы продукции и выпуском продукции<sup>1</sup> на тех же производственных площадях является обратной.

Т а б л и ц а 6

Выпуск тыс. шт. $x$	Себестои- мость 1 шт. в руб. $y$	$\frac{1}{x}$	$\frac{1}{x^2}$	$\frac{y}{x}$	$\bar{y}_x$
5,669	17,72	0,17640	0,0311170	3,12581	18,14
4,453	18,45	0,22457	0,0504317	4,14322	18,33
3,843	18,94	0,26021	0,0677092	4,92838	18,48
3,745	18,75	0,26702	0,0712997	5,00663	18,50
3,726	18,26	0,26838	0,0720278	4,90062	18,51
2,177	19,12	0,45935	0,2110024	8,78277	19,27
	111,24	1,65593	0,5035878	30,88753	

Решая систему уравнений,

$$6a_0 + 1,65593a_1 = 111,24$$

$$1,65593a_0 + 0,5035878a_1 = 30,88753,$$

находим  $a_0 = 17,43$  и  $a_1 = 4,02$ .

Следовательно, корреляционное уравнение:

$$\bar{y}_x = 17,43 + \frac{4,02}{x}$$

(см. график 23).

**§ 5. Выравнивание динамических рядов.** При выравнивании динамических рядов в качестве аналитической функции для выравнивания могут применяться разные функции. Выбор зависит от характера движения динамического ряда.

Приведем пример применения показателей функции вида:

$$y_x = a_0 a_1^x, \quad (7)$$

где параметры  $a_0$  и  $a_1$  находятся решением системы нормальных уравнений.

$$\left. \begin{aligned} n \lg a_0 + \lg a_1 \Sigma x &= \Sigma \lg y \\ \lg a_0 \Sigma x + \lg a_1 \Sigma x^2 &= \Sigma x \lg y \end{aligned} \right\} \quad (8)$$

За  $x$  берется фактор времени, который приводится к такому виду, чтобы  $\Sigma x = 0$ .

**Пример 4.** Имея данные об индексах объема продукции минералов и гидроэнергии на 1 жителя США за период 1915—1929 гг., можно выравнивать этот ряд по показательной кривой.

<sup>1</sup> См. Р о д и н, Практика математических приемов в статистике, Л., 1935.

Таблица 7

Годы	$x$	$x^2$	Индекс объема продукции* $y$	$\lg y$	$x \lg y$	$\bar{y}_x$
1915	-7	49	67,14	1,82698	-12,78886	69,23
1916	-6	36	79,17	1,89856	-11,39136	71,17
1917	-5	25	82,75	1,91777	-9,58885	73,16
1918	-4	16	81,88	1,91318	-7,65272	75,18
1919	-3	9	71,85	1,85643	-5,56929	77,26
1920	-2	4	80,44	1,90547	-3,81094	79,45
1921	-1	1	62,10	1,79309	-1,79309	81,70
1922	0	0	70,80	1,85603	0	83,98
1923	1	1	93,45	1,97058	1,97058	86,33
1924	2	4	87,45	1,94176	3,88352	88,77
1925	3	9	91,42	1,96104	5,88312	91,29
1926	4	16	98,52	1,99352	7,97408	93,81
1927	5	25	99,06	1,99590	9,979 0	96,41
1928	6	36	100,68	2,00294	12,01764	99,10
1929	7	49	108,56	2,03567	14,24969	101,87
	0	280		28,86292	3,36302	

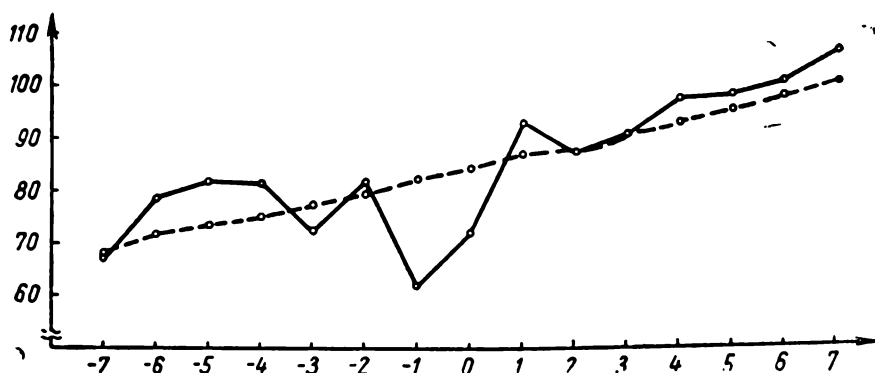


График 24. Индексы объема продукции США (1915–1929 гг.).

Так как  $\lg a_1 \Sigma x = 0$  и  $\lg a_0 \Sigma x = 0$ , система нормальных уравнений упростится:

$$15 \lg a_0 = \Sigma \lg y, \text{ или } 15 \lg a_0 = 28,86292;$$

$$\lg a_1 \Sigma x^2 = \Sigma x \lg y, \text{ или } 280 \lg a_1 = 3,36302;$$

$$a_0 = 83,98, a_1 = 1,028.$$

Уравнение будет таким:

$$\bar{y}_x = 83,98 (1,028)^x$$

\* Крейнин Г. С., Статистические методы и практика их применения в капиталистическом хозяйстве, М., 1938.

§ 6. Корреляционное отношение. При изучении статистических связей большую роль играет разложение дисперсий согласно правилу сложения вариации (см. раздел I). По этому правилу общая дисперсия может рассматриваться как сумма межгрупповой дисперсии и внутригрупповой дисперсии. Межгрупповая дисперсия представит собой вариацию, возникающую за счет признака группировки, а внутригрупповая — вариацию за счет других факторов:

межгрупповая дисперсия

$$\bar{\delta}^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 m_i}{\sum m_i};$$

внутригрупповая дисперсия

$$\bar{\sigma}^2 = \frac{\sum \sigma_i^2 m_i}{\sum m_i},$$

следовательно, по правилу сложения вариаций:

$$\sigma^2 = \bar{\delta}^2 + \bar{\sigma}^2.$$

Показателем тесноты связи служит корреляционное отношение, квадрат которого показывает, какую часть общей дисперсии составляет межгрупповая дисперсия. Если корреляционное отношение выражено в процентах, то оно покажет, на сколько процентов общая вариация зависит от вариации фактора группировки:

$$\eta = \sqrt{\frac{\bar{\delta}^2}{\sigma^2}} \cdot 100. \quad (9)$$

Покажем расчет корреляционного отношения на примере.

Пример 5. Даны 18 участков посева, из них 10 чистые, а 8 засоренные. По каждому участку имели место следующие урожайности (в ц с га):

Т а б л и ц а 8

Чистые посевы	Засоренные посевы
14	10
15	14
14	15
16	12
15	13
16	12
17	16
20	12
21	—
18	—
По 10 участкам общий сбор с 1 га 166	По 8 участкам общий сбор с 1 га 104

Средний сбор с 1 га чистых посевов составит:

$$\bar{x}_1 = \frac{166}{10} = 16,6 \text{ ц.}$$

Средний сбор с 1 га засоренных посевов составит:

$$\bar{x}_2 = \frac{104}{8} = 13 \text{ ц.}$$

Определим дисперсию по урожайности в каждой группе посевов (см. табл.).

Таблица 6

## Чистые посевы

## Засоренные посевы

Урожай- ность ( $x$ )	$x - \bar{x}$	$(x - \bar{x})^2$	Урожай- ность ( $x$ )	$x - \bar{x}$	$(x - \bar{x})^2$
14	-2,6	6,76	10	-3,00	9,00
15	-1,6	2,56	14	+1,00	1,00
14	-2,6	6,76	15	+2,00	4,00
16	-0,6	0,36	12	-1,00	1,00
15	-1,6	2,56	13	0,00	0,00
16	-0,6	0,36	12	-1,00	1,00
17	+0,4	0,16	15	+3,00	9,00
20	+3,4	11,56	12	-1,00	1,00
21	+4,4	19,36			
18	+1,4	1,96			
116	0	52,40	104	0	26,00

Дисперсия урожайности в первой группе (по чистым посевам) составит:  $\sigma_1^2 = \frac{\Sigma (x - \bar{x}_1)^2}{n} = \frac{52,40}{10} = 5,24$ , то же по второй группе:  $\sigma_2^2 = \frac{\Sigma (x - \bar{x}_2)^2}{n} = \frac{26}{8} = 3,25$ .

Внутригрупповая дисперсия:  $\bar{\sigma}^2 = \frac{\Sigma \sigma_i^2 m_i}{\Sigma m_i} = \frac{5,24 \cdot 10 + 3,25 \cdot 8}{18} = \frac{78,4}{18} = 4,35$ .

Определим общую среднюю:

$$\bar{x} = \frac{\Sigma \bar{x}_i m_i}{\Sigma m_i} = \frac{16,6 \cdot 10 + 13,8}{18} = \frac{166 + 104}{18} = \frac{270}{18} = 15 \text{ ц.}$$

Межгрупповая дисперсия  $\bar{\sigma}^2 = \frac{\Sigma (\bar{x}_i - \bar{x})^2 m_i}{\Sigma m_i} =$

$$\frac{(16,6 - 15)^2 \cdot 10 + (13 - 15)^2 \cdot 8}{18} = \frac{(1,6)^2 \cdot 10 + (-2)^2 \cdot 8}{18} = \frac{25,6 + 32}{18} = 3,20.$$

Согласно правилу сложения вариации:

$$\sigma^2 = \bar{\sigma}^2 + \bar{\sigma}^2,$$

следовательно, общая вариация  $\sigma^2 = 4,35 + 3,20 = 7,55$  и можно определить величину корреляционного отношения:

$$\eta = \sqrt{\frac{\bar{\sigma}^2}{\sigma^2}} = \sqrt{\frac{3,20}{7,55}} = \sqrt{0,42} = 0,65, \text{ или } 65\%.$$

Корреляционное отношение, равное 0,65, или 65%, указывает на то, что вариация в урожайности на 65% зависит от засоренности или незасоренности посевов. Таким образом, корреляционное отношение является показателем тесноты связи.

§ 7. Дисперсионный анализ. Дисперсионный анализ служит для подтверждения зависимости при небольшом числе наблюдений. Дисперсия генеральной совокупности больше дисперсии выборки в  $\frac{n}{n-1}$  раз, где  $n-1$  — число степеней свободы независи-

мого варьирования. Поэтому для исчисления дисперсии по выборочным данным следует относить сумму квадратов отклонений не к числу наблюдений, а к числу степеней свободы вариации.

Для отношения  $F$  межгрупповой дисперсии к внутригрупповой  $P$ . Фишер нашел закон распределения вероятностей (см. в приложении распределение Фишера). Этот закон позволяет вычислить критическое значение  $F$ , случайное превышение которого при данных числах степеней свободы вариации в межгрупповой и внутригрупповой дисперсиях маловероятно. Сравнивая с ним вычисленное отношение, можно сделать вывод о достоверности зависимости, проявляющейся в увеличении межгрупповой дисперсии.

Если вычисленные данные  $F$  превышают табличные их значения, то зависимость признается достоверной.

Общая схема дисперсионного анализа различна в зависимости от числа группировочных признаков. При группировке по одному признаку дисперсионный анализ производится следующим образом:

1. Группировка.
2. Определение средних по группам и общей средней.
3. Нахождение суммы квадратов отклонений групповых средних от общей средней, взвешенных по численностям групп.
4. Определение сумм внутригрупповых дисперсий, взвешенных по численностям групп.
5. Нахождение числа степеней свободы вариации по группам и внутри групп.
6. Определение межгрупповой и внутригрупповой дисперсий (с учетом чисел степеней свободы) и отношения первой ко второй

$$F = \frac{s^2}{\sigma^2}. \quad (10)$$

7. Нахождение критического  $F$  по таблице Фишера.
8. Сравнение найденного  $F$  с критическим и вывод о достоверности или недостоверности найденной величины.

Общая схема дисперсионного анализа может быть представлена следующей таблицей (при изучении влияния одного фактора):

Т а б л и ц а 10

Виды вариации	Сумма квадратов отклонений	Число степеней свободы	Дисперсия	Отношение дисперсий	Табличное F
Межгрупповая	$\Sigma(\bar{x}_i - \bar{x})^2 m_i$	$n - 1$	$\bar{s}^2 = \frac{\Sigma(\bar{x}_i - \bar{x})^2 m_i}{n - 1}$		
Внутригрупповая . . . . .	$\Sigma(x - \bar{x}_i)^2 m$	$\Sigma m - n - 2$	$\bar{\sigma}^2 = \frac{\Sigma(x - \bar{x}_i)^2 m}{\Sigma m - n - 2}$	$\frac{\bar{s}^2}{\bar{\sigma}^2}$	
Общая . . . . .	$\Sigma(x - \bar{x})^2 m$	$\Sigma m - 1$	$\sigma^2 = \frac{\Sigma(x - \bar{x})^2 m}{\Sigma m - 1}$		

Если величина исчисленного отношения дисперсий больше величины табличного (критического) значения, то проверяемое влияние факториального признака на результативный признак в изучаемой совокупности достоверно. Если же величина исчисленного отношения дисперсий окажется меньше табличной величины, то результаты данного статистического наблюдения не могут быть признаны достоверными, и вывод о влиянии факториального признака на результативный не подтверждается данным материалом.

Пр и м е р 6. Данные наблюдения по урожайности разгруппированы по срокам уборки:

Т а б л и ц а 11

Сроки уборки	Урожаи пшеницы (в ц с 1 га) по отдельным участкам
Своевременный . . . . .	16,5; 16,2; 18,9; 20,1; 19,3; 10,1; 12,8; 15,0
С некоторым опозданием . . . . .	16,7; 16,3; 14,0; 15,0; 16,7; 12,4; 7,9; 9,8; 14,4; 10,8; 11,1; 13,0; 10,7
С сильным опозданием	10,7; 9,0; 13,9; 9,4; 11,9; 11,3; 10,5; 9,9; 7,4

Сумма квадратов урожаев  $\Sigma y^2 = 5456,01$ .

Сумма урожаев  $\Sigma y = 391,7$ . Число наблюдаемых участков 30.

Отсюда находим общую сумму квадратов отклонений или

$$\Sigma(y - \bar{y})^2 = 5456,01 - \frac{(391,7)^2}{30} = 341,72.$$

В следующей таблице выявляется колеблемость групповых средних около общей средней — сумма квадратов отклонений,

которая дает вариацию за счет фактора группировки (сроков уборки).

