КЫРГЫЗСКО-РОССИЙСКИЙ СЛАВЯНСКИЙ УНИВЕРСИТЕТ

ЕСТЕСТВЕННО-ТЕХНИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра физики и микроэлектроники

В.С.Энгельшт, Б.Б.Чен

СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Методическое пособие

Издательство Кыргызско-Российского
Славянского университета
Бишкек 2002

Энгельшт В.С., Чен Б.Б.

СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА: Методическое пособие /Издательство Кыргызско-Российского Славянского университета. – Бишкек, 2002. – 46 с.

Приведены классификации погрешностей измерений и методы обработки результатов прямых и косвенных измерений. Изложены правила оценки истинного значения измеряемой физической величины. Приведены примеры обработки результатов работ общего физического практикума. В приложении приводятся таблицы наиболее часто встречающихся расчетных функций.

Предназначено для студентов всех специальностей, изучающих физику при выполнении физического практикума, а также может быть использовано выпускниками и магистрами при выполнении текущих исследований и оформлении дипломных работ и магистерских диссертаций.

Рецензент: профессор О.А.Подрезов

Печатается по решению
кафедры физики и микроэлектроники и РИСО КРСУ

СОДЕРЖАНИЕ

Стр.

1. КЛАССИФИКАЦИЯ ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЯ

Погрешностью измерения называется разность $\xi = y - a$ между результатом измерения y и истинным значением a измеряемой величины. Обычно a и ξ неизвестны. Одной из основных задач математической обработки результатов эксперимента является оценка приближения y к истинному значению a с возможно меньшей погрешностью. Для решения этой задачи нужно знать основные свойства погрешности измерений и уметь ими воспользоваться.

Погрешности измерений бывают трех родов.

Грубые погрешности (или ***промахи***) – это погрешности, обусловленные недосмотром экспериментатора. Внешним признаком промаха является его резкое отличие по величине от результатов остальных измерений. Эти погрешности эффективно бракуются в процессе измерений, либо при математической обработке результатов измерений.

Систематические погрешности – это погрешности, которые дают отклонения измеряемой величины либо с занижением, либо с завышением, т.е. действуют всегда в одну сторону. Причинами систематических погрешностей могут быть неправильная градуировка прибора, либо изменение внешних воздействий. Например, в спектральном анализе основной причиной является несоответствие эталонов и проб по валовому и минералогическому составу (по виду соединений). Систематические погрешности можно выявить и устранить с помощью образцовых приборов, использования правильных эталонов и т.д.

Случайные погрешности – это неустранимые погрешности измерений, обусловленные суммарным действием большого количества различных факторов. В спектральном анализе это могут быть: неравномерность проб, колебания тока, неоднородности фотопластин и т.п. И тем не менее влияние случайных погрешностей на оценку истинного значения измеряемой величины все-таки можно учесть с помощью методов теории вероятностей, что позволяет определить значение a со значительно меньшей погрешностью, чем погрешности отдельных измерений y. Для этого используются законы распределения случайных погрешностей измерений.

2. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ СЛУЧАЙНЫХ ПОГРЕШНОСТЕЙ

Наиболее часто реализуется нормальный закон распределения случайных погрешностей – закон Гаусса. Хотя возможны и другие законы распределения (Пуассона, равномерного распределения и т.д.). Существование закона можно обнаружить при многократном повторении измерений некоторой величины и неизменном режиме. Подсчитывают число m тех результатов, которые попадают в любой выделенный интервал. Отношение этого числа к общему числу n произведенных измерений (относительная частота попадания в отмеченный интервал m/n) при достаточно большом n оказывается постоянным (разумеется, своим для каждого интервала), называемым *вероятностью попадания случайной величины в этот интервал* и обозначаемым $P(\xi)$. Правило, позволяющее находить $P(\xi)$ для любых интервалов, называется *законом распределения вероятностей случайной величины*. Обычно законы распределения записываются в интегральной форме:

$$P(\xi) = \int p(\xi) d\xi ,$$

где $p(\xi)$ – неотрицательная функция , удовлетворяющая условию

$$\int_{-\infty}^{\infty} p(\xi) d\xi = 1 \cdot$$

Эта функция полностью определяет соответствующий закон распределения вероятностей и называется *плотностью распределения*. Для нормального закона

$$p(\xi) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\xi^2/(2\sigma^2)}, \qquad (1)$$

где σ – параметр, характеризующий точность измерений. Его называют средней квадратической погрешностью или стандартом. На рис.1 показаны кривые нормального закона распределения при различных значениях σ. Все кривые положительны и симметричны относительно оси ординат. Половина максимальной плотности распределения соответствует значению σ. При 3σ кривая асимптотически стремится к нулю (см. рис.2).

Вероятность попадания погрешности в интервал [ξ_1, ξ_2] графически изображается площадью соответствующей криволинейной трапеции под кривой Гаусса (рис.2). В частности, вероятность попадания погрешности в симметричный интервал [ξ_1, ξ_2] равна:

$$\int_{-\xi_1}^{\xi_1} p(\xi) d\xi = \frac{2}{\sqrt{2\pi}} \int_0^{\xi_1} e^{-\xi^2/(2\sigma^2)} \frac{d\xi}{\sigma} = \frac{2}{\sqrt{2\pi}} \int_0^{t_\sigma} e^{-t^2/2} dt = 2\Phi(t), \quad t = \xi_1/\sigma \cdot$$

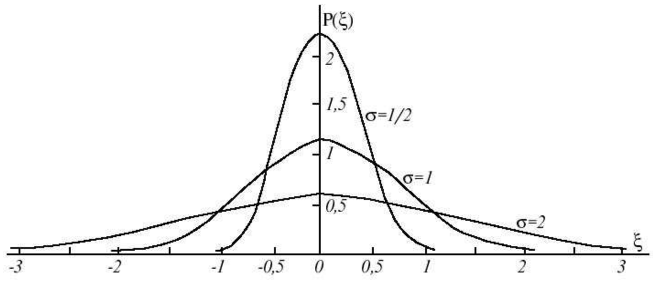


Рис.1.

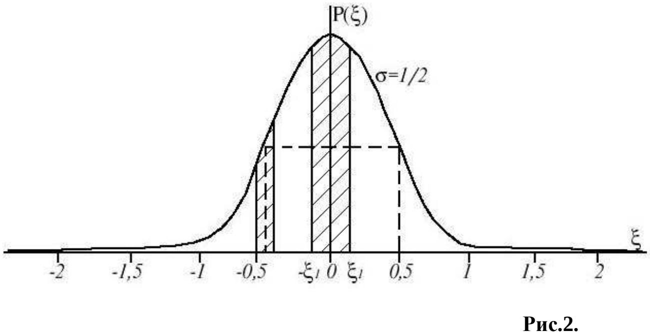


Рис.2.

Другими словами, в эксперименте на интервале [ξ_1, ξ_2] могут встретиться те или иные погрешности с вероятностью $2\Phi(t)$. Функция $\Phi(t)$ называется интегралом вероятности, ее значения обычно табулируются. Вероятность попадания случайной погрешности в любой интервал [ξ_1, ξ_2] для нормального распределения вычисляется по формуле

$$p(\xi) = \Phi(\xi_2/\sigma) - \Phi(\xi_1/\sigma).$$

Вероятность того, что случайная погрешность выйдет за границы $\pm\sigma$ (t>0) равна

$$p(|\xi| > t\sigma) = 1 - 2\Phi(t), \quad t = \xi/\sigma.$$

Если обратиться к фрагменту таблицы интеграла вероятности ([I], стр.171-172, табл.I, П), то можно сказать, что около 70% измерений имеют погрешность $/\xi/\leq\sigma$, около 95% имеют погрешность $/\xi/\leq 2\sigma$.

Таблица 1

$t=\xi/\sigma$	1	2	3
$2\Phi(t)$	0,683	0,954	0,9973
$1-2\Phi(t$	0,317	0,046	0,0027

Лишь 5% (5 измерений из 100) могут иметь погрешность $/\xi/>2\sigma$ и только 0,3% (3 измерения из 1000) – погрешность $/\xi/>3\sigma$. Это говорит о том, что вероятность выхода за 3σ настолько мала, что выход считают практически невозможным (критерий трех сигм), хотя рассматриваемая математическая модель допускает в принципе любые значения погрешностей.

