

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Владимирский государственный университет
имени Александра Григорьевича
и Николая Григорьевича Столетовых»
(ВлГУ)

Э.Ф. Касаткина

ОБРАБОТКА РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

**Методические указания к практическим занятиям по дисциплине
"Метрология, стандартизация, сертификация"**

(электронный ресурс)

ВлГУ – 2018 г.

Касаткина Э.Ф.

Обработка результатов измерений: Методические указания к практическим занятиям по дисциплине "Метрология, стандартизация, сертификация", Владимир: ВлГУ, - 2018. - 42 с.

Методические указания к практическим занятиям служит для изучения дисциплины «Метрология, стандартизация, сертификация» и выполнения расчетно-графической и контрольной работы по этой дисциплине.

Методические указания предназначены для студентов, обучающихся по направлению 27.03.01 «Стандартизация и метрология», 27.03.02 "Управление качеством".

ОГЛАВЛЕНИЕ

Введение	4
1. Прямые многократные равноточные измерения	5
1.1. Определение точечных оценок закона распределения	7
1.2. Исключение грубых погрешностей и промахов	9
1.3. Построение вариационного ряда	13
1.4. Проверка гипотезы о виде распределения эксперименталь- ных данных	20
1.5. Расчет доверительных границ ε (P) погрешности измерения	25
2. Оценка результатов неравноточных измерений	34
3. Косвенные измерения	35
Библиографический список	39
Приложения	40

ВВЕДЕНИЕ

Научно-технический прогресс, происходящий в настоящее время во всех областях науки и техники, во многом базируется на результатах измерений различных физических и нефизических величин, которым предшествует огромная, кропотливая предварительная работа людей различных специальностей. Это объясняется тем, что без опережающего развития теории и практики измерений как единственного способа получения количественной информации о величинах, характеризующих те или иные явления и процессы, невозможен прогресс практически во всех отраслях науки и техники.

С помощью результатов измерений, как результата процесса познания, человечество познаёт окружающий мир и на основании их стремится создать более благоприятные условия своей жизни. В результате этого человечество несёт значительные людские, материальные, финансовые, временные затраты на планирование, постановку, проведение и обработку результатов измерений. Так, на основании данных приведенных в [4], примерно 15% общественного труда затрачивается на проведение измерений, при этом от 3 до 6 % валового национального продукта передовых промышленных стран тратится на измерения и связанные с ними операции. Поэтому важность и значимость измерений трудно как переоценить, так и недооценить.

Известно, что важнейшей задачей, возникающей в процессе измерений, является обеспечение единства измерений и достоверности их результатов. Сотрудничество Российской Федерации с зарубежными странами требует взаимного доверия к измерительной информации, полученной с помощью общепринятых способов, методов и средств измерений. При этом первостепенное значение имеет высокая точность и достоверность результатов измерений, единообразие принципов и способов оценки точности этих результатов, а также условий проведения измерений.

1. ПРЯМЫЕ МНОГОКРАТНЫЕ РАВНОТОЧНЫЕ ИЗМЕРЕНИЯ

Прямые измерения – это такие измерения, когда значение изучаемой величины находят непосредственно из опытных данных. Для нахождения случайных погрешностей измерения необходимо провести несколько раз. Результаты таких измерений имеют близкие значения погрешностей.

Как показано в теории ошибок, наиболее близким к истинному значению Q измеряемой величины X является среднее арифметическое.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Следует иметь в виду, что среднее арифметическое значение рассматривается только как наиболее вероятное значение измеряемой величины.

Результаты отдельных измерений в общем случае отличается от истинного значения. Абсолютные погрешности i -го измерения: $\Delta_i = Q - X_i$ могут принимать как положительные, так и отрицательные значения с равной вероятностью. Суммируя погрешности получаем:

$$\sum_{i=1}^n \Delta_i = nQ - \sum_{i=1}^n X_i.$$

Тогда:

$$Q = \bar{X} + \frac{1}{n} \sum_{i=1}^n \Delta_i.$$

В этом выражении второе слагаемое в правой части при больших n равно нулю, так как всякой положительной погрешности можно поставить в соответствии равную ей отрицательную погрешность, тогда $Q = \bar{X}$. При ограниченном числе измерений будет лишь приближенное равенство: $Q \approx \bar{X}$. Таким образом, \bar{X} можно назвать действительным значением измеряемой величины.

На практике значение Q неизвестно, и следовательно, нельзя определить абсолютную погрешность Δ_i . Есть лишь определенная вероятность того, что Q находится в каком-то интервале вблизи \bar{X} и требуется определить этот интервал, соответствующий некоторой вероятности. В качестве

оценки абсолютной погрешности отдельного измерения используют значение:

$$\Delta X_i = \bar{X} - X_i.$$

Это значение определяет точность конкретного, i -го измерения.

Оценку погрешности ряда измерений (среднего значения) нельзя характеризовать простой суммой отклонений $\sum_{i=1}^n \Delta X_i = \sum_{i=1}^n (\bar{X} - X_i)$ так как она равна нулю, то при обработке результатов измерений используются абсолютные значения разностей. Такую оценку называют средней арифметической погрешностью:

$$\Delta = \frac{1}{n} \sum_{i=1}^n |\Delta X_i| = \frac{1}{n} \sum_{i=1}^n |X - X_i|.$$

Такая погрешность определяет пределы, в которых лежит более половины измерений. Следовательно, значение Q с достаточно большой вероятностью попадает в интервал от $\bar{X} - \Delta$ до $\bar{X} + \Delta$. Результаты измеряемой величины X в этом случае записываются в виде:

$$Q = \bar{X} \pm \Delta.$$

Величина X измерена тем точнее, чем меньше интервал, в котором находится истинное значение Q .

Пример: при измерении длины стержня было проведено шесть опытов, измеренные значения, см: 55,2; 55,0; 55,5; 55,9; 55,7; 55,2. Найти \bar{X} ; Δ ; δ ; γ .

Находи среднее арифметическое:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{55,2 + 55,0 + 55,5 + 55,9 + 55,7 + 55,2}{6} = 55,42 \text{ см.}$$

Погрешности отдельных измерений равны:

$$\Delta_1 = 55,42 - 55,2 = 0,22 \text{ см}; \quad \Delta_4 = 55,42 - 55,9 = - 0,48 \text{ см};$$

$$\Delta_2 = 55,42 - 55,0 = 0,42 \text{ см}; \quad \Delta_5 = 55,42 - 55,7 = - 0,28 \text{ см};$$

$$\Delta_3 = 55,42 - 55,5 = - 0,08 \text{ см}; \quad \Delta_6 = 55,42 - 55,2 = 0,22 \text{ см.}$$

Средняя арифметическая погрешность:

$$\Delta = \frac{1}{n} \sum_{i=1}^n |\Delta X_i| = \frac{1}{n} \sum_{i=1}^n |X - X_i| = \frac{0,22 + 0,42 + 0,08 + 0,48 + 0,28 + 0,22}{6} = 0,28 \text{ см}$$

Таким образом:

$$Q = 55,42 \pm 0,28 \text{ см.}$$

При этом относительная и приведенная погрешности:

$$\delta = \pm \frac{\Delta}{X} \cdot 100\% = \frac{0,28}{55,42} \cdot 100\% = 0,5\% ;$$

$$\gamma = \pm \frac{\Delta}{Q_N} \cdot 100\% = \frac{0,28}{150} \cdot 100\% = 0,19\%$$

При правильных и тщательно выполненных измерениях средняя арифметическая погрешность их результата близка к погрешности прибора.

Если измерения искомой величины X проведены много раз и в результате n измерений величины X , проведенных с одинаковой точностью, получается ряд значений: x_1, x_2, \dots, x_n , называемый выборкой, то такие измерения считаются равноточными.

Равноточными называют измерения, которые проводятся средствами измерений одинаковой точности по одной и той же методике при неизменных внешних условиях.

Обработка результатов многократных измерений заключается в нахождении оценки измеряемой величины и доверительного интервала, в котором находится ее истинное значение. Обработка должна проводиться в соответствии с ГОСТ Р 8.736-2011 «Государственная система обеспечения единства измерений. Измерения прямые многократные. Методы обработки результатов измерений. Основные положения».

Обработка результатов прямых равноточных измерений состоит из нескольких этапов.

1.1. Определение точечных оценок закона распределения

При рассмотрении дискретных величин используются точечные оценки параметров, т.е. оценки, выраженные одним числом. В отличие от числовых характеристик оценки являются случайными величинами, зависящими от числа наблюдений n . Для получения точечных оценок ряд зна-

чений измерений x_i - называемый выборкой должен быть представлена достаточным числом измерений.

Точечные оценки могут быть: состоятельными - при увеличении объема выборки стремятся по вероятности к истинному значению измеряемой величины (т.е. $\bar{X} \rightarrow Q$ при $n \rightarrow \infty$), несмещенными - математическое ожидание которых равно оцениваемой числовой характеристике (т.е. $Q = \bar{X}$), эффективными - та несмещенная оценка, которая имеет наименьшую дисперсию (т.е. $D[x] = \sigma_x^2 \rightarrow \min$).