Т а б л и ц а 12

Сроки уборки	Число наблюдений $m_i$	Средний урожай $\bar{x}_i$	$\bar{x}_i - \bar{x}$	$(\bar{x}_i - \bar{x})^2$	$(\bar{x}_i - \bar{x})^2 m$
Своевременный . . . .	8	16,113	3,06	9,3636	74,9088
С некоторым опозданием .	13	12,985	-0,07	0,0349	0,0637
С сильным опозданием .	9	10,444	2,51	6,3001	56,7009
		13,05			131,6734

$$\Sigma (\bar{x}_i - \bar{x})^2 m = 131,67.$$

Сумма квадратов отклонений внутри групп равна:

$$341,72 - 131,67 = 210,05.$$

Составляем снова таблицу для анализа дисперсий.

Т а б л и ц а 13

Источники вариации	Сумма квадратов отклонений	Число степеней свободы	Дисперсия	Отношение дисперсий	Табличное значение при вероятности	
					0,95	0,99
Межгрупповая . .	131,67	2	65,83	8,46	3,35	5,49
Внутригрупповая .	210,05	27	7,78	1	База сравнения	
Общая . . . .	341,72	29	—	—	—	—

По таблицам Фишера (приложение VI) находим, что отношение дисперсий  $F$  в выборках из той же генеральной совокупности при данном числе степеней свободы не должно превышать 3,35 (вероятность случайного превышения = 0,05) или 5,49 (вероятность 0,01). По данным же нашего примера отношение дисперсий равно 8,46. Таким образом, полученный эффект влияния сроков уборки на размер урожая вполне достоверен.

§ 8. Теснота связи. Наиболее распространенным показателем тесноты связи является линейный коэффициент корреляции, который может быть исчислен по формуле:

$$R = \frac{\bar{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}, \quad (11)$$

где

$$\overline{xy} = \frac{\Sigma xy}{n}; \quad \sigma_x^2 = \overline{x^2} - \bar{x}^2;$$

$$\bar{x} = \frac{\Sigma x}{n}; \quad \sigma_y^2 = \overline{y^2} - \bar{y}^2;$$

$$\bar{y} = \frac{\Sigma y}{n}.$$

Применяется также корреляционное отношение, главным образом для измерения тесноты связи при наличии криволинейной зависимости; но может быть применено и при наличии линейной связи:

$$\tau = \sqrt{1 - \frac{\sigma_{y(x)}^2}{\sigma_y^2}}. \quad (12)$$

**Пример 7.** Даны значения признаков  $x$  и  $y$ , по которым произведен расчет линейного коэффициента корреляции, как показателя тесноты связи.

Таблица 14

$x$	$y$	$x^2$	$xy$	$y^2$
1	3	1	3	9
3	5	9	15	25
5	4	25	20	16
9	12	35	38	50

$$\bar{x} = \frac{\Sigma x}{n} = \frac{9}{3} = 3; \quad \bar{y} = \frac{\Sigma y}{n} = \frac{12}{3} = 4;$$

$$\overline{xy} = \frac{\Sigma xy}{n} = \frac{35}{3} = 12 \frac{2}{3};$$

$$\sigma_x^2 = \frac{\Sigma x^2}{n} - \bar{x}^2 = \frac{35}{3} - 9 = 11 \frac{2}{3} - 9 = 2 \frac{2}{3} = \frac{8}{3};$$

$$\sigma_y^2 = \frac{\Sigma y^2}{n} - \bar{y}^2 = \frac{50}{3} - 16 = 16 \frac{2}{3} - 16 = \frac{2}{3};$$

$$R = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y} = \frac{12 \frac{2}{3} - 3 \cdot 4}{\frac{8}{3} \cdot \frac{2}{3}} = \frac{\frac{2}{3}}{\frac{4}{3}} = 0,5.$$



Пр и м е р 8. Вычислим величину линейного коэффициента для данных об урожайности и количестве внесенных удобрений.

Таблица 15

Удобрения $x$	Урожайн. $y$	$xy$	$x^2$	$y^2$	$\bar{y}_x$	$y$	$y_x$	$(y - y_x)^2$
1	10	10	1	100	9,3		0,7	0,49
2	12	24	4	144	10,2		1,8	3,24
3	11	33	9	121	11,1		-0,1	0,01
4	13	52	16	169	12,0		1,0	1,00
5	12	60	25	144	12,9		-0,9	0,81
6	13	78	36	169	13,8		-0,8	0,64
7	15	105	49	225	14,7		+0,3	0,09
28	86	362	140	1072				6,28

$$\overline{xy} = \frac{362}{7} = 51,714; \quad \bar{x} = \frac{28}{7} = 4; \quad \bar{y} = \frac{86}{7} = 12,3;$$

$$\sigma_x = \sqrt{\frac{140}{7} - (4)^2} = \sqrt{20 - 16} = \sqrt{4,00} = 2,000;$$

$$\sigma_y = \sqrt{\frac{1072}{7} - (12,3)^2} = \sqrt{153,143 - 151,29} = \sqrt{1,852} = 1,361;$$

$$R = \frac{51,714 - 4,12}{2,000 \cdot 3,047} = \frac{3,714}{6,094} = 0,609.$$

Линейный коэффициент корреляции при отсутствии связи равен нулю, а в случае полной функциональной связи равен единице. Отсюда, если среднее произведение равно произведению средних, то связь отсутствует, т. е. если имеет место  $\overline{xy} = \bar{x} \cdot \bar{y}$ , то это указывает на отсутствие корреляционной связи. При полной функциональной связи разность между средним произведением и произведением средних должна равняться произведению средних квадратических отклонений по обоим признакам, т. е.  $\overline{yx} - \bar{y} \cdot \bar{x} = \sigma_x \cdot \sigma_y$ . Положительная величина линейного коэффициента корреляции указывает на прямую связь между признаками, отрицательная — на обратную. Это значит, если  $\overline{xy} > \bar{x} \cdot \bar{y}$  — имеет место прямая связь, при  $\overline{xy} < \bar{x} \cdot \bar{y}$  — имеет место обратная связь.

Для нахождения корреляционного отношения мы должны предварительно найти корреляционное уравнение  $\bar{y}_x = a_0 + a_1x$ , для чего воспользуемся примером о возрасте поросят и их живом весе<sup>1</sup>.

<sup>1</sup> См. В. С. Немчинов, Сельскохозяйственная статистика с основами общей теории, М., 1945

Пример 9.

Таблица 16

Возраст (в неделях) $x$	Живой вес		$y - \bar{y}_x$	$(y - \bar{y}_x)^2$	$y^2$
	наблюден- ный $y$	вычислен- ный $\bar{y}_x$			
0	1,3	0,872	+0,428	0,183184	1,69
1	2,5	2,281	+0,219	0,047961	6,25
2	3,9	3,691	+0,209	0,043681	15,21
3	5,2	5,101	+0,099	0,009801	27,04
4	5,3	6,511	-1,211	1,466521	28,09
5	7,5	7,921	-0,421	0,177241	56,25
6	9,0	9,331	-0,331	0,109561	81,00
7	10,8	10,741	+0,059	0,003481	116,64
8	13,1	12,150	+0,930	0,902500	171,61
58,6				2,954	502,47

Для расчета корреляционного отношения  $\eta$  найдем

$$\sigma_y^2 = \frac{\Sigma y^2}{n} - \bar{y}^2; \quad y = \frac{\Sigma y}{n} = \frac{58,6}{9} = 6,511;$$

$$\sigma_y^2 = \frac{502,47}{9} - 42,25 = 13,58;$$

$$\sigma_{y(x)}^2 = \frac{\Sigma (y - \bar{y}_x)^2}{n} = \frac{2,954}{9} = 0,330;$$

$$\eta = \sqrt{1 - \frac{\sigma_{y(x)}^2}{\sigma_y^2}} = \sqrt{1 - \frac{0,330}{13,58}} = \sqrt{1 - 0,024} = 0,987.$$

Зная величину линейного коэффициента корреляции  $R$ , можно легко определить значения параметров корреляционного уравнения по следующим соотношениям:

$$a_1 = R \frac{\sigma_y}{\sigma_x}, \text{ где } \sigma_x = \sqrt{\bar{x}^2 - \bar{x}^2}; \quad \sigma_y = \sqrt{\bar{y}^2 - \bar{y}^2}.$$

В приведенном выше примере  $\sigma_x = 2,582$ ;  $\sigma_y = 3,685$ ;  $R = 0,987$ ;

$$a_1 = 0,987 \frac{3,685}{2,582} = 1,4098.$$

Величина выравненного значения результативного признака будет равна:

$$\bar{y}_x = a_1(x - \bar{x}) + \bar{y},$$

а в примере:

$$\bar{y}_x = 1,4098(x - 4) + 6,50;$$

$$\bar{y}_x = 1,4098x - 5,639 + 6,511;$$

$$\bar{y}_x = 0,872 + 1,4098x.$$

Найденная величина коэффициента корреляции (пример 8) свидетельствует о значительной зависимости между урожайностью и количеством внесенных удобрений. Однако малое число наблюдений ( $n=7$ ) не позволяет особенно полагаться на этот вывод. Чтобы оценить надежность этого результата, вычисляют среднюю ошибку коэффициента корреляции по формуле:

$$\sigma_R = \frac{1 - R^2}{\sqrt{n}}, \quad (13)$$

которая в первом примере равна

$$\sigma_R = \frac{1 - 0,72}{\sqrt{7}} = \pm \frac{0,28}{2,67} = \pm 0,10.$$

Окончательный ответ  $R = 0,60 \pm 0,24$ , что говорит о вероятности значительных колебаний полученного коэффициента корреляции при аналогичных наблюдениях:

от  $0,60 - 0,24 = 0,36$  до  $0,60 + 0,24 = 0,84$ .

Для примера 9:

$$\sigma_R = \frac{1 - 0,976}{9} = \pm \frac{0,024}{3} = \pm 0,008.$$

Как видно из названия линейного коэффициента корреляции, он применяется только в случае наличия линейной связи между изучаемыми признаками. В случае же криволинейной зависимости расчет линейного коэффициента корреляции для определения тесноты связи приводит к значительным ошибкам. Поэтому применяется корреляционное отношение. При этом расчет корреляционного отношения ведется исходя из найденных выравненных значений признака —  $\bar{y}_x$ . Корреляционное отношение может исчисляться по формуле:

$$\eta = \sqrt{\frac{\overline{\delta_y^2(x)}}{\sigma_y^2}}, \quad (14)$$

где

$$\overline{\delta_y^2(x)} = \frac{\Sigma(\bar{y} - \bar{y}_x)^2}{n} = \sigma_y^2 - \sigma_{y(x)}^2,$$

а  $\sigma_y^2$  — известная нам величина:

$$\sigma_y^2 = \frac{\Sigma(y - \bar{y})^2}{n}.$$

Для примера 9 вычислим корреляционное отношение с помощью следующей таблицы:

Т а б л и ц а 17

$x$	$y$	$\bar{y}_x$	$\bar{y} - \bar{y}_x$	$(\bar{y} - \bar{y}_x)^2$
0	1,3	0,87	5,63	31,70
1	2,5	2,28	4,22	17,81
2	3,9	3,69	2,81	7,89
3	5,2	5,10	1,40	1,96
4	5,3	6,51	0,01	0,00
5	7,5	7,92	1,42	2,01
6	9,0	9,33	2,73	7,44
7	10,8	10,77	4,27	18,22
8	13,1	12,15	5,65	31,92
				119,25

$$\bar{\sigma}_{y(x)}^2 = \frac{119,25}{9} = 13,25;$$

$$\sigma_y^2 = 13,58;$$

$$\eta = \sqrt{\frac{13,25}{13,58}} = 0,987.$$

Величина корреляционного отношения и линейного коэффициента корреляции имеет одинаковый смысл: оба они показывают тесноту связи между изучаемыми признаками.

Если  $\eta$  близко к 1 — связь тесная, если близко к 0 — связь слабая.

**§ 9. Корреляционная таблица.** Основные понятия теории корреляции и расчет показателей тесноты связи и корреляционных уравнений рассматривались до сих пор на примерах небольшого количества наблюдений. Но статистику приходится в большинстве случаев обрабатывать материалы значительного количества наблюдений, в которых каждая пара значений признаков может встречаться некоторое число раз.

Таким образом, возникает необходимость взвешивания этих данных по количествам их повторений. Проблема взвешивания данных известна нам из раздела I. В настоящем разделе принципы взвешивания, сохраняясь в основном, принимают все же специфический вид. Как уже указано в начале настоящего раздела, особенностью весов взаимозависимых признаков является то, что они могут быть удобно расположены в корреляционной таблице. В этом случае расчеты корреляционных уравнений, линейного коэффициента корреляции и корреляционного отношения удобно вести по самой корреляционной таблице.

Пример 10. Группировка заводов по двум признакам

Таблица 18

Среднесуточная переработка смен (тыс. шт) (y)		3,5— 4,5		4,5— 5,5		5,5— 6,5		6,5— 7,5		7,5— 8,5		8,5— 9,5		9,5— 10,5		10,5— 11,5		11,5— 12,5		$m_x$		$\Sigma y_i$		$\overline{y_i}$		$\Sigma x y_i = x \Sigma y_i$	
		4	3	5	4	6	5	7	6	8	7	9	8	10	9	11	10	12	11								
Производственные основные средства (млн. руб.) (x)		Средние интервалы																									
1	2																										
1,0—1,5	1,25	1																		1	4	4,0	5,0				
1,5—2,0	1,75	4		6		9		2												21	114	5,4	199,5				
2,0—2,5	2,25	5		9		15		10		2		1								42	250	6,0	562,5				
2,5—3,0	2,75			4		6		7		7		7		1						25	171	6,8	470,3				
3,0—3,5	3,25	3		3		2		7		8		1								24	161	6,7	523,3				
3,5—4,0	3,75					3		3		2		3		2						13	102	7,9	382,5				
4,0—4,5	4,25							2		2		1		2		2		1		10	93	9,3	395,3				
4,5—5,0	4,75											2						1		3	30	10,0	142,5				
5,0—5,5	5,25					1												1		2	19	9,5	99,8				
$m_y$		13		22		35		32		21		8		5		2		3		141	944	6,7	2780,7				
$y m_y$		52		110		210		224		168		72		50		22		36		944							
$y^2 m_y$		208		550		1260		1568		1344		646		500		242		432		6752							

Возьмем для этого данные о 141 сахаропесочном заводе Украинской ССР, сведенные в корреляционную таблицу по признаку основных производственных средств ( $x$ ) и среднесуточной переработки свеклы ( $y$ ).