3. ЭМПИРИЧЕСКИЙ СТАНДАРТ

Величина σ, характеризующая закон нормального распределения Гаусса – это истинный стандарт, он может быть найден только при бесконечно большом числе измерений или, как говорят, из генеральной совокупности. На практике имеем ограниченное число измерений – n, и по этим измерениям определяем не истинный стандарт, а эмпирический стандарт s – точечная оценка дисперсии случайной погрешности:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2, \qquad (2)$$

где $\bar{y}$ – среднее арифметическое значение $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

Если проводится n-кратное измерение какой-то величины, то погрешность среднего арифметического $\bar{y}$ будет естественно меньше, чем погрешность единичного измерения y и оказывается, что средняя квадратическая погрешность $\bar{s}$ среднего значения $\bar{y}$ при равноточных измерениях (с одной и той же дисперсией) равна

$$\bar{s} = \frac{s}{\sqrt{n}},$$

где s – средняя квадратическая погрешность однократного измерения. Эмпирический стандарт тем ближе к истинному, чем больше n и, наоборот, при малых n отличие s от σ может быть значительным. Для получения s порядка σ необходимо, чтобы среди погрешностей был полный набор по кривой Гаусса (см. выше процентное соотношение погрешностей, соответствующих (σ, 2σ, 3σ). В действительности могут оказаться случаи набора только малых погрешностей (s будет занижена в сравнении с σ) или только больших погрешностей (s будет завышена в сравнения с σ). В связи с этим не указывается точное значение, а дается доверительный интервал:

$$z_1(P)s < \sigma < z_2(P)s.$$

Таблица 2

n	5	10	25	∞
z_1	0,624	0,699	0,785	1
z_2	2,453	1,755	1,381	1

Через P обозначена *доверительная вероятность*. Она характеризует надежность результатов с заданной вероятностью в зависимости от числа измерений n. Например, при P=95% (табл.2), где n – число измерений, по которому определяется эмпирический стандарт s. При очень большом числе измерений $s\approx\sigma$ и $z_1=z_2=1$. При n=5 имеем $0{,}6s<\sigma<2{,}5s$, т.е. истинная погрешность лежит в интервале от 0,6s до 2,5s. Если сделали 10 измерений, то этот доверительный интервал сужается, а при n=25 (это уже достаточно большое число измерений) он достигает размера $0{,}8s<\sigma<1{,}4s$. Это означает, что при погрешности s, допустим интервал равный 10%, истинная погрешность лежит в интервале от 8 до 14%.

Итак, для характеристики измерений следует указывать эмпирический стандарт s, число измерений n и доверительный интервал для σ.

4. ОЦЕНКА ОТКЛОНЕНИЯ ИСТИННОГО ЗНАЧЕНИЯ ФИЗИЧЕСКОЙ ВЕЛИЧИНЫ ОТ СРЕДНЕГО АРИФМЕТИЧЕСКОГО ЗНАЧЕНИЯ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

4.1. Правило трёх сигм

Отклонение истинного значения *a* измеряемой величины от среднего арифметического значения $\bar{y}$ результатов измерений не превосходит утроенной средней квадратической погрешности этого среднего значения. При известном истинном стандарте σ

$$\left|a-\bar{y}\right|<\frac{3\sigma}{\sqrt{n}} \text{ или } \bar{y}-\frac{3\sigma}{\sqrt{n}}<a<\bar{y}+\frac{3\sigma}{\sqrt{n}}. \qquad (3)$$

В случае, когда известен лишь эмпирический стандарт *s*, то

$$\left|a-\bar{y}\right|<\frac{3s}{\sqrt{n}} \text{ или } \bar{y}-\frac{3s}{\sqrt{n}}<a<\bar{y}+\frac{3s}{\sqrt{n}}, \qquad (4)$$

где (σ и *s* – средние квадратические погрешности однократного измерения, $\sigma/\sqrt{n}$ и $s/\sqrt{n}$ – средние квадратические погрешности среднего арифметического из *n* измерений.

Первая из оценок (3) имеет надежность *P=2Ф(3)=0,9973*, независимо от числа измерений *n*. Надежность второй оценки (4) зависит существенно от *n*, см. табл.3 ([I], стр.31, табл.2.2-1).

Таблица 3

n	*5*	*10*	*25*	*∞*
P	0,96	0,985	0,994	0,9973

4.2. Критерий Стьюдента

Для оценки истинного значения *a* через среднее арифметическое $\bar{y}$ из *n* измерений можно пользоваться критерием Стьюдента, согласно которому

$$\left|a-\bar{y}\right|<t(P,n)\frac{s}{\sqrt{n}},$$

где $t(P,n)$ – множитель, зависящий от заданной априори доверительной вероятности *P* и числа измерений *n*. Значения $t(P,n)$ для распределения Стьюдента даны в таблицах. ([1], стр.174, табл.1У). Чтобы воспользоваться этим критерием требуется априори задать *P*, в отношении которого нет никаких объективных рекомендаций. Произвол в выборе *P* может повлечь существенные искажения в оценке результатов. Важно отметить, что если для критерия Стьюдента принять *P*, как это следует из правила трёх сигм (см.табл.3), то оба критерия дают практически одинаковые доверительные интервалы А именно, вместо множителя 3 по трёхсигмовому критерию (4) в критерии Стьюдента (5) будет множитель $t(P,n)$=3,06 независимо от *n*. В связи с изложенным, рекомендуется отдать предпочтение трёхсигмовому критерию, где *P* определяется апостериори по числу измерений (табл.3).

5. ОЦЕНКА ИСТИННОГО ЗНАЧЕНИЯ С ПОМОЩЬЮ ИНТЕРВАЛОВ

Смысл оценки с помощью интервалов заключается в нахождении интервалов, называемых доверительными, между границами которых с определенными вероятностями (доверительными) находятся истинные значения оцениваемых величин.

5.1. Определение доверительного интервала для *среднего арифметического значения* измеряемой величины

Пример. По результатам 5 наблюдений была найдена длина стержня. Итог измерений составляет *L*=15,785 мм; $S_{\bar{L}}=0{,}005$ мм, при этом существуют достаточно обоснованные предположения о том, что распределение результатов наблюдений было нормальным.

а) Требуется оценить вероятность того, что истинное значение длины стержня отличается от среднего арифметического из 5 наблюдений не больше, чем на 0,01 мм.

Вычисляем значение дроби Стьюдента:

$$t_p = \frac{0{,}01}{S_{\bar{L}}} = \frac{0{,}01}{0{,}005} = 2 \qquad (5)$$

и число степеней свободы:

$$k = n - 1 = 5 - 1 = 4. \qquad (6)$$

По данным табл. 1 Приложения находим значение доверительной вероятности для $t_p=2$ и $k=4$. Она равна 88,38%.

Итог измерений записывается в виде:

$$L=15{,}785 \pm 0{,}01\ мм; \quad P=88{,}38\%. \qquad (7)$$

б) В условиях предыдущей задачи найти *доверительную границу погрешности результата измерений* для доверительной вероятности P=99,0%.

По данным табл. 2 Приложения при $k=4$ находим $t_p=4{,}604$. Следовательно, доверительная граница составляет:

$$\delta_{99,0\%} = t_{99,0\%} \cdot S_{\bar{L}} = 4{,}604 \cdot 0{,}005 = 0{,}023\,(\text{мм}). \qquad (8)$$

Итог измерений:

$$L=15{,}785 \pm 0{,}023\ мм; \quad P=99{,}0\%. \qquad (9)$$

5.2. *Определение доверительных интервалов для дисперсии и среднего квадратического отклонения результатов наблюдений*

Пример. Проведены 20 измерений длины l_i, *мм*. В качестве оценки математического ожидания длины принимаем среднее арифметическое

$$\bar{l} = \frac{1}{20}\sum_{i=1}^{20} l_i = 10{,}3078\ \text{мм}.$$

Точечная оценка среднего квадратического отклонения результатов наблюдений составляет:

$$S_l = \sqrt{\frac{1}{20-1}\sum_{i=1}^{20}\left(l_i - \bar{l}\right)^2} = 0{,}0025\ \text{мм}.$$

Приняв уровень доверительной вероятности $\alpha=1-q=90\%=0{,}90$, находим для числа степеней свободы $k=n-1=20-1=19$ в табл. 3 Приложения:

$$\chi^2_{k;\frac{1}{2}q} = \chi^2_{19;0,05} = 10{,}117; \quad \chi_{19;0,05} = 3{,}18, \qquad (10)$$

$$\chi^2_{k;1-\frac{1}{2}q} = \chi^2_{19;0,95} = 30{,}144; \quad \chi_{19;0,95} = 5{,}49. \qquad (11)$$

Границы доверительного интервала для среднего квадратического отклонения результатов наблюдений находим по формуле:

$$\left.\begin{aligned} S_{l_1} &= \frac{\sqrt{n-1}\cdot S_l}{\chi_{k;\frac{1}{2}q}} = \frac{\sqrt{20-1}\cdot 0{,}0025}{3{,}18} = 0{,}0034\ (мм) \\ S_{l_2} &= \frac{\sqrt{n-1}\cdot S_l}{\chi_{k;1-\frac{1}{2}q}} = \frac{\sqrt{20-1}\cdot 0{,}0025}{5{,}49} = 0{,}0020\ (мм) \end{aligned}\right\} \qquad (12)$$

Следовательно, истинное значение среднего квадратического отклонения результатов наблюдений с вероятностью 90% лежит в интервале 0,0020-0,0034 мм.