Состоятельными несмещенными точечными оценками являются: среднее арифметическое значение, дисперсия.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i ; D[X] = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Если при измерении физической величины получают только постоянные значения т.е. $X = \bar{X}$, $D[X] = 0$. Если значения случайной величины X принимают значения не равные \bar{X} , то дисперсия ее отличается от нуля и положительна. Дисперсия таким образом, служит мерой колебаний (флуктуации) значений случайной величины.

Мера рассеяния результатов отдельных измерений X_i от среднего значения \bar{X} должна выражаться в тех же единицах, что и значения измеряемой величины.

Рассеяние результатов в ряду измерений – это несовпадение результатов одной и той же величины в ряду равноточных измерений, как правило, обусловлено действием случайных погрешностей при измерении и носит вероятностный характер. Оценками рассеяния результатов в ряду равноточных измерений одной и той же физической величины могут быть:

1. Размах результатов измерений вычисляется по формуле:
 $R_n = X_{\max} - X_{\min}$ где n - это объем выборки.

2. Средняя квадратическая погрешность результатов измерений:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

На практике широко распространен термин среднее квадратическое отклонение – отклонение результатов в ряду измерений от их среднего арифметического значения. В метрологии это отклонение называют погрешностью измерений, поэтому в соответствии с РМГ 29-99 целесообразно применять термин средняя квадратическая погрешность. При обработке ряда результатов, свободных от систематических погрешностей, средняя квадратическая погрешность и среднее квадратическое отклонение являются одинаковой оценкой рассеяния результатов измерений.

3. Средняя квадратическая погрешность среднего арифметического вычисляется по формуле:

$$\sigma_x = \frac{\sigma_x}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n(n-1)}}$$

4. Доверительные границы погрешности результата измерений – наибольшее и наименьшее значения погрешностей измерений, ограничивающие интервал, внутри которого с заданной вероятностью находится истинное (искомое) значение погрешности результата измерений.

1.2. Исключение грубых погрешностей и промахов

Грубая погрешность (или промах) – это случайная погрешность результата отдельного наблюдения, входящего в ряд измерений, которая для данных условий резко отличается от остальных результатов этого ряда. Грубые погрешности могут сильно исказить \bar{X} , σ_x и доверительный интервал. Промахи легко обнаруживаются в ряде измерений и их результаты отбрасываются.

Источниками грубых погрешностей нередко бывают ошибки, допущенные наблюдателями во время измерений, внезапные и кратковременные изменения условий измерений или оставшиеся незамеченными неисправности в аппаратуре. Наиболее характерны следующие грубые погрешности:

- неправильный отсчет по шкале измерительного прибора, происходящий из-за неверного учета цены малых делений шкалы;
- неправильная запись результата наблюдений, значений отдельных мер использованного набора;

- ошибки при манипуляциях с приборами или частями измерительной установки.

Вопрос о том, содержит ли данный результат наблюдений грубую погрешность решается общими методами статистических гипотез. Проверяемая гипотеза предполагает, что результат наблюдений X_i не содержит грубой погрешности, т.е. является одним из значений случайной величины X с определенным законом распределения. Сомнительным может быть в первую очередь наименьший X_{\min} или наибольший X_{\max} из результатов измерений.

Для проверки гипотезы можно воспользоваться распределением величин:

$$v = \frac{\bar{X} - X_{\min}}{\sigma_x} \quad \text{или} \quad v = \frac{X_{\max} - \bar{X}}{\sigma_x}.$$

Функции их распределения определяются методами теории вероятностей. Они совпадают между собой и для нормального распределения представлены в табл. 1.1. По данным этой таблицы при заданной доверительной вероятности P можно для числа измерений $n = 3 - 15$ предельные значения v_P , которые случайная величина v еще может принимать. Если вычисленное по опытным данным значение $v > v_P$, то результат рассматривается как промах и не принимается во внимание при дальнейшей обработке результатов наблюдений.

Т а б л и ц а 1.1.

n	v_P при P , равной				n	v_P при P , равной			
	0,9	0,95	0,975	0,99		0,9	0,95	0,975	0,99
3	1,406	1,412	1,414	1,414	10	2,097	2,294	2,414	2,540
4	1,645	1,689	1,710	1,723	11	2,146	2,383	2,470	2,606
5	1,731	1,869	1,917	1,955	12	2,190	2,387	2,519	2,663
6	1,894	1,996	2,067	2,130	13	2,229	2,426	2,562	2,714
7	1,974	2,093	2,182	2,265	14	2,297	2,461	2,602	2,759
8	2,041	2,172	2,273	2,374	15	2,326	2,493	2,638	2,808
9	2,097	2,146	2,237	2,464					

В общем случае границы для промахов зависят от объема выборки (СКО) и вида распределения т.е. $(|\bar{X} - X_i|) > t_{гр} \sigma_x$. Коэффициент $t_{гр}$ зависит от уровня значимости α :

При $\alpha < \frac{1}{n+1}$ $t_{гр} = 1,55 + 0,8\sqrt{\varepsilon - 1} \lg(\frac{n}{10})$, где ε – эксцесс распределения.

Правило трех сигм, используется для нормальных законов распределения и числа измерений $n = 20 \dots 50$. В этом случае считается, что результат, возникающий с вероятностью $P \leq 0,003$, не реален и его можно квалифицировать как промах. Сомнительный результат X_i отбрасывается если: $|\bar{X} - X_i| > 3\sigma_x$. Затем среднее значение и СКО вычисляются заново.

Критерий Романовского, при $n < 20$. Вычисляется отношение $\left| \frac{\bar{X} - X_i}{\sigma_x} \right| \geq t_p$ которое сравнивается с табличным (табл. 1.2.) при выбранном уровне значимости, в зависимости от числа наблюдений, если расчетное значение больше табличного, то результат считается промахом.

Т а б л и ц а 1.2

Значения критерия Романовского

P	Число измерений, n						
	n = 4	n = 6	n = 8	n = 10	n = 12	n = 15	n = 20
0,99	1,73	2,16	2,43	2,62	2,75	2,90	3,08
0,98	1,72	2,13	2,37	2,54	2,66	2,80	2,96
0,95	1,71	2,10	2,27	2,41	2,52	2,64	2,78
0,9	1,69	2,00	2,17	2,29	2,39	2,49	2,62

Критерий Шовинэ, при $n < 10$. Результат считается промахом, если разность $|\bar{X} - X_i|$ превышает значения σ_x :

$$|\bar{X} - X_i| > \left\{ \begin{array}{l} 1,4\sigma_x \text{ при } n = 3; \\ 1,7\sigma_x \text{ при } n = 6; \\ 1,9\sigma_x \text{ при } n = 8; \\ 2,0\sigma_x \text{ при } n = 10. \end{array} \right.$$

Исключение результатов выполняется в следующей последовательности. Сначала отбрасывают один результат с наибольшей по модулю разностью $|\bar{X} - X_i|$ и снова рассчитывают σ_x . Если и в этом случае неравенство выполняется, то вновь исключается результат с максимальным отклонением $|\bar{X} - X_i|$.

Критерий Шарлье, при $n > 5$ ($5 \dots 100$) $(|\bar{X} - X_i|) > t_{\text{Ш}} \sigma_x$. Значения критерия определяются по табл.1.3, в случае выполнения неравенства результат отбрасывается.

Т а б л и ц а 1.3.

Значения критерия Шарлье

n	5	10	20	30	40	50	100
$t_{\text{Ш}}$	1,3	1,65	1,96	2,13	2,24	2,32	2,58

Критерий Диксона при $n = 4 \dots 30$. Его особенность заключается в том, что результаты измерений раскладываются в вариационный возрастающий ряд. $X_1 \dots X_n$.

$$K_{\text{Д}} = \frac{X_n - X_{n-1}}{X_n - X_1}.$$

Расчетное значение сравнивается с табличным K_{α} , которое зависит от уровня значимости, в случае выполнения неравенства $K_{\text{Д}} > K_{\alpha}$ результат отбрасывается.

Т а б л и ц а 1.4.

Значения критерия Диксона

n	K_{α} при α , равном			
	0,10	0,05	0,02	0,01
4	0,68	0,76	0,85	0,89
6	0,48	0,56	0,64	0,70
8	0,40	0,47	0,54	0,59
10	0,35	0,41	0,48	0,53
14	0,29	0,35	0,41	0,45
16	0,28	0,33	0,39	0,43
18	0,26	0,31	0,37	0,41
20	0,24	0,30	0,36	0,39
30	0,22	0,26	0,31	0,34

1.3. Построение вариационного ряда

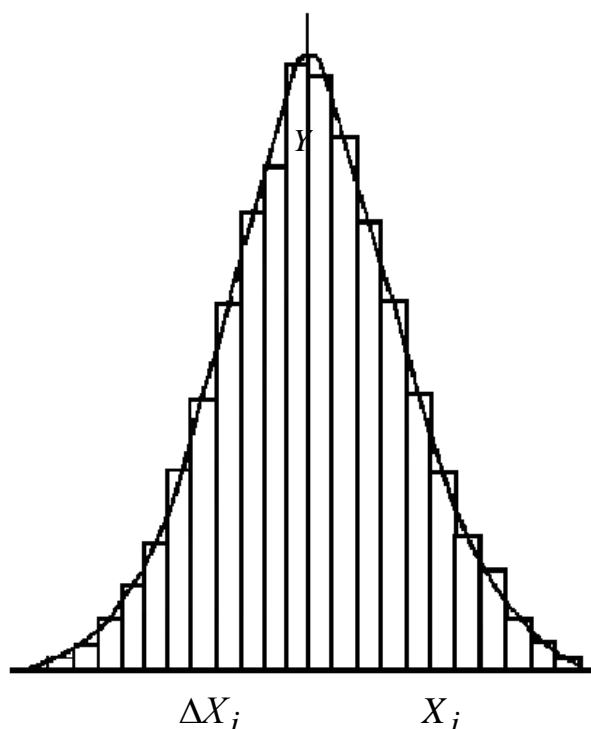


Рис. 1.1. Гистограмма результатов измерений

Построение вариационного ряда, заключающегося в расположении результатов измерений в порядке возрастания. Так как измерения искомой величины X были проведены много раз, то частоты появления того или иного значения X_i можно представить в виде графика, имеющего вид ступенчатой кривой, гистограммы (рис. 1.1), где Y – число отсчетов; $\Delta X_i = X_i - X_{i+1}$. С увеличением числа измерений и уменьшением интервала ΔX_i гистограмма в пределе переходит в кривую, характеризующую плотность распределения вероятности величины X_i оказаться в интервале ΔX_i .