Корреляционная таблица будет иметь следующий вид.

В левой части таблицы расположены значения факториального признака ( $x$ ) — производственные основные средства (в млн. руб.); В графах расположены значения результативного признака ( $y$ ) — среднесуточная переработка свеклы (в тыс. ц). В клетках таблицы располагаются численности заводов, имеющих данную среднесуточную переработку при данных основных производственных средствах. Например, число 10, стоящее в строке 3 графы 6, показывает, что из всех предприятий 10 предприятий имеют стоимость основных производственных средств от 2,0 до 2,5 млн. руб. Всего предприятий этой группы 42, из которых только 10 имеют среднесуточную переработку свеклы от 6,5 до 7,5 тыс. ц.

Расчетная графа 12 ( $m_x$ ) представляет собой построчный итог частот. Графа 13 ( $\Sigma m_y$ ) есть сумма произведений результативного признака ( $y$ ) на частоты соответствующей строки. Для строки 2, например,  $\Sigma m_y$  образуется из суммы произведений:

$$4 \times 4 + 5 \times 6 + 6 \times 9 + 7 \times 2 = 114.$$

В результате деления  $\frac{\Sigma m_y}{\Sigma m_x}$  получаем  $\bar{y}_i$ , т. е.  $\bar{y}_i$  получается делением результата графы 13 на результат графы 12.

Графа 15 получается как произведение результата графы 13 на данные графы 2.

Для корреляционной таблицы (для взвешенных данных) система нормальных уравнений примет вид:

$$\left. \begin{aligned} na_0 + a_1 \Sigma x m_x &= \Sigma m_y \\ a_0 \Sigma x m_x + a_1 \Sigma x^2 m_x &= \Sigma x m_y \end{aligned} \right\} \quad (15)$$

Для составления системы нормальных уравнений следует составить вспомогательную таблицу и дополнительно определить  $\Sigma x m_x$  и  $\Sigma x^2 m_x$ . Для данного примера будем иметь:

Таблица 19

$x$	$m_x$	$x m_x$	$x^2 m_x$
1,25	1	1,3	1,6
1,75	21	36,8	64,3
2,25	42	94,5	212,6
2,75	25	68,8	189,1
3,25	24	78,0	253,5
3,75	13	48,8	182,8
4,25	10	42,5	180,6
4,75	3	14,3	67,7
5,25	2	10,5	55,1
Итого . . . .	141	395,6	1207,3

С помощью основной корреляционной таблицы и данной вспомогательной находим систему уравнений:

$$141a_0 + 395,6a_1 = 944;$$

$$395,6a_0 + 1207,3a_1 = 2780,7,$$

откуда

$$a_0 = 2,9;$$

$$a_1 = 1,4.$$

Значит, корреляционное уравнение будет таким:  $\bar{y}_x = 2,9 + 1,4 x$ . Линейный коэффициент корреляции может быть рассчитан по формуле:

$$R = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{\sqrt{[n\Sigma y^2 - (\Sigma y)^2][n\Sigma x^2 - (\Sigma x)^2]}}. \quad (16)$$

Для взвешенных значений эта формула видоизменяется:

$$R = \frac{n\Sigma x m_{xy} - \Sigma x m_x \cdot \Sigma m_y}{\sqrt{[n\Sigma y^2 m_y - (\Sigma m_y)^2][n\Sigma x^2 m_x - (\Sigma x m_x)^2]}}, \quad (17)$$

а для данного примера составит:

$$R = \frac{141 \cdot 2780,7 - 395,6 \cdot 944}{\sqrt{[141 \cdot 6752 - (944)^2] \cdot [141 \cdot 1207,5 - (395,6)^2]}} = 0,65.$$

Общая дисперсия по корреляционной таблице исчисляется:

$$\sigma_y^2 = \frac{\Sigma y^2 m_y}{n} - \left( \frac{\Sigma m_y}{n} \right)^2 = \frac{6752}{141} - \left( \frac{944}{141} \right)^2 = 3,1.$$

Расчет корреляционного отношения потребует составления другой вспомогательной таблицы.

Т а б л и ц а 20

$x$	$m_x$	$\bar{y}_x$	$(\bar{y} - \bar{y}_x)$	$(\bar{y} - \bar{y}_x)^2$	$(\bar{y} - \bar{y}_x)^2 m_x$
1,25	1	4,7	-2,0	4,0	4,0
1,75	21	5,4	-1,3	1,69	35,5
2,25	42	6,5	-0,6	0,36	15,1
2,75	25	6,8	-0,1	0,01	0,3
3,25	24	7,5	0,8	0,64	15,4
3,75	13	8,2	1,5	2,25	29,3
4,25	10	8,9	2,2	4,84	48,0
4,75	3	9,6	2,9	8,41	25,22
5,25	2	10,3	3,6	12,96	25,92
Итого . . . .	141	6,7	—	—	198,7

Определяем межгрупповую дисперсию

$$\begin{aligned}\overline{\delta_x^2} &= \frac{\Sigma (\bar{y} - \bar{y}_x)^2 m_x}{\Sigma m_x}; \\ \overline{\delta_x^2} &= \frac{198,7}{141} = 1,4.\end{aligned}$$

Зная общую дисперсию и межгрупповую, можно определить корреляционное отношение

$$\eta = \sqrt{\frac{\overline{\delta_x^2}}{\sigma^2}} = \sqrt{\frac{1,4}{3,1}} = 0,66.$$

**§ 10. Множественная корреляция.** До сих пор рассматривались корреляционные связи лишь между двумя признаками. В более общем виде можно рассматривать зависимость между тремя, четырьмя и более признаками. В этом случае возникает необходимость во множественной корреляции, которую называют также совокупной. Поставив вопрос о зависимости урожайности от количества внесенных удобрений и от количества выпавших осадков, придем к понятию множественной зависимости.

Линейное уравнение зависимости результативного признака от двух факториальных признаков определяют по формуле:

$$\bar{y}_{x.v} = a_0 + a_1x + a_2v. \quad (18)$$

Для определения параметров  $a_0$ ,  $a_1$  и  $a_2$  по способу наименьших квадратов составим систему нормальных уравнений:

$$\left. \begin{aligned} na_0 + a_1\Sigma x + a_2\Sigma v &= \Sigma y \\ a_0\Sigma x + a_1\Sigma x^2 + a_2\Sigma xv &= \Sigma yx \\ a_0\Sigma v + a_1\Sigma xv + a_2\Sigma v^2 &= \Sigma yv \end{aligned} \right\} \quad (19)$$

Если же число факториальных признаков возрастет до трех, то линейное корреляционное уравнение будет таким:

$$\bar{y}_{x.v.z} = a_0 + a_1x + a_2v + a_3z, \quad (20)$$

а система нормальных уравнений будет содержать четыре уравнения, из которых находим параметры  $a_0$ ,  $a_1$ ,  $a_2$ ,  $a_3$ . Эта система примет вид:

$$\left. \begin{aligned} na_0 + a_1\Sigma x + a_2\Sigma v + a_3\Sigma z &= \Sigma y \\ a_0\Sigma x + a_1\Sigma x^2 + a_2\Sigma xv + a_3\Sigma xz &= \Sigma yx \\ a_0\Sigma v + a_1\Sigma xv + a_2\Sigma v^2 + a_3\Sigma v z &= \Sigma yv \\ a_0\Sigma z + a_1\Sigma xz + a_2\Sigma vz + a_3\Sigma z^2 &= \Sigma yz \end{aligned} \right\} \quad (21)$$

**Пример 11.** Найдем уравнение связи урожайности пшеницы на Безенчукской опытной станции по данным акад. Немчинова В. С. в зависимости от трех факторов:



# УРОЖАЙ ПШЕНИЦЫ И

## Вычисление вспомога

Годы	$x$	$z$	$v$	$y$	$xz$	$xv$	$zv$
1	2	3	4	5	6	7	8
1905	21	76	211	10,5	1 596	4 431	16 036
1906	1	49	268	6,6	49	268	13 132
1907	44	42	163	10,5	1 848	7 172	6 846
1908	25	36	222	6,1	900	5 550	7 992
1909	37	58	126	11,1	2 146	4 662	7 308
1910	53	58	145	14,7	3 074	7 685	8 410
1911	1	18	310	2,4	18	310	5 580
1912	24	89	175	12,1	2 136	4 200	15 575
1913	27	84	154	10,0	2 268	4 158	12 936
1914	36	52	171	10,8	1 872	6 156	8 892
1915	88	96	159	17,7	8 448	13 992	15 264
1916	14	93	141	8,2	1 302	1 974	13 113
1917	9	20	211	5,4	180	1 899	4 220
1918	14	89	124	11,8	1 246	1 736	11 036
1919	17	109	149	10,2	1 853	2 533	16 241
1920	32	57	214	9,3	1 824	6 848	12 198
1921	1	5	364	1,3	5	364	1 320
1922	19	49	171	9,6	931	3 249	8 879
1923	40	55	178	8,7	2 200	7 120	9 790
1924	8	18	309	3,0	144	2 472	5 562
1925	12	103	167	11,4	1 236	2 009	17 201
1926	66	74	180	13,3	4 884	11 880	13 320
1927	20	18	229	5,5	360	4 580	4 122
1928	66	71	135	14,1	4 686	8 910	9 585
1929	39	40	181	7,0	1 550	7 059	7 240
Сумма . . .	714	1459	4857	231,3	46 766	121 212	251 798
Символ . .	$\Sigma x$	$\Sigma z$	$\Sigma v$	$\Sigma y$	$\Sigma xz$	$\Sigma xv$	$\Sigma zv$

В таблице:  $x$ —осадки в миллиметрах от начала

$z$ —осадки в миллиметрах от начала

$v$ —испарения в миллиметрах от на

# МЕТЕОРОЛОГИЧЕСКИЕ ФАКТОРЫ

тельных величин

Т а б л и ц а 21

$x^2$	$z^2$	$v^2$	$y^2$	$xy$	$zy$	$vy$
9	10	11	12	13	14	15
441	5 776	44 521	110,25	220,5	798,0	2215,5
1	2 401	71 824	43,56	6,6	323,4	1768,8
1 936	1 764	26 569	110,25	462,0	441,0	1711,5
625	1 296	49 284	37,21	152,5	219,6	1354,2
1 369	3 364	15 876	123,21	410,7	643,8	1398,6
2 809	3 364	21 025	216,09	779,1	852,6	2131,5
1	324	96 100	5,76	2,4	43,2	744,0
576	7 921	30 625	146,41	290,4	1076,9	2117,5
729	7 056	23 716	100,00	270,0	840,0	1540,0
1 296	2 704	29 241	116,64	388,8	561,6	1846,8
7 744	9 216	25 281	313,29	1557,6	1699,2	2814,3
196	8 649	19 881	67,24	114,8	762,6	1156,2
81	400	44 521	29,16	48,6	108,0	1139,4
196	7 921	15 376	139,24	165,2	1050,2	1463,2
289	11 881	22 201	104,04	173,4	1111,8	1519,8
1 024	3 249	45 796	86,49	297,6	530,2	1990,2
1	25	132 496	1,69	1,3	6,5	473,2
361	2 401	29 241	92,16	182,4	470,4	1641,6
1 600	3 025	31 684	75,69	348,0	478,5	1548,6
64	324	95 481	9,00	24,0	54,0	927,0
144	10 609	27 889	129,96	136,8	1174,2	1903,8
4 356	5 476	32 400	176,89	877,8	984,2	2394,0
400	324	52 441	30,25	110,0	99,0	1259,5
4 356	5 041	18 225	198,81	930,6	1001,1	1903,5
1 521	1 600	32 761	49,00	273,0	280,0	1267,0
32 116	106 111	1 034 455	2512,29	8224,1	15610,0	40229,7
$\Sigma x^2$	$\Sigma z^2$	$\Sigma v^2$	$\Sigma y^2$	$\Sigma xy$	$\Sigma zy$	$\Sigma vy$

сева до начала кущения;

кущения до начала цветения;

чала кущения до начала цветения.

Статистические данные, полученные в результате наблюдения, и расчеты представлены в табл. 21, откуда извлекаем необходимые данные для составления системы нормальных уравнений:

$$\begin{aligned}25a_0 + 714a_1 + 1459a_2 + 4857a_3 &= 231,3 \\714a_0 + 32116a_1 + 46766a_2 + 121212a_3 &= 8224,1 \\1459a_0 + 47766a_1 + 106111a_2 + 251798a_3 &= 15610,0 \\4857a_0 + 121212a_1 + 251798a_2 + 1034455a_3 &= 40229,7\end{aligned}$$

$$\begin{aligned}a_0 + 28,56a_1 + 58,16a_2 + 194,28a_3 &= 9,24 \\a_0 + 449804a_1 + 65,4986a_2 + 169,7647a_3 &= 11,5183 \\a_0 + 32,0535a_1 + 72,7286a_2 + 172,5826a_3 &= 10,6991 \\a_0 + 24,9561a_1 + 51,8423a_2 + 212,9822a_3 &= 8,2828\end{aligned}$$

$$\begin{aligned}\text{II} - \text{I} \quad 16,4204a_1 + 7,3385a_2 - 24,5153a_3 &= 2,2783 \\ \text{II} - \text{III} \quad 12,9269a_1 - 7,2300a_2 - 2,8179a_3 &= 0,8192 \\ \text{II} - \text{IV} \quad 7,0974a_1 + 20,8863a_2 - 40,3996a_3 &= 2,4163\end{aligned}$$

$$\begin{aligned}a_1 + 0,4469a_2 - 1,4930a_3 &= 0,1387 \\a_1 - 0,5637a_2 - 0,2180a_3 &= 0,0634 \\a_1 + 2,9428a_2 - 5,6924a_3 &= 0,340\end{aligned}$$

$$\begin{aligned}\text{I} - \text{II} \quad 1,0106a_2 - 1,2750a_3 &= 0,0753 \\ \text{III} - \text{II} \quad 3,5065a_2 - 5,9104a_3 &= 0,2770\end{aligned}$$

$$\begin{aligned}a_2 - 1,2616a_3 &= 0,0745 \\ \underline{a_2 - 1,6855a_3} &= 0,0790 \\ -0,4239a_3 &= 0,0045 \\ a_3 &= -0,0106\end{aligned}$$

$$\begin{aligned}a_2 &= 0,0745 - 1,2616 \cdot 0,0106 = 0,0745 - 0,0134 \\ a_2 &= 0,0611 \\ a_1 + 0,0611 \cdot 0,4469 + 1,4930 \cdot 0,0106 &= 0,1389 \\ a_1 &= 0,0273 + 0,0158 = 0,1389 \\ a_1 &= 0,1389 - 0,0431 = 0,0958 \\ a_0 + 28,56 \cdot 0,0958 + 58,16 \cdot 0,0611 - 194,28 \cdot 0,0106 &= 9,24 \\ a_0 + 2,7360 + 3,5536 - 2,0594 &= 9,24 \\ a_0 &= 11,2994 - 6,2896 = 5,0098\end{aligned}$$