Замечание: При $k>30$ можно пользоваться приближенной формулой:

$$\chi_{kp} = \sqrt{k - \frac{1}{2} + \frac{1}{\sqrt{2}}\cdot t_p},$$

где t_p определяется из условия $\Phi(t_p)=P$ по табл. 4 Приложения, в которой помещены значения интегральной функции нормированного нормального распределения.

Тогда границы доверительного интервала для среднего квадратического отклонения результатов наблюдений при доверительной вероятности $\alpha=1-q$ вычисляются по формуле (3.12) при значениях χ_k, равных:

$$\begin{aligned} \chi_{k;\frac{1}{2}q} &= \sqrt{k - \frac{1}{2} + \frac{1}{\sqrt{2}}\cdot t_{\frac{1}{2}q}}, \\ \chi_{k;1-\frac{1}{2}q} &= \sqrt{k - \frac{1}{2} + \frac{1}{\sqrt{2}}\cdot t_{1-\frac{1}{2}q}}. \end{aligned} \qquad (13)$$

5.3. *Проверка нормальности распределения результатов наблюдений*

Если число наблюдений меньше 40 (n<40), то для проверки нормальности можно воспользоваться понятием *статистической функции распределения результатов наблюдений.*

Для ее построения полученные в ходе эксперимента результаты группируют в так называемый *вариационный ряд* $x_1, x_2, \ldots x_n$, члены которого располагаются в порядке их возрастания, т.е. так, что всегда $x_1 \le x_2 \le \ldots \le x_n$.

Статистическую формулу распределения $F_n(x_k)$ определяют по формуле:

$$F_n(x_k) = \frac{k}{n+1}; \quad k = 1, 2, ..., n. \tag{14}$$

Для проверки нормальности распределения результатов наблюдений по табл. 4 Приложения находят z_k, соответствующие полученным значениям $F_n(x_k)$ статистической функции распределения: $\Phi(z_k) = F_n(x_k)$.

Но переменная z определяется через результаты наблюдений как $z_k = \dfrac{x_k - m_x}{\sigma_x}$, и если в координатах (z_k, x_k) нанесены точки z_k,; x_k , то при нормальном распределении они должны располагаться вдоль одной прямой линии. Если получится некоторая кривая, то гипотезу о нормальности распределения придется отвергнуть, как противоречащую опытным данным.

Пример. Проведено 19 измерений длины l_i в мм. Проверить нормальность распределения результатов наблюдений:

18,305	18,304	18,308	18,305	18,309	18,309
18,308	18,306	18,306	18,307	18,308	18,310
18,311	18,310	18,312	18,308	18,307	
18,309	18,303				

Вычисления по изложенной методике сведены в следующую таблицу.

x_k, мм	$F_n(x_k)=\Phi(z_k)$	z_k	x_k, мм	$F_n(x_k)=\Phi(z_k)$	z_k
18,303	0,05	-1,6449	18,308	0,60	0,2533
18,304	0,10	-1,2816	18,309	0,75	0,6745
18,305	0,20	-0,8416	18,310	0,85	1,0364
18,306	0,30	-0,2533	18,311	0,90	1,2816
18,307	0,40	-0,2533	18,312	0,95	1,6449

Затем строится зависимость $z_k(x_k)$. Если точки располагаются очень близко к прямой (в нашем примере так), то распределение результатов наблюдений можно считать нормальным.

При большом числе результатов наблюдений (n>40) задача проверки нормальности распределения результатов наблюдений решается в следующем порядке.

Весь диапазон полученных результатов наблюдений $X_{мах} - X_{min}$ разбивают на r интервалов шириной ΔX_i $(i=1,2,\ldots, r)$ и подсчитывают *частоты* m_i, равные числу результатов, лежащих в каждом *i-м* интервале, то есть меньших или равных его правой и больших левой границы.

Отношения

$$P_i^* = \frac{m_i}{n}, \tag{15}$$

где n- общее число наблюдений, называется *частостями* и представляет собой статистические оценки вероятностей попадания результата наблюдений в *i-й* интервал.

Распределение частостей по интервалам образует *статистическое распределение* результатов наблюдений. Если теперь разделить частость на длину интервала, то получим величины

$$p_i^* = \frac{1}{\Delta X_i} P_i^* = \frac{m_i}{n \Delta X_i}, \tag{16}$$

которые являются ***оценками*** средней плотности распределения в интервале ΔX_i.

Теперь отложим вдоль оси X интервалы ΔX_i в порядке возрастания индекса i и на каждом интервале построим прямоугольник с высотой, равной p_i^*. Полученный график называется *гистограммой статистического распределения*.

Площадь всех прямоугольников равна единице:

$$\sum_{i=1}^{r} p_i^* \Delta X_i = \sum_{i=1}^{r} \frac{m_i}{n} = \frac{1}{n} \sum_{i=1}^{r} m_i = 1. \tag{17}$$

При увеличении числа наблюдений число интервалов можно увеличить. Сами интервалы уменьшаются, а гистограмма все больше приближается к плавной кривой, ограничивающей единичную площадь, – к графику плотности распределения результатов наблюдений.

При построении гистограмм рекомендуется пользоваться следующими правилами.

1. Число интервалов r выбирается в зависимости от числа наблюдений согласно следующим рекомендациям Всероссийского научно-исследовательского института метрологии (ВНИИМ):

n = 40-100 r= 7-9;
n = 100-500 r= 8-12;
n = 500-1000 r= 10-16;
n = 1000-10000 r= 12-22.

2. Длины интервалов удобнее выбирать одинаковыми. Но если распределение крайне неравномерно, то в области максимальной концентрации результатов наблюдений следует выбирать более узкие интервалы.

3. Масштабы по осям гистограммы должны быть такими, чтобы отношение ее высоты к основанию составляло примерно 5:8.

Пример. Было выполнено 100 измерений среднего диаметра изделия. Результаты наблюдений лежат в диапазоне 8,911-8,927 мм, то есть зона распределения резервов составляет 0,016 мм. Весь диапазон удобно разделить на 8 равных интервалов через 0,002 мм. В таблице приведены частоты m_i, частости P_i^* и плотности p_i^* статистического распределения:

i	X_i, *мм*	X_{i+1}, *мм*	m_i	P_i^*	p_i^*, *мм*$^{-1}$
1	8,911	8,913	1	0,01	5
2	8,913	8,915	5	0,05	25
3	8,915	8,917	14	0,14	70
4	8,917	8,919	27	0,27	135
5	8,919	8,921	24	0,24	120
6	8,921	8,923	18	0,18	90
7	8,923	8,925	9	0,09	45
8	8,925	8,927	2	0,02	10

$\overline{X}$ =8,91936 мм, S_X=0,0028 мм.

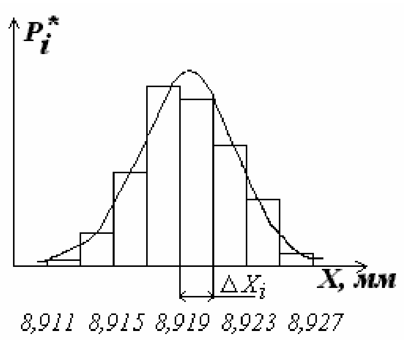


Рис.3

Построенная по данным этой таблицы гистограмма изображена на рис.3.

После построения гистограммы надо подобрать теоретическую плавную кривую распределения, которая, выражая все существенные черты статистического распределения, сглаживала бы все случайности, связанные с недостаточным объемом экспериментальных данных.

Принципиальный вид теоретической кривой выбирают заранее, проанализировав метод измерения или, хотя бы, по внешнему виду гистограммы. Тогда определение аналитического вида кривой распределения сводится к выбору таких значений его параметров, при которых достигается наибольшее соответствие между теоретическим и статистическим распределением.