Под распределением случайной величины понимают совокупность всех возможных значений случайной величины и соответствующих им вероятностей. Законом распределения случайной величины называют всякое соответствие случайной величины возможным значениям их вероятностей. Наиболее общей формулой закона распределения является функция распределения $P(x)$. Тогда функция $p(x) = P'(x)$ – плотность распределения вероятности или дифференциальная функция распределения. В общем случае вероятность может определяться различными законами распределения. До 1974 года было общепризнано, что погрешности должны быть распределены нормально. Затем на основании достаточно большого числа работ был законодательно факт разнообразия законов распределения. Были стандартизированы модели равномерного, трапецеидального, треугольного и нормального распределений. Знание формы закона распределения необходимо для определения связи его параметров. Так, например, значения погрешности с заданной доверительной ве-

роятностью без знания закона распределения не могут быть выражены через среднеквадратичное отклонение. И, наконец, изменение вида закона распределения погрешностей нередко может служить признаком какого-либо резкого изменения условий проведения эксперимента.

Возможность определения формы распределения экспериментальных данных ограничена малостью объема выборки ($n = 20-200$), и особенности распределения оказываются замаскированы случайностью самой выборки. При большом объеме выборки (несколько тысяч наблюдений) построение гистограммы позволяет получить плавную кривую, отражающую особенности наблюдаемого закона. Однако наиболее часто используется нормальный закон распределения:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\bar{X} - X_i)^2}{2\sigma^2}\right],$$

где σ^2 - дисперсия генеральной совокупности. Генеральной совокупностью называют все множество возможных значений измерений X_i или возможных значений погрешности ΔX_i .

Для определения формы распределения выборку необходимо представить в виде гистограммы, состоящей из m столбцов с определенной протяженностью h соответствующих им интервалов. Общепринято делать эти интервалы одинаковыми. Существует оптимальное число интервалов группирования, когда ступенчатая огибающая гистограмма наиболее близка к плавной кривой распределения. Если при группировании данных число интервалов будет слишком большим, то некоторые из них окажутся пустыми и малозаполненными. В этом случае гистограмма будет отличаться от плавной кривой распределения вследствие изрезанности многими всплесками и провалами, т.е. будет иметь "гребенчатый" вид. При малом числе m интервалов гистограмма будет отличаться от действительной кривой распределения вследствие слишком крупной ступенчатости, из-за чего характерные особенности будут потеряны.

Для распределения погрешностей одним из практических признаков приближения к оптимальному числу интервалов может служить исчезновение в гистограмме провалов и близким к оптимальному может считаться наибольшее m , при котором гистограмма еще сохраняет плавный характер.

Оптимальное число интервалов, как показано в некоторых исследованиях, мало зависит также от вида распределения.

В различных источниках можно найти целый ряд рекомендаций по выбору числа интервалов m . При выборе интервалов равной длины определяющим является требование, чтобы количество наблюдений, попавших в интервалы, было не слишком малым и сравнимым. При этом наиболее часто рекомендуется, чтобы количество наблюдений, попавших в интервал, было не менее 10. Отмечается, что на практике допустимо, чтобы количество наблюдений в крайних интервалах было менее 5.

$$m_{\min} = 0,55 \cdot n^{0,4}; m_{\max} = 1,25 \cdot n^{0,4}.$$

Искомое значение m , в этом случае, должно находиться в пределах от m_{\min} до m_{\max} и быть нечетным.

Во многих источниках для определения “оптимального” числа интервалов можно найти упоминание эвристической формулы Старджесса:

$$m = 3,31 \cdot \lg n + 1;$$

формулы Брукса и Каррузера:

$$m = 5 \cdot \lg n;$$

рекомендуемое соотношение:

$$m = n^{1/2}.$$

При больших объемах выборок n разброс значений m , задаваемых различными формулами, достаточно велик. Поэтому на практике при выборе числа интервалов больше руководствуются разумными соображениями, выбирая число интервалов так, чтобы в интервалы попадало число наблюдений не менее 5-10. Так, например, в рекомендациях ВНИИМетрологии в зависимости от n предлагаются следующие величины m представленные в табл. 1.5.

Длина интервала разбиения определяемой по формуле:

$$h = \frac{X_{\max} - X_{\min}}{m}.$$

Установив границы интервалов, подсчитывают число результатов измерений n_k , попавших в каждый интервал.

Т а б л и ц а 1.5.

n	m
40–100	7–9
100–500	8–12
500–1000	10–16
1000–10000	12–22

По полученным значениям рассчитывают вероятности попадания результатов измерений (частности) в каждый из интервал группирования по формуле:

$$p_k = \frac{n_k}{n}, \text{ где } k = 1; 2; \dots; m.$$

Произведенные расчеты позволяют построить гистограмму, полигон, кумулятивную кривую, определить показатели среднего уровня вариационного ряда – среднюю взвешенную, моду и медиану. Полигон представляет собой ломанную кривую, соединяющую середины верхних оснований каждого столбца гистограммы. Кумулятивная кривая – это график статистической функции распределения:

$$F_k = \sum_{k=1}^k p_k.$$

Для построенного интервального вариационного ряда расчет средней арифметической должен быть выполнен по формуле средней арифметической взвешенной, где совокупный объем находится не путем суммирования всех значений признака, а путем перемножения (взвешивания) вариантов признака X_i на их частоты (n_k) и последующего сложения произведений $X_i n_k$, число которых (произведений) равно количеству интервалов m . Следовательно, взвешивание — это лишь технический прием, посредством которого суммирование одинаковых значений заменяется умножением этих значений на их частоты.

$$\bar{X} = \frac{\sum \bar{X}_i n_k}{\sum n_k}$$

В формулах средней арифметической взвешенной, рассчитываемой для интервального вариационного ряда, в качестве \bar{X}_i принято брать середину интервала, исходя из предположения о равномерном распределении единиц совокупности на данном интервала. Середину интервала находится как полусумма значений его нижней и верхней границ (при условии, что верхняя граница данного интервала совпадает с нижней границей следующего интервала). Следовательно, возможно несовпадение средних арифметических, вычисленных на основе исходных данных и на основе вариационного ряда. Это связано с тем, что сделанные допущения о равномерном распределении значений на интервале не всегда выполняется. Среднее, вычисленные на основе интервального ряда, являются приближенными. Степень точности зависит от того, в какой мере распределение единиц внутри интервала приближается к такому распределению, для которого средняя арифметическая взвешенная совпадает с серединой интервала. Точность средней зависит также от длины интервала. Чем уже интервал, тем меньше ошибка, вызванная тем, что середина интервала принимается в качестве среднего его значения.

Достоинства использования среднего арифметического:

- это «центр тяжести» всех данных;
- при его расчете используются все результаты наблюдений;
- нет необходимости в сортировки данных.

Недостатки использования среднего арифметического:

- «резко» выделяющиеся значения могут испортить картину;
- могут потребоваться значительные вычисления;
- среднее значение может не совпадать ни с одним из фактических значений.

Мода — это значение признака, которое чаще всего встречается в вариационном ряду. Во многих случаях эта величина наиболее характерна для ряда распределения и вокруг нее концентрируется большая часть вариантов. При изменении распределения в его концах мода не меняется, т.е. она обладает определенной устойчивостью к вариации признака. Поэтому моду наиболее удобно применять при изучении рядов с неопределенными границами.

Для интервального ряда с равными интервалами длиной h сначала определяется модальный интервал $X_{k-1} \div X_k$, которому соответствует

максимальная частота n_k или частость $p_k = \frac{n_k}{n}$, где $k = 1; 2; \dots; m$. Значение моды внутри модального интервала определяется по интерполяционной формуле как абсцисса точки максимума кривой распределения:

$$M_0 = X_{k-1} + h_k \frac{n_k - n_{k-1}}{(n_k - n_{k-1}) + (n_k - n_{k+1})},$$

где X_{k-1} нижняя граница модального интервала;

h_k длина модального интервала;

n_{k-1}, n_k, n_{k+1} — частота интервала, соответственно предшествующего модальному, модального и следующего за модальным.