Следовательно, корреляционное уравнение будет иметь вид:

$$\bar{y}_{x,z,v} = 5,0098 + 0,0958x + 0,06112z - 0,0106v.$$

При изучении тесноты связи для множественной зависимости можно сначала вычислить ряд простых коэффициентов корреляции. Например, при изучении связи между результативным признаком ( $y$ ) и двумя факториальными признаками ( $x$ ,  $v$ ) можно определить тесноту связи между  $y$  и  $x$ , между  $y$  и  $v$  и между  $x$  и  $v$ , т. е. вычислить линейные коэффициенты корреляции:

$$R_{y(x)}; R_{y(v)}; R_{x(v)}.$$

§ 11. Если стоит задача определения тесноты связи результативного признака от двух факториальных, находят совокупный коэффициент корреляции:

$$R_{y(x,v)} = \sqrt{\frac{R_{y(x)}^2 + R_{y(v)}^2 - 2R_{y(x)}R_{y(v)}R_{x(v)}}{1 - R_{x(v)}^2}}. \quad (22)$$

$$R_{y(x)} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y};$$

$$R_{y(v)} = \frac{\overline{yv} - \bar{y} \cdot \bar{v}}{\sigma_v \sigma_y};$$

$$R_{x(v)} = \frac{\overline{xv} - \bar{x} \cdot \bar{v}}{\sigma_x \sigma_v},$$

где

$$\overline{xy} = \frac{\Sigma xy}{n}; \quad \bar{y} = \frac{\Sigma y}{n};$$

$$\overline{yv} = \frac{\Sigma yv}{n}; \quad \bar{x} = \frac{\Sigma x}{n};$$

$$\overline{xv} = \frac{\Sigma xv}{n}; \quad \bar{v} = \frac{\Sigma v}{n};$$

$$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2}; \quad \sigma_y = \sqrt{\overline{y^2} - \bar{y}^2}; \quad \sigma_v = \sqrt{\overline{v^2} - \bar{v}^2}.$$

Но может возникнуть также задача определить тесноту связи между результативным признаком и одним из факториальных признаков, исключив связь, возникающую за счет второго факториального признака и связь, существующую между самими факториальными признаками. Возникает проблема частной корреляции.

§ 12. Частный коэффициент корреляции в случае изучения зависимости  $y$  от  $x$  при исключении влияния  $v$  определяется по формуле:

$$r_{y(x)} = \frac{R_{y(x)} - R_{y(v)} \cdot R_{x(v)}}{\sqrt{[1 - R_{y(v)}^2][1 - R_{x(v)}^2]}}. \quad (23)$$

Частный коэффициент корреляции в случае изучения зависимости  $y$  от  $v$  при исключении влияния  $x$  определяется:

$$xR_{y(v)} = \frac{R_{y(v)} - R_{y(x)} R_{v(x)}}{\sqrt{[1 - R_{y(x)}^2] [1 - R_{v(x)}^2]}}. \quad (24)$$

В случае зависимости результативного признака  $y$  от трех факториальных признаков  $x, v, z$  совокупный коэффициент корреляции выразится формулой:

$$R_{y(x, v, z)}^2 = 1 - [1 - R_{y(v, z)}^2] [1 - v, z R_{y(x)}^2], \quad (25)$$

где

$$v, z R_{y(x)}^2 = \frac{v R_{y(x)}^2 - v R_{y(z)}^2 v R_{y(z)}^2}{[1 - v R_{y(z)}^2] [1 - v R_{y(v)}^2]}. \quad (26)$$

Зная частные коэффициенты корреляции при множественной зависимости, можно определить совокупный коэффициент корреляции. Например, при зависимости результативного признака от трех факторов  $y, x, z, v$  и найдя  $R_x(z), R_x(v), R_y(x), R_z(x), R_y(z), R_z(v)$ , составим два определителя, из которых найдем совокупный коэффициент корреляции. Он будет равен:

$$R^2 = \begin{vmatrix} 1 & R_{xz} & R_{xv} & R_{yx} \\ R_{xz} & 1 & R_{zv} & R_{yz} \\ R_{xv} & R_{zv} & 1 & R_{yv} \\ R_{yx} & R_{yz} & R_{yv} & 0 \end{vmatrix} : \begin{vmatrix} 1 & R_{xz} & R_{xv} \\ R_{xz} & 1 & R_{zv} \\ R_{xv} & R_{zv} & 1 \end{vmatrix} \quad (27)$$

При зависимости  $y$  от двух факторов  $x, v$  совокупный коэффициент корреляции составит:

$$R^2 = \begin{vmatrix} 1 & R_{xv} & R_{yx} \\ R_{xv} & 1 & R_{yv} \\ R_{yx} & R_{yv} & 0 \end{vmatrix} : \begin{vmatrix} 1 & R_{xv} \\ R_{xv} & 1 \end{vmatrix} \quad (28)$$

**Пример 12.** Проследим это на примере зависимости между объемом ( $y$ ), диаметром ( $x$ ) и высотой ( $z$ ) модельных деревьев, выбранных из Сурской лесной дачи Пинежского бассейна<sup>1</sup>.

<sup>1</sup> См. Дьячков А. Н., Математическая статистика в применении к лесному делу, Архангельск, 1951.

Таблица 22

$x$	$z$	$y$	$x - \bar{x}$	$z - \bar{z}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(z - \bar{z})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$	$(x - \bar{x})(z - \bar{z})$	$(y - \bar{y})(z - \bar{z})$
15,5	14,1	0,14	-3,05	-2,1	-0,149	9,30	4,41	0,022	6,41	0,508	0,313
20,5	18,5	0,52	+1,95	+2,3	+0,231	3,76	5,29	0,053	4,49	0,450	0,531
12,0	10,6	0,07	-6,55	-5,4	-0,219	42,90	29,16	0,046	37,37	1,326	1,183
12,5	12,4	0,08	-6,05	-3,8	-0,209	36,60	14,44	0,044	22,99	1,264	0,794
16,5	16,1	0,19	+2,05	-0,1	-0,099	4,20	0,01	0,010	0,21	0,205	0,010
24,5	19,2	0,45	+5,95	3,0	+0,161	35,48	9,00	0,026	17,85	0,953	0,486
24,5	20,6	0,48	+5,95	+4,4	+0,191	35,48	19,36	0,036	26,18	1,136	0,840
23,5	20,6	0,49	+4,95	4,4	+0,201	23,50	19,36	0,040	21,78	0,995	0,884
12,0	10,9	0,06	-6,55	-5,3	-0,229	42,90	26,09	0,032	34,72	1,500	1,234
24,0	18,8	0,41	+5,45	+2,6	+0,121	29,70	6,76	0,015	14,17	0,569	0,315
185,5	162,0	2,89	—	—	—	263,82	133,88	0,324	186,17	8,596	6,567

$$\sigma_x^2 = 26,382; \quad \sigma_z^2 = 13,388; \quad \sigma = 0,0324;$$

$$\sigma_x = 5,14; \quad \sigma_z = 3,66; \quad \sigma_y = 0,18;$$

$$\bar{x} = 18,55; \quad \bar{z} = 16,20; \quad \bar{y} = 0,289.$$

Найдем коэффициенты корреляции между  $x$  и  $z$ ,  $y$  и  $x$ ,  $y$  и  $z$ .

$$R_{xz} = \frac{\Sigma(x - \bar{x})(z - \bar{z})}{n \sigma_x \sigma_z} = \frac{186,17}{10 \cdot 5,14 \cdot 3,66} = 0,99;$$

$$R_{yx} = \frac{\Sigma(y - \bar{y})(x - \bar{x})}{n \sigma_y \sigma_x} = \frac{8,996}{10 \cdot 5,14 \cdot 0,18} = 0,97;$$

$$R_{yz} = \frac{\Sigma(y - \bar{y})(z - \bar{z})}{n \sigma_y \sigma_z} = \frac{5,567}{10 \cdot 3,66 \cdot 0,18} = 0,99.$$

Найдем совокупный коэффициент корреляционной зависимости объема дерева от диаметра и высоты. Он исчисляется как частное двух определителей, один из которых будет определителем 3-го порядка, а другой — определителем 2-го порядка:

$$R_{y(x,z)}^2 = \begin{vmatrix} 1 & 0,99 & 0,97 \\ 0,99 & 1 & 0,99 \\ 0,97 & 0,99 & 0 \end{vmatrix} : \begin{vmatrix} 1 & 0,99 \\ 0,99 & 1 \end{vmatrix}$$

$$R = 0,99$$

**§ 13. Оценка выборочных коэффициентов корреляции.** При нахождении коэффициентов корреляции по выборочным данным возникает вопрос о связи выборочного коэффициента корреляции с генеральным. Для нормальной корреляции (когда оба признака подчинены закону нормального распределения), связь между ними установлена Р. Фишером, который нашел распределение величины:  $z = \frac{1}{2} \ln \frac{1+r}{1-r}$ , являющейся функцией выборочного коэффициента корреляции. Это распределение имеет вид:

$$D_n(z) = \frac{n-2}{\sqrt{2\pi(n-1)}} e^{-\frac{(n-1)(z-\gamma)^2}{2}} \left[ 1 + \frac{n(z-\gamma)}{2} \right], \quad (29)$$

$$\text{где } \gamma = \frac{1}{2} \ln \frac{1+\rho}{1-\rho},$$

а  $\rho$  — коэффициент корреляции в генеральной совокупности.

Статистическими характеристиками  $z$  является средняя  $\bar{z} = \frac{r}{2(n-1)} + \gamma$  и среднее квадратическое отклонение  $\sigma_z = \frac{1}{\sqrt{n-3}}$ .

Когда выборочный коэффициент корреляции  $r$  находится в пределах  $0 \leq r \leq 1$ , найденная функция изменяется в пределах  $0 \leq z \leq \infty$ .

Практически можно считать распределение Фишера нормальным. Распределение Фишера позволяет решить следующие задачи:

- а) оценить, насколько существенна разница между выборочным ( $r$ ) и генеральным ( $\rho$ ) коэффициентами корреляции;
- б) решить вопрос о существенности расхождения между двумя выборочными коэффициентами корреляции;
- в) вывести наилучший результат из выборочных коэффициентов, исчисленных в выборках из одной генеральной совокупности.

Решая первую задачу, мы должны иметь сведения о выборочном коэффициенте корреляции  $r$ , генеральном  $\rho$  и числе наблюдений.

**Пример 13.** Пусть  $r = 0,55$ ,  $\rho = 0,6$ ,  $n = 25$ . Тогда по приложению V

$$z = \frac{1}{2} \ln \frac{1+0,55}{1-0,55} = 0,618;$$

$$\sigma_z = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{22}} = 0,213;$$

$$\gamma = \frac{1}{2} \ln \frac{1+\rho}{1-\rho} = \frac{1}{2} \ln \frac{1+0,6}{1-0,6} = 0,693.$$

Из нормального распределения критерия следует

$$P[|z - \gamma| \geq t\sigma_z] = 1 - F(t). \quad (29)$$

Найдем

$$t = \frac{|z - \gamma|}{\sigma_z} = \frac{|0,618 - 0,693|}{0,213} = \frac{0,075}{0,213} = 0,35;$$

$$F(0,35) = 0,365; \quad 1 - F(0,35) = 0,635.$$

Следовательно, вероятность того, что существуют случайные значения коэффициента корреляции, отличающиеся от действительного коэффициента корреляции  $\rho = 0,6$  не менее, чем найденный нами коэффициент корреляции  $r = 0,55$ , равна 0,635. Она говорит нам о случайном расхождении между  $r$  и  $\rho$ .

**Пример 14.** Для решения второй задачи допустим, что имеем для двух выборок ( $n_1 = 25$  и  $n_2 = 19$ ) коэффициенты корреляции  $r_1 = 0,5$  и  $r_2 = 0,35$ .

По приложению V найдем для  $r_1$  и  $r_2$  значения  $z_1$  и  $z_2$ :

$$z_1 = 0,549; \quad z_2 = 0,365.$$



В этом случае:

$$\sigma_{z_1} = \frac{1}{\sqrt{25-3}} = 0,213; \quad \sigma_{z_2} = \frac{1}{\sqrt{19-3}} = 0,25.$$

Найдем разность:

$$z_1 - z_2 = 0,184.$$

Эта разность также имеет нормальное распределение, дисперсия которого

$$\sigma^2 = \sigma_{z_1}^2 + \sigma_{z_2}^2 = (0,213)^2 + (0,250)^2 = 0,10.$$

$$P [z_1 - z_2 | \geq t\sigma] = 1 - F(t); \quad c^2 = 0,10; \quad \sigma = 0,31;$$

$$F(0,6) = 0,451; \quad 1 - F(0,6) = 0,549; \quad t = \frac{0,184}{0,31} = 0,6.$$

Вероятность расхождений  $z_1$  и  $z_2$ , превышающих 0,184, имеет достаточно большое значение — 0,549. Поэтому наблюдаемые коэффициенты корреляции мало отличаются друг от друга.

Пример 15. Для решения третьей задачи возьмем те же данные:  $n_1 = 25$ ;  $n_2 = 19$ ;  $r_1 = 0,5$ ;  $r_2 = 0,35$ ; тогда  $z_1 = 0,549$ ;  $z_2 = 0,365$ ;  $\sigma_{z_1}^2 = 0,0441$ ;  $\sigma_{z_2}^2 = 0,0625$ . Среднее значение будет найдено как средняя арифметическая взвешенная:

$$\begin{aligned} \bar{z} &= \frac{z_1 n_1 + z_2 n_2}{n_1 + n_2} = \frac{0,549 \cdot 25 + 0,365 \cdot 19}{25 + 19} = \frac{13,625 + 6,835}{44} = \\ &= \frac{20,46}{44} = 0,46. \end{aligned}$$

По величине  $z = \frac{1}{2} \ln \frac{1+r}{1-r}$  можно вычислить  $r = \frac{e^{2z} - 1}{e^{2z} + 1}$ .