Одним из методов решения этой задачи является *метод моментов*. При его использовании параметрам теоретического распределения придают такие значения, при которых несколько важнейших моментов совпадают с их статистическими оценками. Так, если статистическое распределение, определяемое гистограммой рисунка, мы хотим описать кривой нормального распределения, то естественно потребовать, чтобы математическое ожидание и дисперсия последнего совпадали со средним арифметическим и оценкой дисперсий, вычисленным по опытным данным.

В предыдущем примере $\overline{X}$ =8,91936 мм, S_X=0,0028 мм и уравнение кривой нормального распределения, лучше всего согласующегося со статистическим распределением, должно иметь вид:

$$p_X(x) = \frac{1}{0{,}0028\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-8{,}91936}{0{,}0028}\right)^2} . \qquad (18)$$

Возникает законный вопрос, объясняются ли расхождения между гистограммой и подобранным теоретическим распределением только случайными обстоятельствами, связанными с ограниченным числом наблюдений, или они вызваны тем, что результаты наблюдений в действительности распределены иначе.

Для ответа на этот вопрос используют методы проверки *статистических гипотез*.

Одним из таких методов является <u>*критерий согласия*</u> χ^2.

Проверка нормальности распределения согласно критерию χ^2 сводится к следующему.

1. Данные наблюдений группируют по интервалам, как при построении гистограммы. Подсчитывают частоту m_i. Если в некоторые интервалы попадает меньше 5 наблюдений, то такие интервалы соединяют с соседними. При этом число степеней свободы k, конечно, уменьшается.

2. Вычисляют среднее арифметическое $\overline{X}$ и точечную оценку среднего квадратического отклонения результата наблюдений S_X, которые принимают в качестве параметров теоретического нормального распределения с плотностью $p_x(x)$.

3. Для каждого интервала находят вероятности попадания в них результатов наблюдений либо по общей формуле:

$$p\{x_1 < X \le x_2\} = \Phi(t_2) - \Phi(t_1) = \Phi\left(\frac{x_2 - Q}{\sigma_x}\right) - \Phi\left(\frac{x_1 - Q}{\sigma_x}\right),$$

либо приближенно как произведение плотности теоретического распределения в середине интервала на его длину:

$$p_i \approx p_x\left(\frac{X_i + X_i + 1}{2}\right) \cdot \Delta X_i .$$

4. Для каждого интервала вычисляют величины $\chi_i^2 = \frac{(m_i - nP_i)^2}{nP_i}$ (i=1, 2, …, r) и суммируют их по всем i, в результате получают меру расхождения $\chi_k^2 = \sum\limits_{i=1} \chi_i^2$.

5. Определяют число степеней свободы $k=r-3$. Задаваясь уровнем значимости $q=1-\alpha$, находят по табл.3 Приложения значения $\chi^2_{k;\frac{1}{2}q}$ и $\chi^2_{k;1-\frac{1}{2}q}$. Если $\chi^2_{k;\frac{1}{2}q} < \chi_k^2 \le \chi^2_{k;1-\frac{1}{2}q}$, то распределение результатов считают нормальным.

<u>Пример.</u> Проверить нормальность распределения приведенных данных в предыдущей таблице, для которых $\overline{X}$ =8,91936 мм, S_X=0,0028 мм. Вычисления удобно свести в следующую таблицу:

Интервал составляет ΔX_i=0,002 мм. Плотности нормированного нормального распределения $p(t_i)$ взяты из табл.5 Приложения. Число степеней свободы k=8-2-3=3, так как 4 интервала были объединены в 2.

Задаваясь уровнем значимости q=0,10, находим в табл.3 Приложения $\chi^2_{3;0,05}$=0,352; $\chi^2_{3;0,95}$=7,815. Следовательно, распределение опытных данных можно считать нормальным.

6. ИСКЛЮЧЕНИЕ ГРУБЫХ ПОГРЕШНОСТЕЙ

Исключение грубых погрешностей наиболее эффективно проводить в процессе измерений. Если, однако, это не сделано, то они отбраковываются по трёх сигмовому критерию. Полагают, что результат измерений y_x – брак, если

$$\left|y_x - \bar{y}\right| > \frac{3\sigma}{\sqrt{n}} = 3\bar{s} \qquad (19)$$

с вероятностью P из табл.3.

7. АППРОКСИМАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ. ПРОВЕДЕНИЕ КРИВЫХ ПО ЭКСПЕРИМЕНТАЛЬНЫМ ТОЧКАМ

Функциональные зависимости из эксперимента получают в виде набора дискретных точек, которые затем пытаются представить в аналитической форме, чтобы иметь непрерывную функцию более или менее точно описывающую истинное физическое явление. Не следует стремиться к тому, чтобы приближение точно совпадало с табличными данными, ибо они искажены погрешностями эксперимента. Здесь важнее найти эмпирические формулы, которые сглаживают случайные погрешности и наиболее близки к истинным значениям изучаемой зависимости. Опытные экспериментаторы интуитивно проводят плавные линии по точкам правильно. Математически это позволяет сделать метод наименьших квадратов (мнк). Преимущество математической обработки результатов эксперимента заключается в возможности получать дополнительную информацию о результатах кроме графика. Часто вид функции приближения известен заранее (говорят: "будем строить приближение в виде прямой, или в виде квадратичной параболы" или "воспользуемся экспоненциальной зависимостью" и т.д.), следует определить лишь параметры, входящие в эту функцию приближения.

Рассмотрим простейшую аппроксимацию экспериментальных данных прямой линией. Требуется группу точек $\{x_i, y_i\}, i=\overline{1,n}$, полученных в эксперименте, описать прямой линией (рис.4)

$$f(x)=a+bx,$$

где a и b – постоянные параметры и геометрический смысл их таков: a показывает отрезок, отсекаемый от оси ординат, b – тангенс угла α наклона графика прямой к оси x.

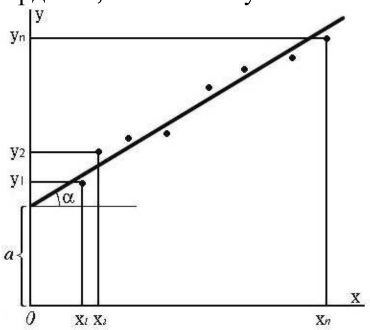


Рис.4.

Проведение прямой в соответствии с мнк означает следующее: нужно найти параметры a и b из условия минимума функционала $F=\sum_{i=1}^{n}(f_i - y_i)^2 = \min$ по параметрам a и b. Для этого необходимо решить систему двух уравнений $\frac{\partial F}{\partial a}=0,\ \frac{\partial F}{\partial b}=0.$ В результате решения получаем

$$a=\frac{\overline{x^2}\,\bar{y}-\bar{x}\,\overline{xy}}{\overline{x^2}-(\bar{x})^2}, \qquad b=-\frac{\bar{x}\,\bar{y}-\overline{xy}}{\overline{x^2}-(\bar{x})^2}, \qquad (20)$$

где $\bar{x}$ - среднее арифметическое значение $\{x_i\}$: $\bar{x}=\frac{1}{n}\sum_i^n x_i$; $\bar{y}$ – среднее арифметическое значение $\{y_i\}$: $\bar{y}=\frac{1}{n}\sum_i^n y_i$; $\overline{x^2}$ – среднее арифметическое значение $\{x_i^2\}$: $\overline{x^2}=\frac{1}{n}\sum_i^n x_i^2$; $\overline{xy}$ – среднее арифметическое значение произведения $\{x_i y_i\}$: $\overline{xy}=\frac{1}{n}\sum_i^n x_i y_i$.

Итак, параметры a и b найдены. Функция приближения $f(x)$ определена.

Дополнительная информация о результатах аппроксимации из математической обработки вытекает следующая. Во-первых, оказывается $f(x)$ можно записать в виде:

$$f(x)=\bar{y}+b(x-\bar{x}).$$

Это означает, что прямая обязательно проходит через точку $(\bar{x},\bar{y})$. Во-вторых, можно указать погрешности:

$s_y^2 = \frac{1}{n-2}\sum_{i=1}^{n}(f_i - y_i)^2$ - погрешность исходных данных $\{y_i\}$,

$s_a^2 = \frac{\overline{x^2}}{\overline{x^2} - (\overline{x})^2}\frac{s_y^2}{n}$ - погрешность параметра *a* из формулы *f(x)=a+ bx,*

$s_b^2 = \frac{1}{\overline{x^2} - (\overline{x})^2}\frac{s_y^2}{n}$ - погрешность параметра *b* из формулы *f(x)=a+ bx,*

$s_{f_i}^2 = \left[1 + \frac{(x_i - \overline{x})^2}{\overline{x^2} - (\overline{x})^2}\right]\frac{s_y^2}{n}$ - погрешность определения аппроксимирующей функции *f(x).*

8. ЭМПИРИЧЕСКИЕ ПРЯМЫЕ РЕГРЕССИИ И ЭМПИРИЧЕСКИЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Регрессии. Уравнение *y=f(x)* называется уравнением регрессии *y* на *x*, а соответствующая линия – линией регрессией *y* на *x*. Регрессия *y* на *x* показывает, как изменяется в среднем величина *y* при изменении величины *x*. Изучение линий и функций регрессии составляет предмет корреляционного анализа.