Графически моду определяют по гистограмме распределения. Для этого выбирают самый высокий прямоугольник, который и является модальным, далее верхнюю правую вершину модального прямоугольника соединяют с верхней правой вершиной предшествующего прямоугольника, а верхнюю левую вершину модального прямоугольника с верхней левой вершиной последующего прямоугольника. Абсцисса точки пересечения этих отрезков и будет модой распределения.

Достоинства использования моды:

- «резко» выделяющиеся значения не влияют на результат;
- значение можно визуально определить по графику.

Недостатки использования моды:

- данные могут и не иметь моды.

При обработке результатов измерений или в статистическом анализе часто применяют структурные, или порядковые, средние, например медиану.

В отличие от средней арифметической, на которую оказывают влияние все значения вариационного ряда, структурные средние совершенно не зависят от крайних значений.

Медианой называют такое значение признака, которое приходится на середину ранжированного ряда. Таким образом, в ранжированном ряду распределения одна половина ряда имеет значения признака, превышающие медиану, другая — меньше медианы.

Ряд с четным числом членов делит пополам не одна, а две единицы совокупности. Так, например, в распределении из 50 измерений в середине

ряда расположены единицы совокупности под номерами 25 и 26, а для 100 измерений – 50 и 51.

$$\text{Тогда } Me = \frac{X_{25} + X_{26}}{2} \quad \text{или} \quad Me = \frac{X_{50} + X_{51}}{2}.$$

Однако на практике для простоты счета номер медианы при четном числе членов ряда определяется как: $\frac{1}{2} \sum n_k$.

$$\text{Номер медианы для ряда с нечетным числом членов равен: } \frac{n+1}{2}.$$

В случае интервального вариационного ряда медиану определяют в такой последовательности. Прежде всего находят медианный интервал. Для этой цели используются накопленные частоты. Соответственно номер медианы определяют по указанной выше формуле, а нахождение медианы на данном интервале осуществляется по следующей интерполяционной формуле:

$$M_e = X_{k-1} + h_k \frac{\frac{1}{2} \sum n_k - p_{k-1}}{p_k},$$

где X_{k-1} — нижняя граница медианного интервала;

h_k — длина медианного интервала;

p_{k-1} — накопленная частота интервала, предшествующего медианному;

p_k — накопленная частота медианного интервала.

Из определения медианы следует, что она не зависит от тех значений признака, которые расположены по обе стороны от нее. В связи с этим медиана является лучшей характеристикой центральной тенденции в тех случаях, когда концы распределений расплывчаты (например, границы крайних интервалов открыты) или в ряду распределения имеются чрезмерно большие или малые значения.

В интервальном ряду медиану можно определить графически. Медиана рассчитывается по кумулятивной кривой. Для этого из точки на шкале накопленных частот (частостей), соответствующей $\frac{1}{2} \sum n_k$, проводится прямая, параллельная оси абсцисс, до пересечения с кумулятой. За-

тем из точки пересечения указанной прямой с кумулятой опускается перпендикуляр на ось абсцисс. Абсцисса точки пересечения и является медианой.

Достоинства использования медианы:

- позволяет представить, где расположена наибольшая часть данных.

Недостатки использования медианы:

- используются не все данные;

- «резко» выделяющиеся значения могут быть существенными.

Для нормального распределения среднее значение равно моде и равно медиане.

1.4. Проверка гипотезы о виде распределения экспериментальных данных

Все предположения о характере распределения являются гипотезами, а не категорическими утверждениями. Следовательно, они должны быть подвергнуты статистической проверке с помощью так называемых *критериев согласия*. Критерии согласия, опираясь на установленный закон распределения, дают возможность установить, когда расхождения между теоретическими и эмпирическими (опытными) данными следует признать незначительными (случайными), а когда — существенными (неслучайными). Критерии согласия позволяют отвергнуть или подтвердить правильность выдвинутой гипотезы о характере распределения в вариационном ряду и дать ответ, можно ли принять для данного вариационного ряда модель, выраженную некоторым теоретическим законом распределения. Существует ряд критериев согласия. Чаще других применяют критерии Пирсона, Романовского и Колмогорова.

Критерий согласия Пирсона χ^2 (хи-квадрат) — один из основных критериев согласия. Критерий используется для оценки случайности (существенности) расхождений между частотами эмпирического и теоретического распределений и заключается в расчете коэффициента χ^2 по формуле:

$$\chi^2 = \sum_{i=1}^m \frac{(n_k - N_i)^2}{N_i},$$

где m – число групп, на которые разбито эмпирическое распределение;

n_k – наблюдаемая частота признака в k -й группе;

N_i – теоретическая частота, рассчитанная по предполагаемому распределению.

Для определения расчетного значения χ^2 вычисляют число наблюдений для каждого из интервалов теоретически соответствующий нормальному закону распределения, для этого от реальных середин интервалов X_{i0} переходят к нормированным по формуле:

$$Z_i = \frac{X_{i0} - \bar{X}}{\sigma_x} \quad (i = 1 \dots m).$$

Для каждого значения Z_i находят значение функции плотности вероятностей $f(z_i)$:

$$f(z_i) = \frac{1}{\sqrt{2 \cdot \pi}} e^{-\frac{z_i^2}{2}} \quad (i = 1 \dots m).$$

По найденному значению $f(z_i)$ определяется та часть N_i общего числа имеющихся наблюдений, которая теоретически должна быть в каждом из интервалов:

$$N_i = \frac{n \cdot h \cdot f(z_i)}{\sigma_x}.$$

Для распределения χ^2 составлены таблицы, где указано критическое значение критерия согласия χ^2 для выбранного уровня значимости α и данного числа степеней свободы ν .

Уровень значимости α — вероятность ошибочного отклонения выдвинутой гипотезы, т.е. вероятность того, что будет отвергнута правильная гипотеза. В статистических исследованиях в зависимости от важности и ответственности решаемых задач пользуются следующими тремя уровнями значимости:

- 1) $\alpha = 0,10$, тогда $P = 0,90$;
- 2) $\alpha = 0,05$, тогда $P = 0,95$;
- 3) $\alpha = 0,01$, тогда $P = 0,99$.

Например, вероятность 0,01 означает, что в одном случае из 100 может быть отвергнута правильная гипотеза. В экономических исследованиях считается практически приемлемой вероятностью ошибки 0,05, т.е. в 5 случаях из 100 может быть отвергнута правильная гипотеза.

Кроме того, χ^2 критерий, определяемый по таблице (см. табл. 1.6.), зависит и от числа степеней свободы. Число степеней свободы ν определяется как число групп в ряду распределения m минус число связей z .

$$\nu = k - z.$$

Под числом связей обычно понимается число показателей эмпирического (вариационного) ряда, использованных при исчислении теоретических частот, т.е. показателей, связывающих эмпирические и теоретические частоты ($\bar{X}; \sigma; \sum_i n_k$)

Так, например, в случае выравнивания по кривой нормального распределения имеется три связи:

$$\bar{X}_{\text{эмп}} = \bar{X}_{\text{теор}}; \sigma_{\text{эмп}} = \sigma_{\text{теор}}; \sum_k n_{k\text{эмп}} = \sum_i N_{i\text{теор}}.$$

Поэтому при выравнивании по кривой нормального распределения число степеней свободы определяется как $\nu = m - 3$, где m — число групп в ряду. В случае выравнивания по кривой Пуассона, из-за особенности распределения ($\bar{X} = D_x$), имеется две связи:

$$\bar{X}_{\text{эмп}} = \bar{X}_{\text{теор}}; \sum_k n_{k\text{эмп}} = \sum_i N_{i\text{теор}}.$$

При полном совпадении теоретического и эмпирического распределений $\chi^2 = 0$, в противном случае $\chi^2 > 0$. Если $\chi_{\text{расч}}^2 > \chi_{\text{табл}}^2$ при заданном уровне значимости α и числе степеней свободы ν гипотезу о несущественности (случайности) расхождений отклоняем.

В случае если $\chi_{\text{расч}}^2 \leq \chi_{\text{табл}}^2$ заключаем, что эмпирический ряд хорошо согласуется с гипотезой о предполагаемом распределении и с вероятностью $(1 - \alpha)$ можно утверждать, что расхождение между теоретическими и эмпирическими частотами случайно.

Используя критерий согласия χ^2 , необходимо соблюдать следующие условия:

1) объем исследуемой совокупности должен быть достаточно большим ($n > 50$), при этом частота или численность каждой группы должна быть не менее 5. Если это условие нарушается, необходимо предварительно объединить маленькие частоты;

2) эмпирическое распределение должно состоять из данных, полученных в результате случайного отбора, т.е. они должны быть независимыми.

Т а б л и а 1.6

Значения χ^2_α при различном уровне значимости

ν	χ^2_α при уровне значимости α , равном								
	0,99	0,95	0,9	0,8	0,5	0,2	0,1	0,05	0,02
2	0,02	0,1	0,21	0,45	1,39	3,22	4,61	5,99	7,82
4	0,3	0,71	1,06	1,65	3,36	5,99	7,78	9,49	11,67
6	0,87	1,63	2,20	3,07	5,35	8,56	10,65	12,59	15,03
8	1,65	2,73	3,49	4,59	7,34	11,03	13,36	15,51	18,17
10	2,56	3,94	4,87	6,18	9,34	13,44	15,99	18,31	21,16
12	3,57	5,23	6,30	7,81	11,34	15,81	18,55	21,03	24,05
14	4,66	6,57	7,79	9,47	13,34	18,15	21,06	23,69	26,87
16	5,81	7,96	9,31	11,2	15,34	20,46	23,54	26,3	29,63
20	8,26	10,85	12,44	14,58	19,34	25,04	28,41	31,41	35,02
25	11,52	14,61	16,47	18,94	24,34	30,68	34,38	37,65	41,57
30	14,95	18,46	20,60	23,36	29,34	36,25	40,26	43,77	47,96

Критерий Романовского с основан на использовании критерия χ^2 Пирсона, т.е. уже найденных значений χ^2 , и числа степеней свободы ν :

$$c = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}}$$

Он весьма удобен при отсутствии таблиц для χ^2 .