Подставив значение  $z=0,46$ , найдем  $r=0,43$ . Значит, если  $r_1$  и  $r_2$  получились в результате выборок из одной генеральной совокупности, то наилучшим результатом обладает коэффициент корреляции, равный 0,43.

## **ПРАКТИЧЕСКИЕ ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ**

### **Задание I по теме «Вариационный ряд»**

По данным о замерах 150 ампул с № ... по № .... (приложение I, табл. 1):

1. Построить комбинационную таблицу распределения ампул по объему и большому диаметру.

Примечание. По объему взять дискретные значения признака, по большому диаметру произвести группировку материала с тем, чтобы число групп было не менее 6 и не более 7.

2. По каждой группе ампул, отличающихся большим диаметром, исчислить средний объем и дисперсию (с точностью до 0,0001).

3. Исчислить среднюю из дисперсий объема по группам, отличающихся большим диаметром.

4. Определить общую среднюю объема ампул по групповым средним с учетом численности групп.

5. Найти межгрупповую дисперсию по объему.

6. Вычислить корреляционное отношение и сделать вывод о размерах влияния большого диаметра на объем пополнения ампул.

### **Задание II по теме «Выборочный метод»**

По данным об объеме наполнения 150 ампул, приведенных в задании I:

1. Построить ряд распределения ампул по объему.

2. По этому ряду распределения исчислить средний объем, дисперсию, среднее квадратическое отклонение, долю стандартных ампул (объемом до 1,2 см<sup>3</sup> включительно), дисперсию альтернативного признака (соответствия стандарту).

3. Отобрать механическим способом 30 ампул и исчислить для них выборочную среднюю по объему.

4. Вычислить предельную ошибку выборки, если вероятность возникновения этой ошибки равна 0,997.

5. Найти границы, в которых лежит выборочная средняя, и сопоставить их с результатом, достигнутым в п. 3 настоящего задания.

6. Уменьшить полученную в п. 4 предельную ошибку выборки в два раза и определить при тех же условиях необходимую численность выборки.

### **Задание III по теме «Кривые распределения»**

По данным о наружном диаметре 100 плашек с № ... по № ... (приложение I, табл. 2):

1. Составить эмпирическое распределение наружного диаметра плашек с числом групп 8—10.

2. Исчислить теоретические частоты наружного диаметра плашек, полагая, что эмпирические данные следуют закону нормального распределения.

3. Построить эмпирическую кривую распределения наружного диаметра.

4. Построить теоретическую кривую распределения на том же графике.

5. Определить согласие эмпирического и теоретического распределений по критерию Б. С. Ястремского.

### **Задание IV по теме «Корреляция»**

По данным о 150 ампулах, приведенных в I задании (приложение I):

1. Построить корреляционную таблицу по объему и длине ампул.

Примечание. По объему берутся дискретные значения признака, по длине производится группировка с выделением 6—7 групп.

2. По данным корреляционной таблицы исчислить среднюю величину объема ампул по группам с данной их длиной и результат исчисления нанести на график.

3. Произвести необходимые расчеты по корреляционной таблице и найти корреляционное уравнение зависимости объема ампул от длины.

4. Рассчитать по корреляционному уравнению величину объема при данной длине и нанести на график.

5. Исчислить линейный коэффициент корреляции и сделать выводы о размерах зависимости объема ампул от их длины.

---

## ЗАМЕРЫ АМПУЛ \*

№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм	№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм
1	1,1	23	10	37	1,1	19,5	10,4
2	1,1	21	10,5	38	1,2	20,0	10,3
3	1,0	18	10,5	39	1,2	22,5	10,3
4	1,1	21	10,0	40	1,0	20,0	10,0
5	1,2	21,9	10,3	41	1,1	22,0	10,0
6	1,1	24	9,6	42	1,0	21,0	9,8
7	1,1	20	10,3	43	1,0	21,0	10,0
8	1,1	20,5	10,1	44	1,0	20,4	9,9
9	1,1	20,0	10,0	45	1,0	21,5	9,8
10	1,0	19,5	10,0	46	1,2	21,3	10,2
11	1,1	21,5	10,0	47	1,0	22,1	10,0
12	1,2	23,0	10,0	48	1,1	21,4	10,6
13	1,2	22,5	10,0	49	1,0	21,4	9,9
14	1,2	21,5	10,4	50	1,1	20,8	10,6
15	1,0	20,5	9,8	51	1,1	22,0	10,1
16	1,2	22,5	10,0	52	1,1	22,1	10,0
17	1,2	21,0	10,0	53	1,0	21,5	9,8
18	1,2	21,0	10,0	54	1,0	21,0	10,3
19	1,0	22,0	9,5	55	1,0	20,2	10,0
20	1,1	19,0	10,3	56	1,0	19,0	10,0
21	1,0	20,6	9,8	57	1,2	21,0	10,0
22	1,2	21,5	9,9	58	1,3	24,8	10,3
23	1,1	22,0	9,7	59	1,1	21,5	10,5
24	1,0	20,4	10,1	60	1,2	23,0	10,0
25	1,2	23,5	10,0	61	1,1	20,8	10,0
26	1,0	19,0	10,0	62	1,2	20,2	10,2
27	1,2	20,5	10,2	63	1,0	21,0	9,8
28	0,9	20,0	9,5	64	1,1	21,2	10,0
29	1,2	22,5	10,3	65	0,8	19,2	9,6
30	0,9	18,3	9,9	66	1,1	21,1	10,1
31	1,1	22,0	10,0	67	1,0	20,4	10,1
32	0,9	20,2	9,6	68	1,1	21,0	10,2
33	1,2	19,5	10,0	69	1,0	25,5	10,3
34	1,3	22,5	10,3	70	1,3	25,2	10,3
35	1,0	18,2	10,3	71	1,3	25,5	10,1
36	1,0	20,3	9,9	72	1,2	22,0	10,1

\* Из материалов НИС МЭСИ.

№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм	№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм
73	1,1	24,0	9,5	115	1,1	22,0	10,2
74	1,0	20,0	10,0	116	1,0	20,5	10,6
75	1,1	21,5	10,0	117	1,2	22,4	10,3
76	1,1	23,5	10,0	118	1,1	20,0	10,2
77	1,0	20,0	9,9	119	1,0	19,7	10,1
78	1,1	20,5	10,0	120	1,0	19,6	9,8
79	1,2	20,0	10,0	121	1,0	20,1	10,0
80	1,1	19,6	9,9	122	1,0	18,0	10,3
81	0,9	22,5	9,0	123	1,1	21,3	10,0
82	1,1	21,0	10,0	124	1,0	21,5	9,8
83	1,1	19,5	10,3	125	1,0	21,0	9,9
84	1,1	22,5	10,0	126	1,1	23,5	9,8
85	1,0	20,0	10,0	127	1,1	23,0	10,0
86	1,0	20,5	9,6	128	1,1	20,5	10,4
87	1,1	21,0	10,2	129	1,1	19,7	10,2
88	1,2	21,5	10,2	130	1,3	25,0	10,2
89	1,0	19,5	9,8	131	1,1	25,0	10,0
90	1,0	21,0	9,8	132	1,0	20,5	10,0
91	1,3	26,0	10,0	133	1,0	19,5	10,3
92	1,0	21,0	9,8	134	1,1	22,3	10,0
93	1,0	20,0	9,6	135	1,2	21,0	10,0
94	1,3	22,3	10,1	136	1,1	23,5	9,8
95	1,2	22,0	9,8	137	1,0	19,6	10,0
96	1,0	20,5	10,0	138	1,0	21,9	9,7
97	1,1	22,6	10,0	139	1,1	24,0	10,0
98	1,0	20,7	10,0	140	1,0	19,8	10,0
99	1,0	22,2	9,9	141	1,0	23,0	10,0
100	1,0	19,5	10,2	142	1,2	23,5	10,0
101	1,1	23,0	9,7	143	1,2	21,5	10,3
102	1,1	20,9	10,0	144	1,2	24,0	10,5
103	1,0	22,7	9,5	145	0,9	22,0	9,3
104	1,2	23,6	10,3	146	1,1	21,3	10,0
105	1,1	20,6	10,5	147	1,2	22,7	10,0
106	1,1	23,7	9,7	148	1,1	21,0	10,1
107	1,0	22,1	10,0	149	1,1	21,3	10,0
108	1,1	21,5	10,3	150	1,1	21,5	9,8
109	1,2	22,6	10,5	151	1,0	20,5	10,0
110	1,0	21,0	10,0	152	1,0	24,0	9,8
111	1,1	19,5	10,1	153	1,0	23,2	9,5
112	1,0	20,6	10,0	154	1,0	20,1	9,6
113	1,3	23,5	10,4	155	1,1	20,7	10,2
114	1,2	20,0	10,2	156	1,0	21,5	9,8

№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм	№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм
157	1,1	23,0	9,7	199	1,2	21,7	10,6
158	1,0	22,7	9,6	200	1,2	23,0	10,0
159	1,2	21,0	10,0	201	1,0	20,0	10,5
160	1,2	23,7	9,9	202	1,0	19,0	10,0
161	1,1	21,6	10,3	203	1,2	23,7	9,8
162	0,8	21,0	9,8	204	1,2	24,5	10,0
163	1,0	19,2	10,0	205	1,1	21,5	10,0
164	1,1	18,6	10,4	206	0,9	22,0	9,0
165	1,1	23,0	9,6	207	1,1	23,5	10,0
166	1,1	20,7	10,0	208	1,1	21,0	10,2
167	1,2	22,7	10,3	209	1,0	19,7	10,1
168	1,2	21,0	10,0	210	1,2	21,5	10,0
169	1,2	22,4	10,0	211	0,9	21,0	9,5
170	1,1	22,1	9,9	212	1,0	23,6	9,8
171	1,0	21,6	9,8	213	1,0	21,0	9,8
172	1,0	20,4	10,0	214	1,0	20,3	10,0
173	1,3	23,5	10,7	215	1,0	21,3	10,0
174	1,2	23,7	10,3	216	1,2	23,5	10,3
175	1,1	25,0	9,8	217	1,1	22,2	10,0
176	1,1	23,0	10,0	218	1,0	20,0	9,8
177	1,2	24,0	10,0	219	0,9	20,5	9,5
178	1,0	21,8	9,8	220	1,2	24,5	10,3
179	1,2	22,0	10,0	221	1,2	24,2	10,2
180	1,0	19,5	10,5	222	1,1	22,0	10,0
181	1,0	22,5	10,0	223	1,0	19,5	10,2
182	1,1	23,8	9,7	224	1,2	24,0	9,8
183	1,2	24,0	9,8	225	1,0	23,5	9,8
184	1,2	23,2	10,0	226	1,0	21,4	10,0
185	1,2	24,5	10,0	227	1,2	20,5	10,2
186	1,0	23,2	9,5	228	1,0	20,4	10,0
187	1,2	23,5	10,0	229	1,0	22,6	9,5
188	1,1	22,3	10,0	230	1,0	21,0	9,8
189	1,1	21,5	10,0	231	1,3	25,0	10,3
190	1,2	23,0	10,0	232	1,2	21,5	10,0
191	1,1	22,0	9,8	233	1,2	23,3	9,9
192	1,0	22,0	9,7	234	1,2	20,5	10,2
193	1,2	25,3	10,1	235	1,0	20,7	10,3
194	1,2	23,5	10,0	236	1,1	22,5	10,0
195	1,0	23,5	9,5	237	1,0	22,7	9,8
196	1,1	23,3	9,8	238	1,1	23,0	9,7
197	1,0	20,4	10,0	239	1,0	20,0	10,0
198	1,2	23,5	10,0	240	1,1	23,3	9,9

№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм	№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм
241	1,0	22,3	9,6	283	1,1	21,0	10,4
242	1,1	20,5	10,0	284	1,0	23,4	9,5
243	1,2	23,3	10,1	285	1,2	24,0	10,2
244	1,0	20,5	9,9	286	1,1	22,0	9,7
245	1,1	23,5	9,7	287	1,2	21,5	10,0
246	1,1	20,5	10,3	288	1,1	20,4	9,8
247	1,2	23,5	10,5	289	1,0	22,5	9,8
248	1,2	24,5	10,0	290	1,0	18,5	9,9
249	1,2	20,1	10,0	291	1,1	19,6	9,9
250	1,2	23,7	10,1	292	1,0	21,5	10,0
251	1,3	25,0	10,3	293	1,0	22,0	9,5
252	1,2	23,3	10,0	294	1,2	23,7	10,0
253	1,2	23,5	10,3	295	1,2	22,3	10,0
254	1,1	20,8	10,3	296	1,0	22,3	9,9
255	1,0	20,0	10,0	297	1,1	20,5	10,0
256	1,2	24,6	10,0	298	1,0	23,0	9,8
257	1,2	21,5	10,3	299	1,2	19,0	9,8
258	1,0	19,5	10,0	300	1,2	23,0	10,2
259	0,8	20,5	9,0	301	1,1	22,5	10,3
260	1,0	19,0	10,2	302	1,0	21,0	10,1
261	0,9	19,1	10,0	303	1,0	23,0	9,7
262	1,1	20,0	10,3	304	1,1	19,0	10,5
263	1,0	21,0	9,8	305	1,0	20,0	10,2
264	1,2	22,0	10,0	306	1,0	23,0	9,8
265	1,1	22,5	9,9	307	1,0	19,0	10,0
266	1,2	25,0	10,0	308	1,2	23,0	10,0
267	1,0	23,2	9,5	309	1,0	21,0	10,0
268	1,2	21,5	10,0	310	1,0	20,5	9,8
269	1,0	19,0	9,9	311	1,1	22,0	9,8
270	1,3	24,0	10,3	312	1,1	21,5	10,0
271	1,1	20,5	10,0	313	1,0	21,8	10,0
272	1,2	21,7	10,0	314	1,2	23,0	10,0
273	1,1	20,0	10,5	315	1,2	21,5	10,3
274	1,1	22,0	10,0	316	1,1	20,5	10,0
275	0,9	20,2	9,7	317	1,0	22,5	10,0
276	1,2	25,0	10,0	318	1,1	21,0	10,0
277	1,2	22,5	10,0	319	1,2	23,5	10,0
278	1,2	23,4	10,0	320	1,1	23,0	10,0
279	1,2	23,8	10,0	321	1,0	20,0	9,8
280	1,0	20,6	10,0	322	1,0	20,0	10,0
281	1,1	20,5	10,0	323	1,2	22,5	10,0
282	1,2	22,5	10,0	324	1,0	21,0	10,0