Линейная корреляция. Если обе функции регрессии *f(x)* и *g(y)* линейны, то говорят, что между величинами *x* и *y* имеет место линейная корреляционная зависимость (или, просто, линейная корреляция). В этом случае обе линии регрессии будут прямыми линиями. Они называются прямыми регрессии.

Корреляция между коэффициентами уравнения регрессии. Пусть в результате обработки получено уравнение прямой линии $y(x) = f(x) = a + bx = \overline{y} + b(x_i - \overline{x})$. Коэффициенты в уравнении *a* и *b* есть:

$$a = \frac{\overline{x^2}\,\overline{y} - \overline{x}\,\overline{xy}}{\overline{x^2} - \overline{x}^2}, \qquad b = -\frac{\overline{x}\,\overline{y} - \overline{xy}}{\overline{x^2} - \overline{x}^2},$$

где $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$; $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$; $\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$.

Точечные оценки дисперсии есть:

$$s_y^2 = \frac{1}{n-2}\sum_{i=1}^{n}(t_i - y_i)^2 ; \; s_a^2 = \frac{\overline{x^2}}{\overline{x^2} - \overline{x}^2}\frac{s_y^2}{n};$$

$$s_b^2 = \frac{1}{\overline{x^2} - \overline{x}^2}\frac{s_y^2}{n}; \; s_{f_i}^2 = \left[1 + \frac{(x_i - \overline{x})^2}{\overline{x^2} - \overline{x}^2}\right]\frac{s_y^2}{n}.$$

Значение прямой линии в точке пересечения с осью *x* (рис.5): $\hat{x} = -\frac{a}{b}$, при $f(\hat{x}) = 0$. Точечная оценка дисперсии (погрешности) $\hat{x}$:

$$s_{\hat{x}}^2 = \hat{x}^2\left(\frac{s_a^2}{a^2} + \frac{s_b^2}{b^2} - 2\rho_{ab}\frac{s_a}{a}\frac{s_b}{b}\right),$$

где ρ_{ab} – коэффициент корреляции между *a* и *b*.

$$\rho_{ab} = -\frac{\overline{x}}{\sqrt{\overline{x^2}}} ; \; \hat{x} = \overline{x} - \frac{\overline{y}}{b};$$

$$\frac{s_{\hat{x}}^2}{\hat{x}^2} = \frac{s_a^2}{a^2} + \frac{s_b^2}{b^2} + \frac{2\overline{x}}{\sqrt{\overline{x^2}}}\frac{s_a}{a}\frac{s_b}{b} ;$$

$$s_{f(\hat{x})}^2 = \left[1 + \frac{\overline{y}^2}{b^2(\overline{x^2} - \overline{x}^2)}\right]\frac{s_y^2}{n}.$$

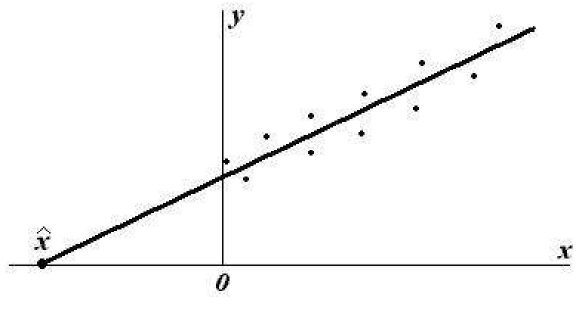


Рис.5.

Пример. Было выполнено 10 пар измерений *{x_i; y_i}*. Результаты измерений представлены в табл.4. Требуется написать уравнение линии регрессии и выявить тесноту связи между коэффициентами этого уравнения.

Таблица 4

x	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
y	5,2	5,6	5,2	8,1	9,9	8,8	12,5	12,7	14,1	13,0
	0,2	-0,4	-1,8	0,1	0,9	-1,2	1,5	0,7	1,1	-1,0
$f_{ист.}$	5	6	7	8	9	10	11	12	13	14

$f_{ист.}$=5+x; y=f+3; n=10.

$\overline{x}$=4,5; $\overline{x}^2$=20,25; $\overline{x^2}$=28,5; $\overline{y}$=9,51; $\overline{y}^2$=90,44; $\overline{xy}$=51,8.

s_y^2=11,93; s_y=3,45; a=4,6; s_a^2=0,44; s_a=0,66; b=1,085; s_b^2=1,54·10^{-2}; s_b=0,124.

$\hat{x}$=-4,25; $s_{\hat{x}}^2$=1,14; $s_{\hat{x}}$=1,07; $s_{f(\hat{x})}^2$=1,41; $s_{f(\hat{x})}$=1,19.

$\rho_{ab}\approx$-0,85.

Таким образом, уравнение аппроксимирующей прямой линии есть *f(y)=4,6+1,085x*. Отсюда следует, что полученная расчетная величина a=4,6 занижена, а b=1,085 завышена, т.е. ρ_{ab}<0 (рис.6).

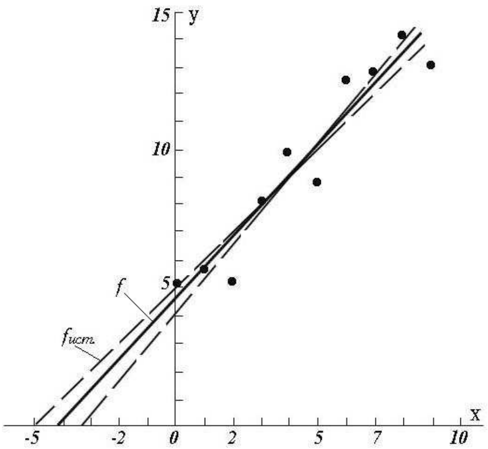


Рис.6.

Коэффициент корреляции. С помощью полученных выше эмпирических оценок – эмпирических средних и стандартов (эмпирических дисперсий s_x^2 и s_y^2) образуют так называемый эмпирический коэффициент корреляции:

$$r_n = \frac{\sum\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)}{\sqrt{\sum\left(x_i-\overline{x}\right)^2}\sqrt{\sum\left(y_i-\overline{y}\right)^2}}, \qquad (21)$$

который и применяется в качестве оценки коэффициента корреляции

$$r(x,y)\approx r_n .$$

Эта оценка является уже несколько смещенной, но она удобна для расчетов эмпирических прямых регрессии. Уравнения эмпирических прямых регрессии получают из уравнений прямых регрессий путем замены в них всех параметров соответствующими оценками. Так, *эмпирическая прямая регрессии y на x* имеет уравнение

$$y-\overline{y}=r_n\frac{s_y}{s_x}\left(x-\overline{x}\right), \qquad (22)$$

а *эмпирическая прямая регрессии x на y* – уравнение

$$x-\overline{x}=r_n\frac{s_x}{s_y}\left(y-\overline{y}\right). \qquad (23)$$

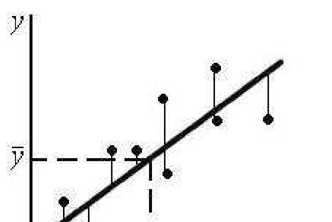
Эмпирические прямые регрессии (22) и (23) являются прямыми наилучшего среднеквадратического приближения к эмпирическим точкам {x_i; y_i} в следующем смысле: сумма квадратов отклонений эмпирических значений y_i от эмпирической прямой регрессии (22) меньше, чем сумма квадратов отклонений их от любой другой прямой:

$$\sum_{i=1}^{n}\left[y_i-\overline{y}-r_n\frac{s_y}{s_x}\left(x_i-\overline{x}\right)\right]^2\le\sum_{i=1}^{n}\left[y_i-\left(Ax_i+B\right)\right]^2,$$

и аналогично сумма квадратов отклонений эмпирических значений x_i от эмпирической прямой регрессии (23) меньше, чем сумма квадратов отклонений их от любой другой прямой:

$$\sum_{i=1}^{n}\left[x_i-\overline{x}-r_n\frac{s_x}{s_y}\left(y_i-\overline{y}\right)\right]^2\le\sum_{i=1}^{n}\left[x_i-\left(Ay_i+B\right)\right]^2.$$

На рис.7 и 8 показано, какие отклонения имеются в виду в обоих случаях.