Если $c < 3$, то расхождения между теоретическим и эмпирическим распределением случайны, если же $c > 3$, то не случайны и, соответствен-

но, теоретическое распределение не может служить моделью для изучаемого эмпирического распределения (вариационного ряда).

Критерий Колмогорова λ основан на определении максимального расхождения между накопленными частотами или частостями (суммарными частотами) эмпирических и теоретических распределений:

$$\lambda = \frac{D}{\sqrt{n}},$$

где D максимальная разность между накопленными (кумулятивными) частотами ($F_k - F'_k$) эмпирического и теоретического рядов распределений;

n - число единиц в совокупности.

Рассчитав значение λ , по таблице $P(\lambda)$ табл. 1.7 определяют вероятность, с которой можно утверждать, что отклонения эмпирических частот от теоретических случайны. Вероятность $P(\lambda)$ может изменяться от 0 до 1. При $P(\lambda) = 1$ происходит полное совпадение частот, при $P(\lambda) = 0$ — полное расхождение. Если X принимает значения до 0,3, то $P(\lambda) = 1$.

Т а б л и ц а 1.7

Значения функции $P(\lambda)$

λ	P	λ	P
0,30	1,0000	1,10	0,1777
0,35	0,9997	1,20	0,1122
0,40	0,9972	1,30	0,0681
0,45	0,9874	1,40	0,0397
0,50	0,9639	1,50	0,0222
0,55	0,9228	1,60	0,0120
0,60	0,8643	1,70	0,0062
0,65	0,7920	1,80	0,0032
0,70	0,7112	1,90	0,0015
0,75	0,6272	2,00	0,0007
0,80	0,5441	2,10	0,0003
0,85	0,4653	2,20	0,0001
0,90	0,3927	2,30	0,0001
0,95	0,3275	2,40	0,0000
1,00	0,2700	2,50	0,0000

Основное условие для использования критерия Колмогорова — достаточно большое число наблюдений.

1.5 Расчет доверительных границ $\varepsilon(P)$ погрешности измерения

При нормальном законе распределения, доверительные границы вычисляются по формуле:

$$\varepsilon(P) = t_{\alpha} \cdot \sigma_{\bar{x}},$$

где t_{α} - коэффициент Стьюдента.

Доверительный интервал может быть получен двумя способами: с использованием коэффициента Стьюдента и Функции Лапласа. В обоих случаях необходимо найти точечные оценки: среднее значение и среднее квадратичное отклонение, выбрать доверительную вероятность P .

Затем при использовании коэффициента Стьюдента необходимо его определить по табл. в зависимости от числа измерений n и выбранной P . Результат измерения может быть записан в виде:

$$Q = \bar{X} \pm t_{\alpha} \frac{\sigma_x}{\sqrt{n}} = \bar{X} \pm t_{\alpha} \sigma_{\bar{x}}; P[\bar{X} - t_{\alpha} \sigma_{\bar{x}} < X < \bar{X} + t_{\alpha} \sigma_{\bar{x}}].$$

При использовании функции Лапласа интервал приобретет вид:

$$Q = \bar{X} \pm z_p \frac{\sigma_x}{\sqrt{n}} = \bar{X} \pm z_p \sigma_{\bar{x}};$$

$$P[\bar{X} - z_p \sigma_{\bar{x}} < X < \bar{X} + z_p \sigma_{\bar{x}}] = 2F(z_p)$$

Доверительный интервал по коэффициенту Стьюдента рассчитывается при числе измерений меньше 20, так как при числе измерений $20 \leq n \leq 30$ распределение становится нормальным.

Если закон распределения параметра неизвестен и нет оснований утверждать, что он близок к нормальному, то используется функция Лапласа.

В общем виде результат может быть записан: $Q = \bar{X} \pm t \frac{\sigma_x}{\sqrt{n}} = \bar{X} \pm t \sigma_{\bar{x}}$ где t

□ положительное число, зависящее от n .

Расчет доверительных границ суммарной неисключенной систематической составляющей погрешности измерения. $\theta(P) = K \cdot \sqrt{\sum_{j=1}^k \theta_j^2}$, где K

- коэффициент соответствующей выбранной доверительной вероятности (при $P = 0,95$ $K = 1,1$).

Вычислим отношение $\frac{\theta(P)}{\sigma_{\bar{X}}}$.

В зависимости от отношения неисключенной систематической составляющей погрешности и оценки СКО результат измерения записывается:

Значение отношения $\frac{\theta(P)}{\sigma_{\bar{X}}}$	Результат измерения
$\theta(P)/\sigma_{\bar{X}} < 0,8$	$\bar{X} \pm \varepsilon(P)$
$0,8 \leq \theta(P)/\sigma_{\bar{X}} \leq 8$	$\bar{X} \pm k_p [\varepsilon(P) + \theta(P)]$
$\theta(P)/\sigma_{\bar{X}} > 8$	$\bar{X} \pm \theta(P)$

k_p – определяется по табл. в зависимости от вероятности и значения соотношения, изменяется в пределах: $0,7 \div 0,85$.

Доверительные границы общей погрешности измерения:

$$\Delta p = k \cdot \sigma_{\Sigma} ;$$

$$k = \frac{\theta(P) + \varepsilon(P)}{\sigma_{\bar{X}} + \sqrt{\frac{1}{3}\theta(P)^2}} = \frac{\theta(P) + \varepsilon(P)}{\sigma_{\bar{X}} + \theta(P)/\sqrt{3}} ; \quad \sigma_{\Sigma} = \sqrt{\frac{1}{3}\theta(P)^2 + \sigma_{\bar{X}}^2} .$$

где k - коэффициент, σ_{Σ} - суммарное среднее квадратичное отклонение результата измерения.

Следовательно, результат измерений можно записать в виде:

$$X = \bar{X} \pm \Delta p$$

Если величина $\varepsilon(P)$ окажется сравнимой с абсолютной погрешностью погрешность средства измерения, то в качестве доверительного интервала следует взять величину: $\Delta a = \sqrt{\varepsilon(P)^2 + (\frac{t}{3} \cdot \Delta)^2}$, где Δ - погрешность средства измерения. Окончательный результат записывается в виде:

$$\bar{X} \pm \varepsilon(P) \quad \text{или} \quad \bar{X} \pm \Delta a$$

Относительная погрешность результата серии измерений выразится как:

$$\delta = \pm \frac{\varepsilon(P)}{\bar{X}} \text{ или } \delta = \pm \frac{\Delta a}{\bar{X}}$$

Пример: Объем выборки для обработки является ряд из n результатов измерений ($n = 150$). Результаты наблюдений представлены в табл. 1.8.

Т а б л и ц а 1.8

Результаты наблюдений

479	495	497	491	501	492	505	490	506	490
483	505	495	498	499	506	502	501	499	495
497	501	508	510	505	493	491	509	508	512
495	492	499	479	496	502	512	480	494	498
500	509	511	515	505	482	495	500	514	510
488	500	494	502	495	517	505	515	505	490
502	482	514	505	515	503	487	495	517	515
496	512	515	490	483	479	497	517	497	501
505	493	480	495	517	505	517	496	487	502
499	501	516	521	502	517	499	514	517	492
490	503	501	495	492	485	523	490	493	521
506	488	507	515	506	499	502	508	515	501
510	513	491	500	481	501	504	501	495	514
501	498	514	498	509	501	491	512	515	482
494	505	497	493	496	494	502	499	505	500

Определяем точечные оценки закона распределения результатов измерения:

– среднее арифметическое значение \bar{X} измеряемой величины по

формуле: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i; \bar{X} = 500,6$

– среднее квадратичное отклонение (СКО) результата измерения:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}; \sigma = 10,07$$

Исключение грубых погрешностей по критерию “трех сигм“ применяется для результатов измерений, распределенных по нормальному закону. По этому критерию считается, что результат, возникающий с вероятностью $q \leq 0,003$, маловероятен и его можно считать промахом, если $|\bar{X} - X_i| > 3 \cdot \sigma$.

Грубые погрешности и промахи исключаются, после чего проводится повторный расчет \bar{X} и σ .

$$X_{\min} = 479; X_{\max} = 523; 3 \cdot \sigma_x = 30,22466$$

$$\bar{X} - 3\sigma_x = 470,3753 < X_{\min} = 479$$

$\bar{X} + 3\sigma_x = 530,8247 > X_{\max} = 523$, следовательно, грубых погрешностей и промахов в результатах наблюдений нет.