№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм	№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм
325	1,1	20,3	10,2	367	1,1	22,0	10,0
326	1,1	23,0	10,0	368	1,1	22,0	9,8
327	1,0	22,5	9,7	369	1,0	20,5	9,9
328	1,1	22,0	10,2	370	1,0	19,0	10,0
329	1,0	21,0	10,0	371	1,0	23,0	9,6
330	1,2	21,0	10,0	372	1,0	22,0	10,0
331	0,9	23,0	9,0	373	1,0	21,5	10,0
332	1,1	21,2	10,0	374	1,1	21,0	10,0
333	1,0	20,0	9,8	375	1,0	20,0	10,0
334	1,0	23,0	9,5	376	0,8	23,0	9,7
335	1,0	22,5	9,6	377	1,0	20,5	10,0
336	1,0	19,7	10,0	378	1,1	23,5	10,0
337	1,1	20,4	10,0	379	1,2	23,0	10,4
338	1,0	22,5	9,8	380	1,0	19,5	10,3
339	1,2	23,0	10,0	381	1,1	20,0	10,3
340	1,1	21,5	10,0	382	1,0	19,3	10,0
341	1,0	20,0	10,0	383	1,0	22,3	10,0
342	1,1	23,0	10,0	384	0,9	20,5	9,6
343	1,5	24,0	10,6	385	1,2	23,7	10,2
344	1,1	19,8	10,5	386	1,0	20,6	9,9
345	1,0	18,5	9,9	387	1,2	21,5	10,2
346	1,0	22,8	9,8	388	1,2	20,0	10,0
347	1,2	21,0	10,0	389	1,0	22,0	9,8
348	1,2	22,5	10,0	390	1,1	20,5	10,3
349	1,0	22,0	9,8	391	1,0	20,7	10,0
350	1,0	22,5	9,5	392	1,1	23,2	9,9
351	1,1	23,0	10,0	393	1,2	23,0	9,9
352	1,1	23,0	9,8	394	1,2	23,5	10,3
353	0,9	20,2	9,9	395	1,0	20,5	9,8
354	0,9	20,0	9,8	396	1,1	21,5	10,2
355	1,0	20,0	10,0	397	1,0	20,4	10,0
356	1,2	23,0	10,3	398	1,0	21,0	10,1
357	1,0	20,5	9,8	399	1,0	20,0	10,5
358	1,1	20,0	10,4	400	1,0	20,5	10,0
359	1,1	21,0	10,0	401	1,0	23,0	9,5
360	1,1	21,0	10,0	402	1,1	21,0	10,5
361	1,2	25,0	10,0	403	1,1	22,0	10,0
362	1,0	21,0	10,1	404	1,2	23,0	10,3
363	1,0	21,5	9,8	405	1,2	21,8	10,0
364	1,2	23,5	10,0	406	1,0	19,4	10,2
365	1,1	19,6	10,3	407	1,0	21,5	9,9
366	1,0	20,0	10,0	408	1,2	23,0	10,0



## Продолжение

№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм	№ п/п	Объем в см <sup>3</sup>	Длина в мм	Большой диаметр в мм
409	1,2	22,5	10,0	419	1,1	21,7	9,8
410	1,2	24,0	10,0	420	1,1	24,0	9,5
411	1,2	23,6	10,4	421	1,1	22,0	10,0
412	1,0	21,5	9,8	422	1,2	24,0	10,0
413	1,1	21,5	10,0	423	1,3	24,4	10,0
414	1,1	19,7	10,0	424	1,1	19,6	10,3
415	1,2	23,5	10,3	425	1,0	21,2	10,8
416	1,2	21,0	9,8	426	1,2	23,0	10,3
417	1,1	22,4	10,0	427	1,0	21,5	10,0
418	1,0	20,0	9,8	428	1,2	22,0	10,0

Таблица 2

## РЕЗУЛЬТАТЫ ИЗМЕРЕНИЯ НАРУЖНОГО ДИАМЕТРА ПЛАШЕК

Номинал — 37,2

Технический допуск — 0,1

№ п/п	Размеры	№ п/п	Размеры	№ п/п	Размеры	№ п/п	Размеры	№ п/п	Размеры
1	37,95	41	37,92	81	37,81	121	37,81	161	37,82
2	37,80	42	37,94	82	37,85	122	37,86	162	37,87
3	37,95	43	37,90	83	37,90	123	37,87	163	37,89
4	37,97	44	37,87	84	37,86	124	37,83	164	37,90
5	37,93	45	37,89	85	37,85	125	37,86	165	37,85
6	37,00	46	37,93	86	37,83	126	37,88	166	37,88
7	37,97	47	37,91	87	37,84	127	37,89	167	37,87
8	37,97	48	37,91	88	37,85	128	37,86	168	37,85
9	37,95	49	37,95	89	37,85	129	37,86	169	37,89
10	37,95	50	37,85	90	37,83	130	37,89	170	37,87
11	37,93	51	37,95	91	37,87	131	37,89	171	37,86
12	37,94	52	37,91	92	37,86	132	37,90	172	37,82
13	37,93	53	37,92	93	37,88	133	37,90	173	37,84
14	37,94	54	37,93	94	37,84	134	37,82	174	37,87
15	37,94	55	37,91	95	37,83	135	37,82	175	37,81
16	37,96	56	37,89	96	37,80	136	37,84	176	37,80
17	37,95	57	37,82	97	37,85	137	37,89	177	37,78
18	37,98	58	37,92	98	37,88	138	37,88	178	37,82
19	37,99	59	37,92	99	37,82	139	37,86	179	37,83
20	37,95	60	37,92	100	37,78	140	37,88	180	37,82
21	37,93	61	37,92	101	37,81	141	37,85	181	37,80
22	37,95	62	37,93	102	37,83	142	37,87	182	37,87
23	37,91	63	37,87	103	37,82	143	37,90	183	37,80
24	37,92	64	37,86	104	37,81	144	37,80	184	37,78
25	37,94	65	37,86	105	37,85	145	37,80	185	37,83
26	37,94	66	37,93	106	37,84	146	37,87	186	37,83
27	37,90	67	37,84	107	37,84	147	37,75	187	37,85
28	37,92	68	37,90	108	37,85	148	37,91	188	37,82
29	37,93	69	37,80	109	37,88	149	37,90	189	37,83
30	37,92	70	37,85	110	37,85	150	37,82	190	37,81
31	37,91	71	37,89	111	37,89	151	37,80	191	37,87
32	37,96	72	37,90	112	37,91	152	37,87	192	37,88
33	37,93	73	37,94	113	37,81	153	37,88	193	37,87
34	37,94	74	37,92	114	37,85	154	37,82	194	37,84
35	37,92	75	37,86	115	37,87	155	37,87	195	37,85
36	37,93	76	37,84	116	37,88	156	37,91	196	37,85
37	37,90	77	37,84	117	37,84	157	37,90	197	37,82
38	37,94	78	37,85	118	37,86	158	37,87	198	37,84
39	37,94	79	37,82	119	37,76	159	37,82	199	37,90
40	37,93	80	37,82	120	37,87	160	37,88	200	37,87

Таблица 3

## РЕЗУЛЬТАТЫ ИЗМЕРЕНИЯ ТОЛЩИНЫ ПЛАШЕК ПОСЛЕ ШЛИФОВКИ

Номинал — 14,20

Технический допуск — 0,05

№ п/п	Замеры	№ п/п	Замеры	№ п/п	Замеры	№ п/п	Замеры	№ п/п	Замеры
1	14,13	41	14,13	81	14,14	121	14,17	161	14,19
2	14,13	42	14,12	82	14,20	122	14,16	162	14,18
3	14,18	43	14,7	83	14,17	123	14,21	163	14,20
4	14,13	44	14,13	84	14,15	124	14,13	164	14,20
5	14,20	45	14,21	85	14,15	125	14,19	165	14,17
6	14,20	46	14,12	86	14,19	126	14,16	166	14,13
7	14,14	47	14,13	87	14,13	127	14,14	167	14,12
8	14,21	48	14,18	88	14,15	128	14,13	168	14,19
9	14,12	49	14,14	89	14,15	129	14,23	169	14,18
10	14,18	50	14,11	90	14,19	130	14,14	170	14,18
11	14,19	51	14,13	91	14,19	131	14,16	171	14,19
12	14,17	52	14,13	92	14,18	132	14,15	172	14,14
13	14,17	53	14,17	93	14,20	133	14,14	173	14,15
14	14,18	54	14,21	94	14,18	134	14,17	174	14,15
15	14,18	55	14,19	95	14,15	135	14,18	175	14,21
16	14,17	56	14,14	96	14,20	136	14,21	176	14,16
17	14,15	57	14,14	97	14,18	137	14,17	177	14,20
18	14,20	58	14,22	98	14,15	138	14,17	178	14,18
19	14,21	59	14,13	99	14,15	139	14,15	179	14,18
20	14,18	60	14,12	100	14,17	140	14,12	180	14,14
21	14,12	61	14,21	101	14,15	141	14,15	181	14,17
22	14,18	62	14,15	102	14,20	142	14,14	182	14,16
23	14,20	63	14,22	103	14,15	143	14,20	183	14,17
24	14,16	64	14,21	104	14,16	144	14,17	184	14,20
25	14,22	65	14,13	105	14,13	145	14,18	185	14,18
26	14,14	66	14,17	106	14,14	146	14,18	186	14,19
27	14,20	67	14,12	107	14,14	147	14,19	187	14,18
28	14,19	68	14,16	108	14,13	148	14,18	188	14,18
29	14,13	69	14,12	109	14,15	149	14,20	189	14,18
30	14,13	70	14,13	110	14,15	150	14,20	190	14,17
31	14,13	71	14,12	111	14,13	151	14,18	191	14,18
32	14,20	72	14,12	112	14,14	152	14,18	192	14,20
33	14,18	73	14,21	113	14,20	153	14,20	193	14,18
34	14,15	74	14,13	114	14,13	154	14,18	194	14,19
35	14,17	75	14,15	115	14,18	155	14,17	195	14,14
36	14,21	76	14,21	116	14,13	156	14,15	196	14,16
37	14,13	77	14,13	117	14,16	157	14,13	197	14,19
38	14,22	78	14,13	118	14,19	158	14,14	198	14,18
39	14,35	79	14,12	119	14,13	159	14,17	199	14,18
40	14,13	80	14,20	120	14,12	160	14,13	200	14,13

ТАБЛИЦА ЗНАЧЕНИЙ ФУНКЦИИ  $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

1	0	1	2	3	4	5	6	7	8	9
0,0	3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0855	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0203	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
4,0	0001	0001	0001	0000	0000	0000	0000	0000	0000	0000

Все значения умножены на 10000

ТАБЛИЦА ЗНАЧЕНИЙ ФУНКЦИИ

$$F(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt \quad \text{при разных значениях } t.$$

<i>t</i>	<i>F(t)</i>	<i>t</i>	<i>F(t)</i>	<i>t</i>	<i>F(t)</i>	<i>t</i>	<i>F(t)</i>
0,00	0,00000	35	27366	0,70	0,51607	1,05	70628
01	00798	36	28115	71	52230	06	71086
02	01596	37	28862	72	52848	07	71538
03	02393	38	29605	73	53461	08	71986
04	03191	39	30346	74	54070	09	72429
05	03988	0,40	0,31084	75	54675	1,10	0,72867
06	04784	41	31819	76	55275	11	73300
07	05581	42	32552	77	55870	12	73729
08	06376	43	33280	78	56461	13	74152
09	07171	44	34006	79	57047	14	74571
0,10	0,07966	45	34729	0,80	0,57629	15	74986
11	08759	46	35448	81	58206	16	75395
12	09552	47	36164	82	58778	17	75800
13	10348	48	36877	83	59346	18	76200
14	11134	49	37587	84	59909	19	76595
15	11924	0,50	0,38292	85	60468	1,20	0,76986
16	12712	51	38995	86	61021	21	77372
17	13499	52	39694	87	61570	22	77754
18	14285	53	40389	88	62114	23	78130
19	15069	54	41080	89	62653	24	78502
0,20	0,15852	55	41768	0,90	0,63188	25	78870
21	16633	56	42452	91	63718	26	79233
22	17413	57	43132	92	64243	27	79592
23	18191	58	43809	93	64763	28	79945
24	18967	59	44481	94	65278	29	80295
25	19741	0,60	0,45149	95	65789	1,30	0,80640
26	20514	61	45814	96	66294	31	80980
27	21284	62	46474	97	66795	32	81316
28	22052	63	47131	98	67291	33	81648
29	22818	64	47783	99	67783	34	81975
0,30	0,23582	65	48431	1,00	0,68269	35	82298
31	24344	66	49075	01	68750	36	82617
32	25103	67	49714	02	69227	37	82931
33	25860	68	50350	03	69699	38	83241
34	26614	69	50981	04	70166	39	83547

$t$	$F(t)$	$t$	$F(t)$	$t$	$F(t)$	$t$	$F(t)$
1,40	0,83849	1,80	0,92814	2,20	0,97219	2,60	0,99068
41	84146	81	92970	21	97289	61	99095
42	84439	82	93124	22	97358	62	99121
43	84728	83	93275	23	97425	63	99146
44	85013	84	93423	24	97491	64	99171
45	85294	85	93569	25	97555	65	99195
46	85571	86	93711	26	97618	66	99219
47	85844	87	93852	27	97679	67	99241
48	86113	88	93989	28	97739	68	99263
49	86378	89	94124	29	97798	69	99285
1,50	0,86639	1,90	0,94257	2,30	0,97855	2,70	0,99307
51	86696	91	94387	31	97911	71	99327
52	87149	92	94514	32	97966	72	99347
53	87398	93	94639	33	98019	73	99367
54	87644	94	94762	34	98072	74	99386
55	87886	95	94882	35	98123	75	99404
56	88124	96	95000	36	98172	76	99422
57	88358	97	95116	37	98221	77	99439
58	88589	98	95230	38	98269	78	99456
59	88817	99	95341	39	98315	79	99473
1,60	0,89040	2,00	0,95450	2,40	0,98360	2,80	0,99489
61	89260	01	95557	41	98405	81	99505
62	89477	02	95662	42	98448	82	99520
63	89690	03	95764	43	98490	83	99535
64	89899	04	95865	44	98531	84	99549
65	90106	05	95964	45	98571	85	99563
66	90309	06	96060	46	98611	86	99576
67	90508	07	96155	47	98649	87	99590
68	90704	08	96247	48	98686	88	99602
69	90897	09	96338	49	98723	89	99615
1,70	0,91087	2,10	0,96427	2,50	0,98758	2,90	0,99627
71	91273	11	96514	51	98793	91	99639
72	91457	12	96599	52	98826	92	99650
73	91637	13	96683	53	98859	93	99661
74	91814	14	96765	54	98891	94	99672
75	91988	15	96844	55	98923	95	99682
76	92159	16	96923	56	98953	96	99692
77	92327	17	96999	57	98983	97	99702
78	92492	18	97074	58	99012	98	99712
79	92655	19	97148	59	99040	99	99721