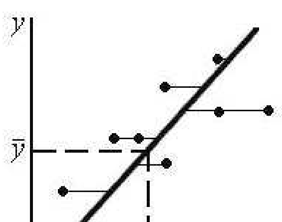


Рис.7. **Рис.8.**

В условиях предыдущего примера рассчитанный эмпирический коэффициент корреляции равен r_n=0,952.

Полученные средние и эмпирические стандарты позволяют написать уравнения эмпирических регрессий y на x:

$$y-9{,}51=0{,}952\frac{3{,}45}{2{,}87}(x-4{,}5)=1{,}144(x-4{,}5),$$

и x на y:

$$x-4{,}5=0{,}952\frac{2{,}87}{3{,}45}(y-9{,}51)=0{,}792(y-9{,}51).$$

Оценка значимости коэффициента корреляции. Обычно решение вопроса о проверке гипотезы некоррелированности величин x и y, несмотря на наличие эмпирического коэффициента корреляции, отличного от нуля, дается при дополнительном предположении о том, что совместное распределение величин x и y является нормальным двумерным распределением (т.е. для случая нормальной корреляции). Тогда при условии равенства нулю истинного коэффициента корреляции $r(x, y)$ распределение эмпирического коэффициента корреляции зависит только от числа и результатов эксперимента, что позволяет установить критическую область для эмпирического коэффициента корреляции в зависимости от заданной надежности Р и числа n.

В табл.5 приводятся критические значения произведения $|r_n|$ на $\sqrt{n-1}$.

Таблица 5

Критические значения $|r_n|\sqrt{n-1}$

P / n	*0,95*	*0,99*	*0,999*	P / n	*0,95*	*0,99*	*0,999*
10	1,89	2,29	2,62	25	1,94	2,47	3,03
11	1,90	2,32	2,68	30	1,94	2,49	3,07
12	1,91	2,35	2,73	35	1,95	2,50	3,10
13	1,91	2,37	2,77	40	1,95	2,51	3,13
14	1,92	2,39	2,81	50	1,950	2,527	3,160
15	1,92	2,40	2,84	60	1,953	2,536	3,184
16	1,93	2,41	2,87	70	1,954	2,541	3,198
17	1,93	2,42	2,90	80	1,955	2,546	3,209
18	1,93	2,43	2,92	90	1,956	2,550	3,219
19	1,93	2,44	2,94	100	1,956	2,553	3,226
20	1,94	2,45	2,96	∞	1,960	2,576	3,291

Если произведение $|r_n|\sqrt{n-1}$ больше критического значения, то с надежностью P утверждается, что истинный коэффициент корреляции $r(x, y)$ отличен от нуля.

Для нашего примера имеем: n=10; r_n=0,952; $|r_n|\sqrt{n-1}$=2,856.

Полученное число 2,856 больше критического значения при надежности P=0,99. Это свидетельствует о наличии линейной корреляции между величинами x и y. Если при том же числе n=10 мы получили бы r_n=0,4 и, значит, $|r_n|\sqrt{n-1}$=1,2, то надежность вывода о линейной корреляции была бы менее 0,95 и вправе было бы поставить этот вывод под сомнение.

9. МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ПРЯМЫХ РАВНОРАССЕЯННЫХ НАБЛЮДЕНИЙ

Прямыми называются измерения, в результате которых находят непосредственно искомые значения величин.

Результаты наблюдений X_1; X_2; X_3; ……X_n, получаемые при прямых измерениях постоянной физической величины Q, называются *равнорассеянными* (равноточными), если они являются независимыми, одинаково распределенными случайными величинами.

Равнорассеянные результаты получают при измерениях, проводимых одним или группой наблюдателей с помощью одних и тех же средств измерений в неизменных условиях внешней среды. Их математическую обработку проводят для определения итога измерений в виде выражения (7) или (9). Результаты обрабатывают по-разному в зависимости от того, мало (n<40) или много (n>40) проведено наблюдений.

При количестве наблюдений $n<40$ результаты обрабатывают в следующем порядке.

1. Определяют точечную оценку истинного значения измеряемой величины – среднее арифметическое результатов наблюдений $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

2. Вычисляют случайные отклонения результатов наблюдений $(X_i - \overline{X})$ и их квадраты $(X_i - \overline{X})^2$.

3. Вычисляют точечную оценку дисперсии случайной погрешности (среднего квадратического отклонения результатов наблюдений) $S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2$.

4. Определяют точечную оценку дисперсии среднего арифметического (среднего квадратического отклонения результата измерений) $S_{\bar{x}}^2 = \frac{1}{n}S_x^2 = \frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \overline{X})^2$.

5. Проверяют нормальность распределения результатов наблюдений.

6. Задаваясь определенными значениями доверительной вероятности, находят доверительную погрешность результата измерений и доверительный интервал для среднего квадратического отклонения.

7. Определяют наличие грубых погрешностей и промахов. Если они обнаружены, то соответствующие результаты отбрасывают и повторяют вычисления.

8. Итог измерений записывают в виде:

$$Q = \overline{X} \pm t_p \cdot S_{\overline{X}} \quad (n=\ldots\ldots;\ P=\ldots..\%), \qquad (24)$$

если распределение случайных погрешностей нормально, и в виде

$$Q = \overline{X}\ ;\ (S_{\overline{X}} = \ldots\ldots;\ n = \ldots.) \qquad (25)$$

в остальных случаях.

Пример. Обработать наблюдения, полученные при измерении ускорения свободного падения:

i	g_i, *м/с*^{2}	$(g_i - \overline{g}) \cdot 10^{-5}$	$(g_i - \overline{g})^2 \cdot 10^{-10}$
1	**2**	**3**	**4**
1	9,81931	+7,7	59,29
2	9,81920	-3,3	10,89
3	9,81919	-4,3	18,49
4	9,81941	+17,7	313,29
5	9,81904	-19,3	372,49
6	9,81923	-0.3	0,09
7	9,81925	+1.7	2,89
8	9,81918	-5.3	28,09
9	9,81932	+8.7	75,69
10	9,81929	+5.7	32,49
11	9,81920	-3.3	10,89

1	2	3	4
12	9,81925	+1,7	2,89
13	9,81916	-7,3	53,29
14	9,81918	-5,3	28,09
15	9,81928	+4,7	22,09
Суммы:	146,28849	-0,5	1030,95

$\overline{g}$ =9,819233 м·с$^{-2}$ S_g=8,6·10^{-5} м·с$^{-2}$.

1. Вычисляем точечную оценку истинного значения измеряемой величины, то есть среднее арифметическое данных наблюдений:

$$\overline{g} = \frac{1}{15}\sum_{i=1}^{15} g_i = 9{,}81923267 \ (м \cdot c^{-2}) .$$

Полученное числовое значение округляем так, чтобы случайные отклонения результатов наблюдений имели не больше трех значащих цифр при ответственных или особо точных измерениях, а также в тех случаях, когда полученные результаты предполагается использовать при дальнейших вычислениях. В противном случае округляют до двух значащих цифр. Поэтому округленно $\overline{g}$ =9,819233 м·с$^{-2}$.

2. Сумма отклонений результатов наблюдений $\sum_{i}^{15} g_i - \overline{g}$ составляет 0,0000005 м·с$^{-2}$, хотя должна равняться нулю. Расхождение объясняется тем, что значение среднего арифметического было округлено.

Результаты 4-го и 5-го наблюдений кажутся сомнительными, поэтому в дальнейшем надо проверить, не содержат ли они грубых погрешностей.

3. Оценка дисперсии случайной погрешности:

$$S_g = \sqrt{\frac{1}{15-1}\sum_{i=1}^{15}(g_i - \overline{g})^2} = 8{,}6 \cdot 10^{-5} \ (м \cdot c^{-2}) .$$

4. Оценка дисперсии среднего арифметического:

$$S_{\overline{g}} = \frac{S_g}{\sqrt{n}} = \frac{8{,}6 \cdot 10^{-5}}{\sqrt{15}} = 2{,}2 \cdot 10^{-5} \ (м \cdot c^{-2}) .$$

5. Для проверки нормальности распределения по формуле (14) находим значения статистической функции распределения и соответствующие им значения z. Результаты вычислений представлены в таблице:

g_k	$F_n(g_k)=\Phi(z_k)$	z_k	g_k	$F_n(g_k)=\Phi(z_k)$	z_k
9,81904	0,0625	-1,534	9,81925	0,6250	0,319
9,81916	0,1250	-1,150	9,81928	0,6771	0,461
9,81918	0,2500	-0,674	9,81929	0,7500	0,674
9,81919	0,3225	-0,461	9,81931	0,8225	0,925
9,81920	0,4475	-0,132	9,81932	0,8750	1,150
9,81923	0,5000	0,000	9,81941	0,9275	1,534

По результатам таблицы строим зависимость $z_k(g_k)$, соответствующую опытным данным. Если разброс отдельных точек относительно прямой $z=\frac{x-m_x}{\sigma_x}$ не велик, то распределение с некоторым приближением можно считать нормальным.