Проводим построение вариационного ряда, заключающегося в расположении результатов измерений в порядке возрастания. Затем вариационный ряд разбивается на оптимальное число m одинаковых интервалов группирования длиной h , определяемой по формуле: $h = \frac{X_{\max} - X_{\min}}{m}$;

$$m_{\min} = 0,55 \cdot n^{0,4}; m_{\min} = 4,08 \quad m_{\max} = 1,25 \cdot n^{0,4}; m_{\max} = 9,28$$

Искомое значение m должно находиться в пределах от m_{\min} до m_{\max} и быть нечетным. Принимаем $m = 7$.

$$\text{Тогда } h = \frac{X_{\max} - X_{\min}}{m} = \frac{523 - 479}{7} = 6,285714.$$

Установив границы интервалов, подсчитываем число результатов измерений n_k , попавших в каждый интервал рис. 1.2. Результаты заносим в табл. 4.5.

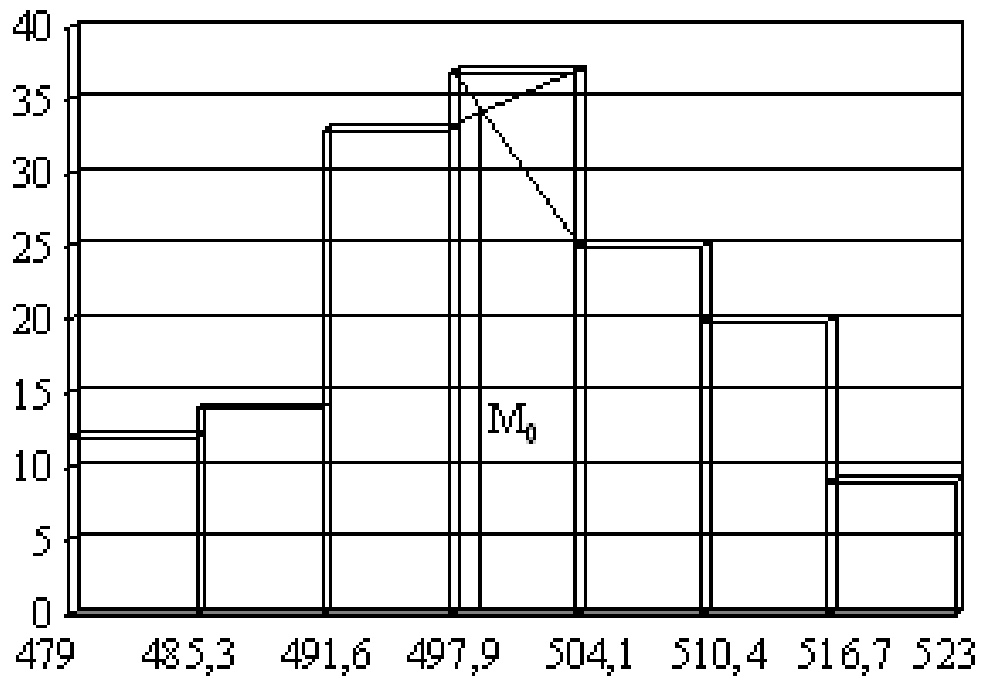


Рис. 1.2 Гистограмма распределения значений

По полученным значениям рассчитываем вероятности попадания результатов измерений (частоты) в каждый из интервал группирования по формуле: $p_k = \frac{n_k}{n}$, где $k = 1; 2; \dots; m$ рис. 4.2. Результаты заносим в табл. 1.9.

Произведенные расчеты позволяют построить гистограмму и полигон. Полигон представляет собой ломанную кривую, соединяющую середины верхних оснований каждого столбца гистограммы.

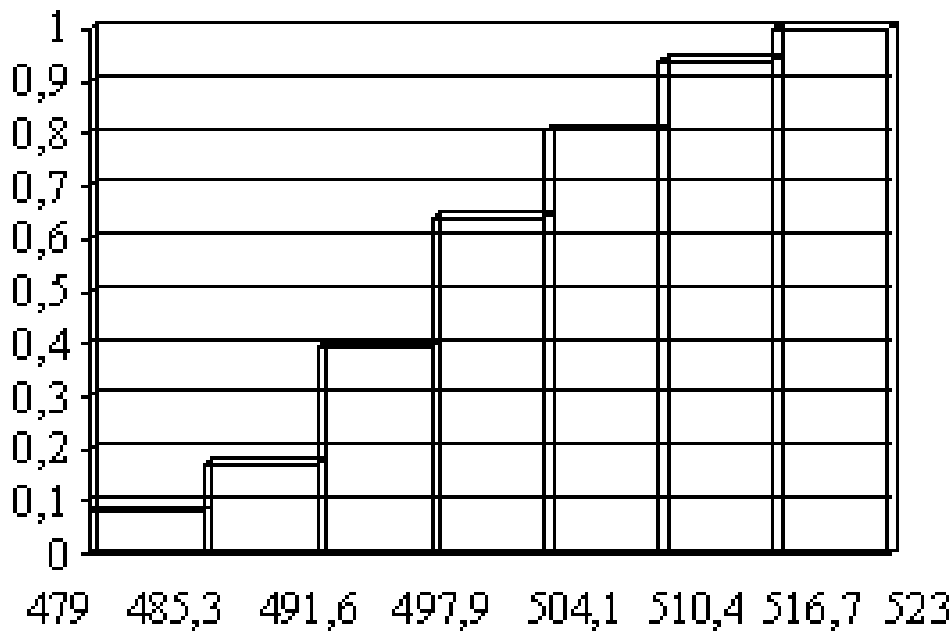


Рис. 1.3. Кумулятивная кривая

$$M_0 = X_{k-1} + h_k \frac{n_k - n_{k-1}}{(n_k - n_{k-1}) + (n_k - n_{k+1})} =$$

$$= 497,86 + 6,29 \frac{37 - 33}{(37 - 33) + (37 - 25)} = 499,43$$

$$M_e = X_{k-1} + h_k \frac{\frac{1}{2} \sum n_k - p_{k-1}}{p_k} = 497,86 + 6,29 \frac{\frac{150}{2} - 59}{96} = 498,91$$

Построение медианы распределения на рис. 1.4.

Проводим проверка гипотезы о виде распределения экспериментальных данных. При числе экспериментальных данных $n \geq 50$ для проверки критерия согласования теоретического распределения с практическим чаще всего используют критерий Пирсона (χ^2). Идея этого метода состоит в контроле отклонений гистограммы экспериментальных данных от гистограммы с таким же числом интервалов, построенной на основе нормального распределения.

Вычисляем число наблюдений для каждого из интервалов теоретически соответствующих нормальному закону распределения, для этого от ре-

альных середин интервалов X_{i0} переходим к нормированным по формуле:

$$Z_i = \frac{X_{i0} - \bar{X}}{\sigma} \quad (i = 1 \dots m).$$

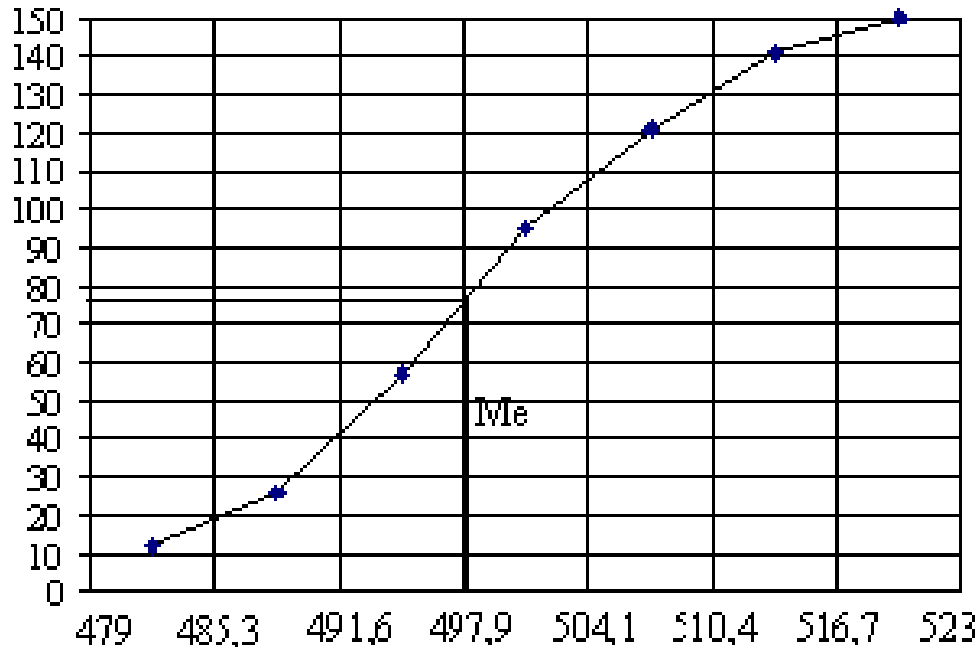


Рис. 1.4. Построение медианы

Для каждого значения Z_i находим значение функции плотности вероятностей $f(Z_i)$:

$$f(z_i) = \frac{1}{\sqrt{2 \cdot \pi}} e^{-\frac{z_i^2}{2}} \quad (i = 1 \dots m).$$

По найденному значению $f(Z_i)$: определяем ту часть N_i общего числа имеющихся наблюдений, которая теоретически должна быть в каждом из интервалов:

$$N_i = \frac{n \cdot h \cdot f(z_i)}{\sigma}$$

Вычисляем величину χ^2 .
$$\chi^2 = \sum_{i=1}^m \frac{(n_k - N_i)^2}{N_i}$$

Если в какой-то из интервалов, теоретически попадает меньше 5 наблюдений, то в обеих гистограммах его соединяем с соседним интервалом. После этого определяем число степеней свободы:

$\nu = m - 1 - r$, где m - общее число интервалов; r - число определяемых по статистике параметров, необходимых для совмещения модели и гистограммы ($r = 2$).