$t$	$F(t)$	$t$	$F(t)$	$t$	$F(t)$	$t$	$F(t)$
3,00	0,99730	25	99885	3,50	0,99953	75	99982
01	99739	26	99889	51	99955	76	99983
02	99747	27	99892	52	99957	77	99984
03	99755	28	99896	53	99958	78	99984
04	99763	29	99900	54	99960	79	99985
05	99771	3,30	0,99903	55	99961	3,80	0,99986
06	99779	31	99907	56	99963	81	99986
07	99786	32	99910	57	99964	82	99987
08	99793	33	99913	58	99966	83	99987
09	99800	34	99916	59	99967	84	99988
3,10	0,99806	35	99919	3,60	0,99968	85	99988
11	99813	36	99922	61	99969	86	99989
12	99819	37	99925	62	99971	87	99989
13	99825	38	99928	63	99972	88	99990
14	99831	39	99930	64	99973	89	99990
15	99837	3,40	0,99933	65	99974	3,90	0,99990
16	99842	41	99935	66	99975	91	99991
17	99848	42	99937	67	99976	92	99991
18	99853	43	99940	68	99977	93	99992
19	99858	44	99942	69	99978	94	99992
3,20	0,99863	45	99944	3,70	0,99978	95	99992
21	99867	46	99946	71	99979	96	99992
22	99872	47	99948	72	99980	97	99993
23	99876	48	99950	73	99981	98	99993
24	99880	49	99952	74	99982	99	99993

$S(t)$  в распределении Стюдента

$n \backslash t$	1	2	3	4	5	6—7	8—10	11— —15	16— —24	25— —35	$\infty$
0,0	500	500	500	500	500	500	500	500	500	500	500.000.0
0,1	532	535	537	537	538	538	539	539	539	539	539.827.8
0,2	563	570	573	574	576	576	578	578	578	578	579.259.7
0,3	593	606	608	610	612	613	615	616	616	616	617.911.4
0,4	621	636	642	645	647	649	651	652	653	654	655.421.7
0,5	648	667	674	678	681	683	685	687	689	689	691.462.5
0,6	672	695	705	710	713	715	718	721	722	724	725.746.9
0,7	694	723	733	739	742	746	749	752	754	756	758.036.3
0,8	715	746	759	766	770	774	778	781	783	785	788.144.6
0,9	733	768	783	790	795	800	801	808	811	813	815.939.9
1,0	750	789	804	813	818	823	823	832	835	838	841.344.7
1,1	765	807	842	834	839	844	850	854	858	860	864.333.9
1,2	779	824	842	852	858	864	870	874	878	881	884.930.3
1,3	791	838	858	868	875	881	887	892	896	899	903.199.5
1,4	803	852	872	883	890	896	902	907	912	915	919.243.3
1,5	813	864	885	896	903	909	916	921	925	928	933.192.8
1,6	822	875	896	908	915	921	928	933	937	940	945.200.7
1,7	831	884	906	918	925	932	938	943	948	951	955.434.5
1,8	839	893	915	927	934	941	947	952	956	959	964.069.7
1,9	846	901	923	935	942	948	955	960	964	967	971.283.4
2,0	852	908	930	942	949	955	962	967	970	973	977.249.9
2,1	858	915	937	948	955	961	967	972	976	978	982.135.6
2,2	864	921	942	954	960	966	972	977	980	982	986.096.6
2,3	870	926	948	958	965	971	977	981	984	986	989.275.9
2,4	874	931	952	963	969	975	980	984	987	989	991.802.5
2,5	879	935	956	976	973	978	983	987	989	991	993.790.3
2,6	883	939	960	970	976	981	986	989	991	993	995.338.8
2,7	887	943	963	973	979	983	988	991	993	995	996.533.8
2,8	891	946	966	976	981	985	990	993	995	996	997.444.9



## Продолжение

$\begin{matrix} n \\ t \end{matrix}$	1	2	3	4	5	6—7	8—10	11— —15	16— —24	25— —35	$\infty$
2,9	894	949	969	978	983	987	991	994	996	997	998.134.2
3,0	898	952	971	980	985	989	993	995	997	997	998.650.1
3,1	901	955	973	982	987	990	994	996	997	998	999.032.4
3,2	904	967	975	984	988	991	995	997	998	998	999.312.9
3,3	906	960	977	985	989	992	995	997	998	999	999.516.6
3,4	909	962	979	986	990	993	996	998	998	999	999.663.1
3,5	911	964	980	988	991	993	997	998	999	—	999.767.4
3,6	914	965	982	989	992	994	997	998	—	—	999.840.9
3,7	916	967	983	990	993	995	998	999	—	—	999.892.2
3,8	918	969	984	991	994	996	998	999	—	—	999.927.7
3,9	920	970	985	992	994	996	998	999	—	—	999.951.9
4,0	922	971	986	993	995	997	998	—	—	—	999.968.3
4,1	924	973	987	993	995	997	999	—	—	—	999.979.3
4,2	926	974	988	993	996	998	999	—	—	—	999.986.7
4,3	927	975	988	994	996	998	999	—	—	—	999.991.5
4,4	929	976	989	994	996	998	—	—	—	—	999.994.6
4,5	930	977	989	995	997	998	—	—	—	—	999.996.6
4,6	932	978	990	995	997	998	—	—	—	—	999.997.9
4,7	933	979	991	995	997	999	—	—	—	—	999.998.7
4,8	935	980	991	996	998	999	—	—	—	—	999.999.2
4,9	936	980	992	996	998	999	—	—	—	—	999.999.5
5,0	937	981	992	996	998	999	—	—	—	—	999.999.7
5,1	938	982	993	996	998	—	—	—	—	—	999.999.8
5,2	940	982	993	997	998	—	—	—	—	—	999.999.9
5,3	941	983	993	997	998	—	—	—	—	—	999.999.9
5,4	942	984	994	997	998	—	—	—	—	—	—
5,5	943	984	994	997	999	—	—	—	—	—	—
5,6	943	984	994	997	999	—	—	—	—	—	—
5,7	945	985	995	998	999	—	—	—	—	—	—
5,8	946	986	995	998	999	—	—	—	—	—	—
5,9	947	986	995	998	999	—	—	—	—	—	—
6,0	947	987	995	998	—	—	—	—	—	—	—
6,1	948	987	996	999	—	—	—	—	—	—	—

ТАБЛИЦА ЗНАЧЕНИЙ  $z = \frac{1}{2} \ln \frac{1+r}{1-r}$

r	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0100	0,0200	0,0300	0,0400	0,0500	0,0601	0,0701	0,0802	0,0902
0,1	0,1003	0,1104	0,1206	0,1307	0,1409	0,1511	0,1614	0,1717	0,1820	0,1923
0,2	0,2027	0,2132	0,2237	0,2342	0,2448	0,2554	0,2661	0,2769	0,2877	0,2986
0,3	0,3095	0,3205	0,3316	0,3428	0,3541	0,3654	0,3769	0,3884	0,4001	0,4118
0,4	0,4236	0,4356	0,4477	0,4599	0,4722	0,4847	0,4973	0,5101	0,5230	0,5361
0,5	0,5493	0,5627	0,5763	0,5901	0,6042	0,6184	0,6328	0,6475	0,6625	0,6777
0,6	0,6931	0,7089	0,7250	0,7414	0,7582	0,7753	0,7928	0,8107	0,8291	0,8480
0,7	0,8673	0,8872	0,9076	0,9287	0,9505	0,9730	0,9962	1,0203	1,0454	1,0714
0,8	1,0986	1,1270	1,1568	1,1881	1,2212	1,2562	1,2933	1,3331	1,3758	1,4219
0,9	1,4722	1,5275	1,5890	1,6584	1,7380	1,8318	1,9459	2,0923	2,2976	2,6466
0,99	2,6466	2,6996	2,7587	2,8257	2,9031	2,9945	3,1063	3,2504	3,4534	3,8002

**ЗНАЧЕНИЕ  $F$  ПРИ ВЕРОЯТНОСТЯХ 0,95 (верхняя строка) и  
0,99 (нижняя строка)**

Степени свободы вариации для большей дисперсии

	1	3	5	7	9	11	14	20	30	50
1	161 4052	216 5403	230 5764	237 5928	241 6022	243 6082	245 6142	248 6208	250 6258	252 6302
3	10,13 34,12	9,28 29,46	9,01 28,24	8,88 27,67	8,81 27,34	8,76 27,13	8,71 26,92	8,66 26,69	8,62 26,50	8,58 26,35
5	6,61 16,26	5,41 12,06	5,05 10,97	4,88 10,45	4,78 10,15	4,70 9,96	4,64 9,77	4,56 9,55	4,50 9,38	4,44 9,24
7	5,59 12,25	4,35 8,45	3,97 7,46	3,79 7,00	3,68 6,71	3,60 6,54	3,52 6,35	3,44 6,15	3,38 5,98	3,32 5,85
9	5,12 10,56	3,86 6,99	3,48 6,06	3,29 5,62	3,18 5,35	3,10 5,18	3,02 5,00	2,93 4,80	2,86 4,64	2,80 4,51
11	4,84 9,85	3,59 6,22	3,20 5,32	3,01 4,88	2,90 4,63	2,82 4,46	2,74 4,29	2,65 4,10	2,57 3,94	2,50 3,80
13	4,67 9,07	3,41 5,74	3,02 4,86	2,84 4,44	2,72 4,19	2,63 4,02	2,55 3,85	2,46 3,67	2,38 3,51	2,32 3,37
15	4,54 8,68	3,29 5,42	2,90 4,56	2,70 4,14	2,59 3,89	2,51 3,73	2,43 3,56	2,33 3,36	2,25 3,20	2,18 3,07
17	4,45 8,40	3,20 5,18	2,81 4,34	2,62 3,93	2,50 3,68	2,41 3,52	2,33 3,35	2,23 3,16	2,15 3,00	2,08 2,86
19	4,38 8,18	3,13 5,01	2,74 4,17	2,55 3,77	2,43 3,52	2,34 3,36	2,26 3,19	2,15 3,00	2,07 2,84	2,00 2,70
21	4,32 8,02	3,07 4,87	2,68 4,04	2,49 3,65	2,37 3,40	2,28 3,24	2,20 3,07	2,09 2,88	2,00 2,72	1,93 2,58
23	4,28 7,88	3,03 4,76	2,64 3,94	2,45 3,54	2,32 3,30	2,24 3,14	2,14 2,97	2,04 2,78	1,96 2,62	1,88 2,48
25	4,24 7,77	2,99 4,68	2,60 3,86	2,41 3,46	2,28 3,21	2,20 3,05	2,11 2,89	2,00 2,70	1,92 2,54	1,84 2,40
27	4,21 7,68	2,96 4,60	2,57 3,79	2,37 3,39	2,25 3,14	2,16 2,98	2,08 2,83	1,97 2,63	1,88 2,45	1,80 2,33
29	4,18 7,60	2,93 4,54	2,54 3,73	2,35 3,33	2,22 3,08	2,14 2,92	2,05 1,77	1,94 2,57	1,85 2,41	1,77 2,27

$(x')^{1..m}$  или  $(x'')^{1..m}$  для вычисления моментов

$x' \text{ или } x''$ $m$	3	4	5	6	7	8	9	11	12
1	81	256	625	1296	2401	4096	6561	14641	20736
2	162	512	1250	2592	4802	8192	13122	29282	41472
3	243	768	1875	3888	7203	12288	19683	43923	62208
4	324	1024	2500	5184	9604	16384	26244	58564	82944
5	405	1280	3125	6840	12005	20480	32805	73205	103680
6	486	1536	3750	7776	14406	24576	39366	87846	124416
7	567	1792	4375	9072	16807	28672	45927	102487	145152
8	648	2048	5000	10368	19208	32768	52488	117128	165888
9	729	2304	5625	11664	21609	36864	59049	131769	186624
10	810	2560	6250	12960	24010	40960	65610	146410	207360
11	891	2816	6875	14256	26411	45056	72171	161051	228096
12	972	3072	7500	15552	28812	49152	76732	175692	248832
13	1053	3328	8125	16848	31213	53248	85293	190333	269568
14	1134	3584	8750	18144	33614	57344	91854	204974	290304
15	1215	3840	9375	19440	36015	61440	98415	219615	311040
16	1296	4096	10000	20736	38416	65536	104976	234256	331776
17	1377	4352	10625	22032	40817	69632	111537	248897	352512
18	1458	4608	11250	23328	43218	73728	118098	263538	372248
19	1539	4864	11875	24624	45619	77824	124659	278179	393984
20	1620	5120	12500	25920	48020	81920	131220	292820	414720
21	1701	5376	13125	27216	50421	86016	137781	307461	435456
22	1782	5632	13750	28512	52822	90112	144342	322102	456192
23	1863	5888	14375	29808	55223	94208	150903	336743	476928
24	1944	6144	15000	31104	57624	98304	157464	351384	497664
25	2025	6400	15625	32400	60025	102400	164025	366025	518400