6. Задаваясь доверительной вероятностью α=1-q=96% для дисперсии случайной погрешности и p=95% для среднего арифметического, при k=n-1=14 из табл. 2 и 3 Приложения находим $t_{0,95}$=2,145; $\chi^2_{14;0,02}$=5,368; $\chi_{14;0,02}$=2,32; $\chi^2_{14;0,98}$=26,873; $\chi_{14;0,98}$=5,19.

Доверительная погрешность результата измерений:

$$\delta_P=\delta_{0,95}=t_{0,95}\cdot S_{\bar{g}}=2{,}145\cdot 2{,}2\cdot 10^{-5}=4{,}7\cdot 10^{-5}\ (\text{м}\cdot\text{с}^{-2}),$$

а границы доверительного интервала для дисперсии случайной погрешности следующие:

$$S_{g_1}=\frac{\sqrt{15-1}\cdot S_g}{\chi_{14;0,02}}=\frac{\sqrt{14}\cdot 8{,}6\cdot 10^{-5}}{2{,}32}=13{,}9\cdot 10^{-5}\ (\text{м}\cdot\text{с}^{-2}),$$

$$S_{g_2}=\frac{\sqrt{15-1}\cdot S_g}{\chi_{14;0,98}}=\frac{\sqrt{14}\cdot 8{,}6\cdot 10^{-5}}{5{,}19}=6{,}2\cdot 10^{-5}\ (\text{м}\cdot\text{с}^{-2}).$$

7. Проверим, не содержит ли грубую погрешность пятый результат $g_5=g_{min}$=9,81904 м·с$^{-2}$ с наибольшей по абсолютной величине погрешностью.

Для этого, задавшись уровнем значимости, например, $q=1-\alpha$=0,05 из табл.6 Приложения находим $\nu_{0,95}$=2,493. Но $\nu=\frac{\bar{g}-g_{\min}}{S_g}=\frac{9{,}819233-9{,}81904}{8{,}6\cdot 10^{-5}}=2{,}245.$ Так как $\nu<\nu_{0,95}$, то достаточных оснований считать результат g_5=9,81904 м·с$^{-2}$, содержащим грубую погрешность, нет.

Окончательно итог измерений записываем в виде:

$$g \approx 9{,}819233 \text{ м}\cdot\text{с}^{-2} \ (S_{\bar g} = 0{,}000022 \text{ м}\cdot\text{с}^{-2};\ n=15), \text{ или}$$

$$g = 9{,}819233 \pm 0{,}000047 \text{ м}\cdot\text{с}^{-2} \ (n=15;\ P=95\%).$$

При числе наблюдений $n \geq 40$ обработка результатов по вышеприведенной методике связана с большой вычислительной работой. В этом случае применяют способ обработки, основанный на группировании данных, как при построении гистограмм.

10. МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ КОСВЕННЫХ ИЗМЕРЕНИЙ

При *косвенных* измерениях значение искомой величины получают на основании известной зависимости, связывающей ее с другими величинами, которые подвергаются прямым измерениям.

Рассмотрим простейший случай, когда искомая величина Q_z определяется как сумма двух величин Q_x и Q_y:

$$Q_z = Q_x + Q_y \,. \qquad (26)$$

Так как результаты прямых измерений величин Q_x и Q_y после исключения систематических погрешностей включают в себя некоторые случайные погрешности, то формулу косвенного измерения суммы (26) можно переписать в виде:

$$\overline{Z} - \lambda_z = \overline{X} - \lambda_x + \overline{Y} - \lambda_y, \qquad (27)$$

где $\overline{X}$ и $\overline{Y}$ – средние арифметические (или средние взвешенные), полученные при обработке результатов прямых измерений величин Q_x и Q_y; λ_x и λ_y – случайные погрешности средних; $\overline{Z}$ и $\lambda_я$ – оценка истинного значения косвенно измеряемой величины и ее случайная погрешность.

Из (27) вытекает справедливость двух следующих равенств:

$$\overline{Z} = \overline{X} + \overline{Y}; \quad \lambda_z = \lambda_x + \lambda_y, \qquad (28)$$

то есть оценкой истинного значения косвенно измеряемой величины служит сумма оценок истинных значений исходных величин, случайные погрешности которых складываются.

Дисперсия оценки $\overline{Z}$ составляет:

$$\sigma^2_{\overline{Z}} = \sigma^2_{\overline{X}} + \sigma^2_{\overline{Y}} + 2M[\lambda_X \cdot \lambda_Y]. \qquad (29)$$

Третье слагаемое в (29) есть математическое ожидание произведения случайных погрешностей и называется *корреляционным моментом*, которое определяет степень “тесноты” линейной зависимости между погрешностями.

Часто вместо корреляционного момента пользуются безразмерной величиной – *коэффициентом корреляции*

$$r_{\overline{X}\cdot\overline{Y}} = \frac{M[\lambda_X \cdot \lambda_Y]}{\sigma_{\overline{X}} \cdot \sigma_{\overline{Y}}}. \qquad (30)$$

Отсюда, в частности, следует, что коэффициент корреляции между погрешностями λ_x и λ_y средних арифметических равен коэффициенту корреляции между погрешностями δ_x и δ_y результатов отдельных измерений величин Q_x и Q_y: $r_{\overline{X}\cdot\overline{Y}} = r_{X\cdot Y}$.

Дисперсия результата косвенных измерений (оценки истинного значения косвенно измеряемой величины) с учетом коэффициента корреляции:

$$\sigma^2_{\overline{Z}} = \sigma^2_{\overline{X}} + \sigma^2_{\overline{Y}} + 2 \cdot r_{XY} \cdot \sigma_{\overline{X}} \cdot \sigma_{\overline{Y}}. \qquad (31)$$

Если погрешности измерения величин Q_x и Q_y не коррелированы, то (31) упрощается:

$$\sigma^2_{\overline{Z}} = \sigma^2_{\overline{X}} + \sigma^2_{\overline{Y}}; \quad \sigma_{\overline{Z}} = \sqrt{\sigma^2_{\overline{X}} + \sigma^2_{\overline{Y}}}. \qquad (32)$$

В случаях, когда теоретические дисперсии распределения результатов прямых измерений неизвестны, определяется оценка дисперсии результата косвенных измерений $S^2_{\overline{Z}}$ через оценки дисперсий $S^2_{\overline{X}}$ и $S^2_{\overline{Y}}$:

$$S^2_{\overline{Z}} = S^2_{\overline{X}} + S^2_{\overline{Y}} + 2 \cdot \bar{r}_{XY} \cdot S_{\overline{X}} \cdot S_{\overline{Y}}. \qquad (33)$$

Оценки коэффициента корреляции $\bar{r}_{XY}$ вычисляют на основании результатов прямых измерений исходных величин:

$$\bar{r}_{XY} = \frac{1}{S_{\overline{X}} \cdot S_{\overline{Y}}} \sum_{i=1}^{m} (X_i - \overline{X})(Y_i - \overline{Y}), \qquad (34)$$

где $m = min\,(n_x, n_y)$ – наименьшее из чисел наблюдений n_x и n_y.

Распределение результата косвенных измерений будет нормальным, если нормальны распределения результатов прямых измерений. В этих условиях для построения доверительного интервала, накрывающего истинное значение измеряемой величины, следует применить нормированную функцию нормального распределения, если число измерений достаточно велико.

В противном случае (если объемы рядов прямых измерений недостаточно велики) можно воспользоваться распределением Стьюдента с некоторым “эффективным” числом степеней свободы, которое при независимости погрешностей измерения (r_{xy}=0) подсчитывается по формуле

$$k_{эфф} = \frac{(S^2_{\overline{X}} + S^2_{\overline{Y}})^2}{\dfrac{S^4_{\overline{X}}}{n_x - 1} + \dfrac{S^4_{\overline{Y}}}{n_y - 1}} = \frac{S^4_{\overline{Z}}}{\dfrac{S^4_{\overline{X}}}{n_x - 1} + \dfrac{S^4_{\overline{Y}}}{n_y - 1}}, \qquad (35)$$

где n_x и n_y – числа прямых наблюдений величин Q_x и Q_y.