Выбираем (по таблице) уровень значимости q , который должен быть небольшим. По уровню значимости и числу степеней свободы находим границу критической области χ_q^2 (табл.). Если $\chi_q^2 > \chi^2$, то гипотеза о нормальном значении принимается. 44считанные значения представлены в табл. 1.9:

Т а б л и ц а 1.9

	Интервал	n_k	p_k	X_{i0}	Z_i	$f(Z_i)$	N_i	χ^2
1	479 ÷ 485,2857	12	0,08	482,1429	-1,83199	0,074494	7	3,627065
2	485,2857÷ 491,5714	14	0,173333	488,4286	-1,2081	0,192302	18	0,887545
3	491,5714÷ 497,8571	33	0,393333	494,7143	-0,5842	0,336357	31	0,073595
4	497,8571÷ 504,1429	37	0,64	501	0,039703	0,398628	37	0,002503
5	504,1429÷ 510,4286	25	0,806667	507,2857	0,663602	0,3201	30	0,820088
6	510,4286÷ 516,7143	20	0,94	513,5714	1,287501	0,174162	16	0,840407
7	516,7143÷ 523	9	1	519,8571	1,9114	0,064206	6	1,489181
Σ		150					145	7,740383

$\chi_{0,1}^2 = 7,78$; $7,78 > 7,74 \Rightarrow$ Гипотеза о совпадении экспериментально-го и выбранного теоретического (нормального) распределений принимается.

Если табличные значения для критерия Пирсона отсутствуют можно воспользоваться критерием Романовского:

$$c = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}} = \frac{|7,74 - 4|}{\sqrt{2 \cdot 4}} = 1,32.$$

Значение $c < 3$, следовательно, расхождения между теоретическим и эмпирическим распределением случайны, и теоретическое распределение может служить моделью для изучаемого эмпирического распределения (вариационного ряда).

Обработка результатов прямых равноточных измерений.

Определяем СКО среднего арифметического:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (X_i - \bar{X})^2} = 0,8221.$$

Проводим расчет доверительных границ ε погрешности измерения. При нормальном законе распределения, доверительные границы вычисляются по формуле: $\varepsilon(P) = t \cdot \sigma_{\bar{X}}$, где t - коэффициент Стьюдента, $t = 1,96$.

При нормальном законе распределения результатов измерений истинное значение измеряемой величины X , с доверительной вероятностью P ($P=0,95$), находится в пределах: $[\bar{X} \mp t \cdot \sigma_{\bar{X}}]$ т.е. $[500,6 \pm 1,96 \cdot 0,8221]$ или $[498,9877; 502,2123]$.

Расчет доверительных границ суммарной неисключенной систематической составляющей погрешности измерения.

$$\theta(P) = K \cdot \sqrt{\sum_{j=1}^k \theta_j^2} = 1,1 \cdot 1 = 1,1.$$

В данном случае неисключенная систематическая погрешность измерения обусловлена одной составляющей $\theta_j = 1$.

$$\text{Вычислим соотношение } \frac{\theta(P)}{\sigma_{\bar{X}}} = \frac{1,1}{0,8221} = 1,338.$$

Полученное значение лежит в пределах $0,8 \div 8,0$, следовательно, ни одной из составляющих погрешности измерения пренебречь нельзя, тогда, погрешность результата будет содержать как случайную так и не случайную составляющие погрешности.

Доверительные границы общей погрешности измерения:

$$\sigma_{\Sigma} = \sqrt{\frac{1}{3} \theta(P)^2 + \sigma_{\bar{X}}^2} = \sqrt{\frac{1}{3} \cdot 1,1^2 + 0,8221^2} = 1,03884$$

$$k = \frac{\theta(P) + \varepsilon(P)}{\sigma_{\bar{x}} + \sqrt{\frac{1}{3}\theta(P)^2}} = \frac{\theta(P) + \varepsilon(P)}{\sigma_{\bar{x}} + \theta(P)/\sqrt{3}} = \frac{1,1 + 1,61132}{0,8221 + 1,1/\sqrt{3}} = \frac{2,71132}{1,45719} = 1,86$$

$$\Delta p = k \cdot \sigma_{\Sigma} = 1,86 \cdot 1,03884 = 1,93292 = 1,9.$$

Следовательно, результат измерений можно записать в виде:

$$X = \bar{X} \pm \Delta p \text{ или } 501,0 \pm 1,9.$$

2. ОЦЕНКА РЕЗУЛЬТАТОВ НЕРАВНОТОЧНЫХ ИЗМЕРЕНИЙ

Неравноточные измерения – ряд измерений какой-либо величины, выполненных различными по точности средствами измерений и (или) в разных условиях. Такие измерения нельзя обрабатывать как равноточные, так как информация об искомой величине будет неправильно интерпретирована. Отбрасывать результаты измерений, доверие к которым меньше, нельзя, так как в этом случае часть информации будет потеряно. Следовательно, необходимо использовать все измерения, с учетом их точности.

В случае неравноточных измерений вместо среднего арифметического значения и средней квадратической погрешности одного измерения используются среднее арифметическое значение и средняя квадратическая погрешность взвешенного среднего значения.

Значение измеренной величины наиболее близкое к ее истинному значению определяется по формуле:

$$X_0 = \frac{\bar{X}_1 P_1 + \bar{X}_2 P_2 + \dots + \bar{X}_m P_m}{P_1 + P_2 + \dots + P_m},$$

где X_0 - средневзвешенное значение, $\bar{X}_1, \dots, \bar{X}_m$ среднее значения группы измерений, P_1, \dots, P_m их вес или степень доверия к результатам измерения. Вес группы обратно пропорционален дисперсии, т.е.:

$$P_1 + \dots + P_m = \frac{1}{D_1} + \dots + \frac{1}{D_m}.$$

Окончательный результат: $X_0 \pm \varepsilon(P)$ или $X_0 \pm \Delta a$.

3. КОСВЕННЫЕ ИЗМЕРЕНИЯ

Очень часто при проведении физического эксперимента встречается ситуация, когда искомые величины $U(X_i)$ непосредственно определить невозможно, однако можно измерить величины X_i . Величины X_i , содержат, как обычно, случайные погрешности, т.е. наблюдают величины $X'_i = X_i \pm \Delta X_i$. На практике всегда стараются выбрать такой метод измерения и такую точность, чтобы погрешности результатов были малы по сравнению с их значениями. Иначе измерения теряют смысл или в лучшем случае их можно использовать лишь для установления наличия ожидаемого эффекта и определения порядка измеряемой величины. Будем считать, что X_i распределены по нормальному закону:

$$f(X'_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(X'_i - X_i)^2}{2\sigma^2}\right]$$

Для простоты будем считать погрешности равными.

Задачу о нахождении значений погрешностей косвенных измерений по измеренным значениям X_i можно решить с помощью метода наименьших квадратов. Однако в связи с тем, что сложность вычислений сравнительно велика, остановимся на упрощенной трактовке определения погрешностей косвенных измерений. Кроме того, будем рассматривать исследования физических величин, остающихся при проведении серии измерений в неизменном состоянии, например объема прямоугольной пластины по ее сторонам.

Пусть $U = f(X)$. Точное значение результата косвенных измерений обозначим $U_0 = f(\bar{X} \pm \Delta X)$, так как $X = \bar{X} + \Delta X$ и $\Delta X \ll X$, тогда $U = f(\bar{X})$, поэтому можно записать:

$$\Delta U_u = f(X) - f(\bar{X}).$$

Используем связь дифференциала функции df с бесконечно малым изменением аргумента

$$df(X) = \frac{df}{dx} dx$$

В этом случае абсолютная погрешность

Косвенного измерения может быть найдена из выражения:

$$\Delta U_u = df(X) = f(\bar{X} + \Delta X) - f(\bar{X}) = \frac{df}{dx}(\bar{X})\Delta X.$$

Относительная погрешность результата косвенных измерений:

$$\delta_u = \frac{\Delta U}{U} = \frac{f'(\bar{X})}{f(\bar{X})}\Delta X = d(\ln \bar{X}).$$

Найдем абсолютную и относительную погрешности, когда значение искомой величины определяется двумя значениями, измеряемыми в эксперименте.