$x'$ или $x''$		Продолжение									
$m$		3	4	5	6	7	8	9	11	12	
26	2 106	6 656	16 250	33 696	62 426	106 496	170 586	380 666	539 136		
27	2 187	6 912	16 875	34 992	64 827	110 592	177 147	395 307	559 872		
28	2 268	7 168	17 500	36 288	67 228	114 688	183 708	409 948	580 608		
29	2 349	7 424	18 125	37 584	69 629	118 784	190 269	424 589	601 344		
30	2 430	7 680	18 750	38 880	72 030	122 280	196 830	439 230	622 080		
31	2 511	7 936	19 375	40 176	74 431	126 976	203 391	453 871	642 816		
32	2 592	8 192	20 000	41 472	76 832	131 072	209 952	468 512	663 552		
33	2 673	8 448	20 625	42 768	79 233	135 168	216 513	483 153	684 288		
34	2 754	8 704	21 250	44 064	81 634	139 264	223 074	497 794	705 024		
35	2 835	8 960	21 875	45 360	84 035	143 360	229 635	512 435	725 760		
36	2 916	9 216	22 500	46 656	86 436	147 456	236 196	527 076	746 496		
37	2 997	9 472	23 125	47 952	88 837	151 552	242 757	541 717	767 232		
38	3 078	9 728	23 750	49 248	91 238	155 648	249 318	556 358	787 968		
39	3 159	9 984	24 375	50 544	93 639	157 744	255 879	570 999	808 704		
40	3 240	10 240	25 000	51 840	96 040	163 840	262 440	585 640	829 440		
41	3 321	10 496	25 625	53 136	98 441	167 936	269 001	600 281	850 176		
42	3 402	10 752	26 250	54 432	100 842	172 032	275 562	614 922	870 912		
43	3 483	11 008	26 875	55 728	103 243	176 128	282 123	629 563	891 648		
44	3 564	11 264	27 500	57 024	105 644	180 224	288 684	644 204	912 384		
45	3 645	11 520	28 125	58 320	108 045	184 320	295 245	658 845	933 120		
46	3 726	11 776	28 750	59 616	110 446	188 416	301 806	673 486	953 856		
47	3 807	12 032	29 375	60 912	112 847	192 512	308 367	688 127	974 592		
48	3 888	12 288	30 000	62 208	115 248	196 608	314 928	702 768	995 328		
49	3 969	12 544	30 625	63 504	117 649	200 704	321 489	714 409	1 016 064		
50	4 050	12 800	31 250	64 800	120 050	204 800	328 050	732 050	1 036 800		

**ТАБЛИЦА ВЕРОЯТНОСТЕЙ  $P(\lambda)$ , СООТВЕТСТВУЮЩИХ  
РАЗНЫМ ЗНАЧЕНИЯМ  $\lambda$**

$\lambda$	$P(\lambda)$	$\lambda$	$P(\lambda)$
0,30	1,0000	1,10	0,1777
0,35	0,9997	1,20	1122
0,40	9972	1,30	0681
0,45	9874	1,40	0397
0,50	9639	1,50	0222
0,55	9228	1,60	0120
0,60	8643	1,70	0062
0,65	7920	1,80	0032
0,70	7112	1,90	0015
0,75	6272	2,00	0007
0,80	5441	2,10	0003
0,85	4653	2,20	0001
0,90	3927	2,30	0001
0,95	3275	2,40	0000
1,00	2700	2,50	0000

ТАБЛИЦА ВЕРОЯТНОСТЕЙ

$\chi^2 \backslash k$	1	2	3	4	5	6	7	8	9	10
1	0,3173	0,6065	0,8013	0,9098	0,9626	0,9856	0,9948	0,9982	0,9994	0,9998
2	1574	3679	5724	7318	8491	9197	9598	9810	9915	9963
3	0833	2231	3916	5578	7000	8088	8850	9344	9643	9814
4	0455	1353	2615	4060	5494	6767	7798	8571	9114	9473
5	0254	0821	1718	2873	4159	5438	6600	7576	8343	8912
6	0143	0498	1116	1991	3062	4232	5398	6472	7399	8153
7	0081	0302	0719	1359	2206	3208	4289	5366	6371	7254
8	0047	0183	0460	0916	1562	2381	3326	4335	5341	6288
9	0027	0111	0293	0611	1091	1736	2527	3423	4373	5321
10	0016	0067	0186	0404	0752	1247	1886	2650	3505	4405
11	0009	0041	0117	0266	0514	0884	1386	2017	2757	3575
12	0005	0025	0074	0174	0348	0620	1006	1512	2133	2851
13	0003	0015	0046	0113	0234	0430	0721	1119	1626	2237
14	0002	0009	0029	0073	0156	0296	0512	0818	1223	1730
15	0001	0006	0018	0047	0104	0203	0360	0591	0909	1321
16	0001	0003	0011	0030	0068	0138	0251	0424	0669	0996
17	0000	0002	0007	0019	0045	0093	0174	0301	0487	0744
18	—	0001	0004	0012	0029	0062	0120	0212	0352	0550
19	—	0001	0003	0008	0019	0042	0082	0149	0252	0403
20	—	0000	0002	0005	0013	0028	0056	0103	0179	0293
21	—	—	0001	0003	0008	0018	0038	0071	0126	0211
22	—	—	0001	0002	0005	0012	0025	0049	0089	0151
23	—	—	0000	0001	0003	0008	0017	0034	0062	0107
24	—	—	—	0001	0002	0005	0011	0023	0043	0076
25	—	—	—	0001	0001	0003	0008	0016	0030	0053
26	—	—	—	0000	0001	0002	0005	0010	0020	0037
27	—	—	—	—	0001	0001	0003	0007	0014	0026
28	—	—	—	—	0000	0001	0002	0005	0010	0018
29	—	—	—	—	—	0001	0001	0003	0006	0012
30	—	—	—	—	—	0000	0001	0002	0004	0009

$P(\chi^2)$  ДЛЯ КРИТЕРИЯ  $\chi^2$ 

11	12	13	14	15	16	17	18	19	20
0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
9985	0,9994	0,9998	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
9907	9955	9979	9991	0,9996	0,9998	0,9999	1,0000	1,0000	1,0000
9699	9834	9912	9955	9977	9989	9995	9998	0,9999	1,0000
9312	9580	9752	9858	9921	9958	9978	9989	9994	0,9997
8734	9161	9462	9665	9797	9881	9932	9962	9979	9989
7991	8576	9022	9347	9576	9733	9835	9901	9942	9967
7133	7851	8436	8893	9238	9489	9665	9786	9867	9919
6219	7029	7729	8311	8775	9134	9403	9597	9735	9829
5304	6160	6939	7622	8197	8666	9036	9319	9539	9682
4433	5289	6108	6860	7526	8095	8566	8944	9238	9462
3626	4457	5276	6063	6790	7440	8001	8472	8856	9161
2933	3690	4478	5265	6023	6728	7362	7916	8386	8774
2330	3007	3738	4497	5255	5987	6671	7291	7837	8305
1825	2414	3074	3782	4514	5246	5955	6620	7226	7764
1411	1912	2491	3134	3821	4530	5238	5925	6573	7166
1079	1496	1993	2562	3189	3856	4544	5231	5899	6530
0816	1157	1575	2068	2627	3239	3888	4557	5224	5874
0611	0885	1231	1649	2137	2687	3285	3918	4568	5218
0453	0671	0952	1301	1719	2202	2742	3328	3946	4579
0334	0504	0729	1016	1368	1785	2263	2794	3368	3971
0244	0375	0554	0786	1078	1432	1847	2320	2843	3405
0177	0277	0417	0603	0841	1137	1498	1906	2373	2888
0127	0203	0311	0458	0651	0895	1194	1550	1962	2424
0091	0148	0231	0346	0499	0698	0947	1249	1605	2014
0065	0107	0170	0259	0380	0540	0745	0998	1302	1658
0046	0077	0124	0193	0287	0415	0581	0790	1047	1353
0032	0055	0090	0142	0216	0316	0449	0621	0834	1094
0023	0039	0065	0104	0161	0239	0345	0484	0660	0878
0016	0028	0047	0076	0119	0180	0263	0374	0518	0699



## Продолжение

$\chi^2 \backslash k$	21	22	23	24	25	26	27	28	29
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
2	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
3	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
4	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
5	0,9999	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
6	9994	9997	0,9999	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000
7	9981	9990	9995	9997	0,9999	0,9999	1,0000	1,0000	1,0000
8	9951	9972	9984	9991	9995	9997	0,9999	0,9999	1,0000
9	9892	9933	9960	9976	9986	9992	9995	9997	0,9999
10	9789	9863	9913	9945	9967	9980	9988	9993	9996
11	9628	9747	9832	9890	9929	9955	9972	9983	9990
12	9396	9574	9705	9799	9866	9912	9943	9964	9977
13	9086	9332	9520	9661	9765	9840	9892	9929	9954
14	8696	9015	9269	9466	9617	9730	9813	9872	9914
15	8230	8622	8946	9208	9414	9573	9694	9784	9850
16	7696	8159	8553	8881	9148	9362	9529	9658	9755
17	7111	7634	8093	8487	8818	9091	9311	9486	9622
18	6490	7060	7575	8080	8424	8758	9035	9261	9443
19	5851	6453	7012	7520	7971	8364	8700	8981	9213
20	5213	5830	6419	6968	7468	7916	8308	8645	8929
21	4589	5207	5811	6387	6926	7420	7863	8253	8591
22	3995	4599	5203	5793	6357	6887	7374	7813	8202
23	3440	4017	4608	5198	5776	6329	6850	7330	7765
24	2981	3472	4038	4616	5194	5760	6303	6815	7289
25	2472	2971	3503	4058	4624	5190	5745	6278	6782
26	2064	2517	3009	3532	4076	4631	5186	5730	6255
27	1709	2112	2560	3045	3559	4093	4638	5182	5717
28	1402	1757	2158	2600	3079	3585	4110	4644	5179
29	1140	1449	1803	2201	2639	3111	3609	4125	4651
30	0920	1185	1494	1848	2243	2676	3142	3632	4140

$$F(x) = e^{-x}$$

$x$	$e^{-x}$	$x$	$e^{-x}$	$x$	$e^{-x}$	$x$	$e^{-x}$
1	2	1	2	1	2	1	2
0,00	1,0000	0,24	0,7866	48	0,6188	0,72	0,4868
01	0,9900	25	0,7788	49	0,6126	73	0,4819
02	0,9802	26	0,7711	50	0,6065	74	0,4771
03	0,9704	27	0,7634	51	0,6005	75	0,4724
04	0,9608	28	0,7558	52	0,5945	76	0,4677
05	0,9512	29	0,7483	53	0,5886	77	0,4630
06	0,9418	0,30	0,7408	54	0,5827	78	0,4584
07	0,9324	31	0,7334	55	0,5769	$\frac{1}{4}\pi = 0,7854$	0,4559
08	0,9231	32	0,7261	56	0,5712	79	0,4538
09	0,9139	33	0,7189	57	0,5655	0,80	0,4493
0,10	0,9048	34	0,7118	58	0,5599	81	0,4449
11	0,8958	35	0,7047	59	0,5543	82	0,4404
12	0,8869	36	0,6977	0,60	0,5488	83	0,4360
13	0,8781	37	0,6907	61	0,5434	84	0,4317
14	0,8694	38	0,6839	62	0,5379	85	0,4274
15	0,8607	39	0,6771	63	0,5326	86	0,4232
16	0,8521	0,40	0,6703	64	0,5273	87	0,4190
17	0,8437	41	0,6637	65	0,5220	88	0,4148
18	0,8353	42	0,6570	66	0,5169	89	0,4107
19	0,8270	43	0,6505	67	0,5117	0,90	0,4066
0,20	0,8187	44	0,6440	68	0,5066	91	0,4025
21	0,8106	45	0,6376	69	0,5016	92	0,3985
22	0,8025	46	0,6313	0,70	0,4966	93	0,3946
23	0,7945	47	0,6250	71	0,4916	94	0,3906

$$f(x) = e^{-x}$$

$x$	$e^{-x}$	$x$	$e^{-x}$	$x$	$e^{-x}$	$x$	$e^{-x}$
1	2	1	2	1	2	1	2
0,95	0,3867	1,21	0,2982	1,47	0,2299	1,72	0,1791
96	0,3829	22	0,2952	48	0,2276	73	0,1773
97	0,3791	23	0,2923	49	0,2254	74	0,1755
98	0,3753	24	0,2894	1,50	0,2231	75	0,1738
99	0,3716	25	0,2865	51	0,2209	76	0,1720
1,00	0,3679	26	0,2837	52	0,2187	77	0,1703
01	0,3642	27	0,2808	53	0,2165	78	0,1686
02	0,3606	28	0,2780	54	0,2144	79	0,1670
03	0,3570	29	0,2753	55	0,2122	1,80	0,1653
04	0,3535	1,30	0,2725	56	0,2101	81	0,1637
05	0,3499	31	0,2698	57	0,2080	82	0,1620
06	0,3465	32	0,2671	$\frac{1}{2}\pi = 1,5708$	0,207	83	0,1604
07	0,3430	33	0,2645	58	0,2060	84	0,1588
08	0,3396	34	0,2618	59	0,2039	85	0,1572
09	0,3362	35	0,2592	1,60	0,2019	86	0,1557
1,10	0,3329	36	0,2567	61	0,1999	87	0,1541
11	0,3296	37	0,2541	62	0,1979	88	0,1526
12	0,3263	38	0,2516	63	0,1959	89	0,1511
13	0,3230	39	0,2491	64	0,1940	1,90	0,1496
14	0,3198	1,40	0,2466	65	0,1920	91	0,1481
15	0,3166	41	0,2441	66	0,1901	92	0,1466
16	0,3135	42	0,2417	67	0,1882	93	0,1451
17	0,3104	43	0,2393	68	0,1864	94	0,1437
18	0,3073	44	0,2369	69	0,1845	95	0,1423
19	0,3042	45	0,2346	1,70	0,1827	96	0,1409
1,20	0,3012	46	0,2322	71	0,1809	97	0,1395
						98	0,1381
						99	0,1367
						2,00	0,1353

## ОГЛАВЛЕНИЕ

	<i>Стр.</i>
Раздел I. Вариационный ряд и его характеристики .	4
Раздел II. Основные сведения из теории вероятностей . . . . .	62
Раздел III. Выборочный метод . . . . .	95
Раздел IV. Кривые распределения . . . . .	121
Раздел V. Корреляция . . . . .	146
Практические задания для самостоятельной работы . .	177
Приложения . . . . .	179

---

*Илья Григорьевич Венецкий, Григорий Семенович Кильдишев*

**Пособие по математической статистике**

Редактор *Е. М. Шенцис*  
Техн. редактор *В. А. Виноградова*

Художник *В. В. Лепорк*  
Корректор *Л. И. Никифорова*

Сдано в набор 1/VIII 1956 г.

Подписано к печати 1/XII 1956 г.

Формат 60×92<sup>1</sup>/<sub>16</sub>  
А 11767

Объем 12,75 печ. л.  
Заказ 2543

Уч.-изд. л. 12,15  
Тираж 5 000  
Цена 4 р. 65 к.

Государственное статистическое издательство  
Москва, ул. Разина, 3, пом. 69

---

Типография № 5 Углетехиздата, Москва, Южно-портовый 1-й пр., 17

Цена ~~4~~ р. 65 к.