В общем случае, разлагая $U = f(X, Y)$ в ряд Тейлора и оставляя только члены с нулевыми и первыми степенями ΔX и ΔY , получаем:

$$f(X, Y) = f(\bar{X} + \Delta X, \bar{Y} + \Delta Y) = f(\bar{X}, \bar{Y}) + \frac{df}{dx}\Delta X + \frac{df}{dy}\Delta Y.$$

Тогда

$$\begin{aligned} (\Delta U)^2 &= (\overline{U_0 - U})^2 = [f(X_0, Y_0) - f(\bar{X}, \bar{Y})]^2 = \\ &= \left(\frac{df(\bar{X}, \bar{Y})}{dx}\Delta X - \frac{df(X, Y)}{dy}\Delta Y \right)^2 \approx \left[\frac{df(\bar{X}, \bar{Y})}{dx} \right]^2 (\Delta X)^2 + \left[\frac{df(\bar{X}, \bar{Y})}{dy} \right]^2 (\Delta Y)^2 = \\ &= \left(\frac{df}{dx} \right)^2 (\Delta X)^2 + \left(\frac{df}{dy} \right)^2 (\Delta Y)^2. \end{aligned}$$

Таким образом, для любой функции от двух (или более) переменных средний квадрат отклонения

$$(\Delta U)^2 = \left(\frac{df}{dx} \right)^2 (\Delta X)^2 + \left(\frac{df}{dy} \right)^2 (\Delta Y)^2,$$

или

$$\Delta U = \sqrt{\left(\frac{df}{dx} \right)^2 (\Delta X)^2 + \left(\frac{df}{dy} \right)^2 (\Delta Y)^2}.$$

Все производные вычисляются для $X = \bar{X}$ и $Y = \bar{Y}$, а значения ΔX и ΔY вычисляются, как и для результатов прямых измерений (считаем все слагаемые положительными, чтобы быть уверенными в том, что погрешность функции не превзойдет $\pm dU$). Поэтому берем не просто сумму частных дифференциалов, а сумму их абсолютных значений, какие бы знаки

не получились у частных производных. Эту замену знаков производим после того, как сделаны все преобразования: приведены все члены с одним дифференциалом dx или dy в одну группу и получено их общее выражение. Относительная погрешность находится так же, как и для функции одного переменного, т. е.

$$\delta_u = \frac{\Delta U}{U} = \sqrt{\left(\frac{d \ln f}{dx}\right)^2 (\Delta X)^2 + \left(\frac{d \ln f}{dy}\right)^2 (\Delta Y)^2}.$$

Приведем несколько частных случаев зависимости $U = f(X, Y)$:

1. $U = X + Y$, тогда:

$$\Delta U = \sqrt{(\Delta X)^2 + (\Delta Y)^2} \text{ и } \delta_u = \frac{\Delta U}{U} = \frac{\sqrt{(\Delta X)^2 + (\Delta Y)^2}}{\bar{X} + \bar{Y}}.$$

2. $U = X - Y$, тогда

$$\Delta U = \sqrt{(\Delta X)^2 + (\Delta Y)^2} \text{ и } \delta_u = \frac{\Delta U}{U} = \frac{\sqrt{(\Delta X)^2 + (\Delta Y)^2}}{\bar{X} - \bar{Y}}.$$

Сравнивая выражения два последних выражения, видим, что при одних и тех же погрешностях ΔX и ΔY относительная погрешность разности может быть значительно больше относительной погрешности суммы, если значения X и Y близки. Следовательно, нельзя добиться хорошей точности измерений величины, если она находится как небольшая разность результатов независимых измерений двух величин, существенно превышающих искомую.

3. $U = XY$, тогда:

$$\Delta U = \sqrt{\bar{Y}^2 (\Delta X)^2 + \bar{X}^2 (\Delta Y)^2} \text{ и } \delta_u = \sqrt{\delta_X^2 + \delta_Y^2} = \sqrt{\left(\frac{\Delta X}{\bar{X}}\right)^2 + \left(\frac{\Delta Y}{\bar{Y}}\right)^2}.$$

4. $U = X/Y$, тогда:

$$\Delta U = \sqrt{\frac{(\Delta X)^2}{\bar{Y}^2} + \frac{\bar{X}^2 (\Delta Y)^2}{\bar{Y}^4}} \text{ и } \delta_u = \sqrt{\delta_X^2 + \delta_Y^2} = \sqrt{\left(\frac{\Delta X}{\bar{X}}\right)^2 + \left(\frac{\Delta Y}{\bar{Y}}\right)^2}.$$

В табл. 4.6 приведены формулы для вычисления абсолютных и относительных погрешностей некоторых часто встречающихся на практике функций.

Способы нахождения погрешностей косвенных измерений тем точнее, чем больше число измерений. При наиболее ответственных измерениях необходимо использовать сложные методы математической статистики, например проверку гипотез, дисперсионный анализ и т. п.

Т а б л и ц а 3.1

Формулы для вычисления погрешностей функции

№ п/п	Функция	Абсолютная погрешность	Относительная погрешность
1.	X^k	$k\bar{X}^{k-1}\Delta X$	$k \frac{\Delta X}{\bar{X}}$
2.	$\sqrt[k]{X}$	$\frac{1}{k} \bar{X}^{\frac{1}{k}-1} \cdot \Delta X$	$\frac{1}{k} \cdot \frac{\Delta X}{\bar{X}}$
3.	$\frac{X}{1 \pm X}$	$\frac{\Delta X}{\bar{X}(1 \pm \bar{X})^2}$	$\frac{\Delta X}{\bar{X}(1 \pm \bar{X})}$
4.	e^X	$e^{\bar{X}} \Delta X$	ΔX
5.	$\ln X$	$\frac{\Delta X}{\bar{X}}$	$\frac{\Delta X}{\bar{X} \ln \bar{X}}$
6.	$\sin X$	$\cos \bar{X} \cdot \Delta X$	$ctg \bar{X} \cdot \Delta X$
7.	$\cos X$	$\sin \bar{X} \cdot \Delta X$	$tg \bar{X} \cdot \Delta X$
8.	$tg X$	$\frac{\Delta X}{\cos^2 \bar{X}}$	$\frac{2 \cdot \Delta X}{\sin 2\bar{X}}$
9.	$ctg X$	$\frac{\Delta X}{\sin^2 \bar{X}}$	$\frac{2 \cdot \Delta X}{\sin 2\bar{X}}$

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. ГОСТ Р 8.736-2011 Государственная система обеспечения единства измерений. Измерения прямые многократные. Методы обработки результатов измерений. Основные положения. - М.: Изд-во Стандартиформ, 2013- 20 с.

2. Математическое моделирование динамической прочности конструкционных материалов: Учебное пособие. - М.: Изд-во АСВ, 2013- 562 с. - ISBN 978-5-93093-981-1.

3. Метрология, стандартизация, подтверждение соответствия: учебное пособие / Н.Б. Камардин, И.Ю. Суркова. - Казань : Изд-во КНИТУ, 2013. - -240 с. - ISBN 978-5-7882-1401-6

4. . Метрология, стандартизация и сертификация: Учеб. для вузов / Я.М. Радкевич, А.Г. Схиртладзе, Б.И. Лактионов. - М. : Абрис, 2012. - - 791 с.: ил. ISBN 978-5-4372-0064-3.

5. Романов В. Н. Ромодановская М. П. Прикладная метрология: учебное пособие для вузов . Владим. гос.ун-т им Изд-о ВлГУ, 2014. – 188 с SBN 978-5-9984-0488-7

Приложение 1.

Значения функции Лапласа

t	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,008	0,0120	0,0160	0,019	0,023	0,027	0,031	0,0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0,7	2580	2611	2642	2673	2703	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2,0	4772	4778	4783	4788	4793	4798	4803	4808	4813	4817
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2,2	4861	4864	4868	4871	4874	4878	4881	4884	4887	4890
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2,5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2,6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2,7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2,8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2,9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4886
3,0	4986									
3,5	4998									
4,0	4999									

Приложение 2

Значения распределения Стьюдента

n	Доверительная вероятность P				
	0,90	0,95	0,98	0,99	0,999
2	6,31	12,71	31,82	63,68	636,62
3	2,92	4,30	6,97	9,93	31,60
4	2,35	3,18	4,54	5,84	12,92
5	2,13	2,78	3,75	4,60	8,61
6	2,02	2,57	3,37	4,06	6,87
7	1,94	2,45	3,14	3,71	5,96
8	1,90	2,37	3,00	3,50	5,41
9	1,86	2,31	2,90	3,36	5,04
10	1,83	2,26	2,82	3,25	4,78
11	1,81	2,23	2,76	3,17	4,59
12	1,80	2,20	2,72	3,11	4,44
13	1,78	2,18	2,68	3,06	4,32
14	1,77	2,16	2,65	3,01	4,22
15	1,76	2,15	2,62	2,98	4,14
16	1,75	2,13	2,60	2,95	4,07
17	1,75	2,12	2,58	2,92	4,02
18	1,74	2,11	2,57	2,90	3,97
19	1,73	2,10	2,55	2,88	3,92
20	1,73	2,09	2,54	2,88	3,88
∞	1,65	1,98	2,33	2,58	3,29

Приложение 3

Значения критерия Фишера для различных уровней значимости

k_2	F , при k_1 , равном									
	1	2	3	4	5	6	8	12	16	00
$\alpha = 0,05$										
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,43	19,50
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,84	5,63
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,92	3,67
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,20	2,93
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,82	2,54
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,60	2,30
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,44	2,13
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,33	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,25	1,92
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,18	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,99	1,62
00	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,64	1,00
$\alpha = 0,01$										
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,42	99,44	99,50
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	14,15	13,46
6	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,52	6,88
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,67	5,48	4,86
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,52	3,91
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,98	3,36
14	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,62	3,00
16	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,37	2,75
18	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,20	2,57
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	3,05	2,42
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,66	2,01
00	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,99	1,00

Примечание: k_1 — число степеней свободы большей дисперсии; k_2 — число степеней свободы меньшей дисперсии.