При $n_x = n_y = n$ выражение для эффективного числа степеней свободы распределения Стьюдента упрощается

$$k_{эфф} = (n-1)\frac{S^4_{\overline{Z}}}{S^4_{\overline{X}} + S^4_{\overline{Y}}}. \qquad (36)$$

Итоговый результат измерений записываем в виде $Q_Z = \overline{Z} \pm t_p \cdot S_{\overline{Z}}$.

Рассмотрим теперь *общий случай*, когда требуется оценить истинное значение величины Q, связанной с величинами Q_j (j =1,2,3,….,n), которые измеряются прямыми способами, некоторым нелинейным уравнением:

$$Q = F(Q_1; Q_2; \ldots; Q_m). \qquad (37)$$

В качестве наиболее достоверного значения $\overline{X}_Q$ косвенно измеряемой величины Q следует понимать значение, получаемое подстановкой в формулу (37) косвенного измерения средних арифметических $\overline{X}_j$ рядов измерений исходных величин:

$$\overline{X}_j = F(\overline{X}_1; \overline{X}_2; \ldots; \overline{X}_m). \qquad (38)$$

Дисперсия этой оценки определяется из формулы:

$$\sigma^2_{\overline{XQ}} = \sum_{j=1}^{m}\left(\frac{\partial F}{\partial Q_j}\right)^2 \cdot \sigma^2_{\overline{Xj}} + \sum_{i;\,j=1}^{m}\left(\frac{\partial F}{\partial Q_i}\right)\cdot\left(\frac{\partial F}{\partial Q_j}\right)\cdot r_{ij}\cdot\sigma_{\overline{Xi}}\cdot\sigma_{\overline{Xj}}, \qquad (39)$$

причем значения частных производных вычисляются при средних арифметических значениях аргументов $Q_j = \overline{X}_j$.

В (39) первое слагаемое – произведения частных производных уравнения косвенного измерения на среднее квадратическое отклонение результатов измерения соответствующих аргументов *называются частными погрешностями косвенного измерения*:

$$E_j = \left(\frac{\partial F}{\partial Q_j}\right)\cdot\sigma_{\overline{Xj}}. \qquad (40)$$

Если случайные погрешности измерения отдельных аргументов попарно некоррелированы (r_{ij}=0; i; j=1,2,….,m), то дисперсия результатов равна сумме квадратов частных погрешностей:

$$\sigma^2_{\overline{XQ}} = \sum_{j=1}^{m} E_j^2. \qquad (41)$$

Погрешность результата косвенного измерения наряду со случайной содержит еще и систематическую составляющую, которая определяется как

$$\theta = \frac{1}{2}\sum_{j=1}^{m}\left(\frac{\partial^2 F}{\partial Q_j^2}\right)\cdot\sigma^2_{\overline{Xj}}. \qquad (42)$$

Чтобы исключить эту систематическую погрешность, нужно к рассчитанному по формуле (38) результату прибавить суммарную поправку q, равную систематической погрешности по величине и обратную по знаку.

Окончательный итог косвенного измерения:

$$Q = \overline{X}_Q + q \pm t_p \cdot \sigma_{\overline{XQ}}, \qquad (43)$$

если дисперсии исходных величин известны, и

$$Q = \overline{X}_Q + q \pm t_p \cdot S_{\overline{XQ}}, \qquad (44)$$

если теоретические дисперсии неизвестны.

Оценка дисперсии случайной погрешности $S_{\overline{XQ}}$ проводится при этом по вышеприведенным формулам подстановкой вместо теоретических дисперсий их оценок $S^2_{\overline{Xj}}$.

Величина t_p в формулах (43) и (44) вычисляется, как и ранее, для заданной доверительной вероятности, исходя из закона распределения итогового результата косвенного измерения.

Пример. Определить момент инерции круглой платформы на трифилярном подвесе, связанной формулой $I = \frac{gRr}{4\pi^2 l} \cdot mT^2$ со следующими величинами, измеряемыми прямыми способами:

R=(11,50±0,05)·10^{-2} *м* – радиус платформы;

r=(10,00±0,05)·10^{-2} *м* – радиус верхнего диска платформы;

l=(233,0±0,2)·10^{-2} *м* – длина нитей подвеса;

m=(125,7±0,1)·10^{-3} *кг* – масса платформы;

T=(2,81±0,01)·*с* – период малых колебаний платформы;

g= 9,81 *м·с*$^{-2}$ – ускорение свободного падения.

(Все результаты даны с отклонениями, соответствующими t_p=1).

Подставляя в исходную формулу средние арифметические значения измеряемых прямыми способами величин и округленные значения постоянных, получим оценку истинного значения момента инерции платформы:

$$\bar{I} = \frac{9,82 \cdot 11,50 \cdot 10,00}{4 \cdot 3,14^2 \cdot 233} \cdot 125,7 \cdot 2,81^2 \cdot 10^{-5} = 1,22 \cdot 10^{-3} \text{ кг·м}^2,$$

поскольку результат должен быть округлен до трех значащих цифр.

Для оценки точности полученного значения момента инерции вычислим частные производные и частные погрешности косвенного измерения:

$$E_R = \left(\frac{\partial I}{\partial R}\right) \cdot S_{\bar{R}} = \frac{I}{R} \cdot S_{\bar{R}} = \frac{1,22 \cdot 10^{-3}}{11,50} \cdot 0,05 = 0,0053 \cdot 10^{-3} \text{ кг·м}^2;$$

$$E_r = \left(\frac{\partial I}{\partial r}\right) \cdot S_{\bar{r}} = \frac{I}{r} \cdot S_{\bar{r}} = \frac{1,22 \cdot 10^{-3}}{10,00} \cdot 0,05 = 0,0061 \cdot 10^{-3} \text{ кг·м}^2;$$

$$E_l = \left(\frac{\partial I}{\partial l}\right) \cdot S_{\bar{l}} = -\frac{I}{l} \cdot S_{\bar{l}} = -\frac{1,22 \cdot 10^{-3}}{233,0} \cdot 0,2 = -0,0010 \cdot 10^{-3} \text{ кг·м}^2;$$

$$E_m = \left(\frac{\partial I}{\partial m}\right) \cdot S_{\bar{m}} = \frac{I}{m} \cdot S_{\bar{m}} = \frac{1,22 \cdot 10^{-3}}{125,7} \cdot 0,1 = 0,0010 \cdot 10^{-3} \text{ кг·м}^2;$$

$$E_T = \left(\frac{\partial I}{\partial T}\right) \cdot S_{\bar{T}} = 2\frac{I}{T} \cdot S_{\bar{T}} = 2\frac{1,22 \cdot 10^{-3}}{2,81} \cdot 0,01 = 0,0087 \cdot 10^{-3} \text{ кг·м}^2.$$

Среднее квадратическое отклонение косвенного измерения момента инерции платформы рассчитаем по формуле (41):

$$S_I = \sqrt{E_R^2 + E_r^2 + E_l^2 + E_m^2 + E_T^2} = 0,01 \cdot 10^{-3} \text{ кг·м}^2.$$

Окончательно итог косвенного измерения записываем в виде

I= (1,22±0,01)·10^{-3} *кг·м*^{2}.

Далее по формуле (42) следует оценить систематическую погрешность, возникающую при косвенных измерениях:

$\theta = \frac{1}{2}\sum_{j=1}^{m}\left(\frac{\partial^2 F}{\partial Q_j^2}\right) \cdot S_{Q_j}^2 = \frac{1}{2}\left[\left(\frac{\partial^2 I}{\partial l^2}\right) \cdot S_l^2 + \left(\frac{\partial^2 I}{\partial T^2}\right) \cdot S_T^2\right]$, так как вторые производные по остальным аргументам равны нулю.

Тогда

$$\theta=\frac{1}{2}\left(\frac{I}{l^2}\cdot S_l^2+2\frac{I}{T^2}\cdot S_T^2\right)=\frac{1}{2}\left(\frac{1{,}22\cdot10^{-3}}{233^2}\cdot0{,}04+\frac{2\cdot1{,}22\cdot10^{-3}}{2{,}81^2}\cdot0{,}0001\right)=0{,}0016\cdot10^{-3}\ кг\cdot м^2.$$

Полученная величина значительно меньше пяти единиц разряда, следующего за последней значащей цифрой погрешности результата. Если эту погрешность учесть путем введения в итог измерения соответствующей поправки, то она все равно пропадет при округлении. Поэтому принимается θ=0.

ЛИТЕРАТУРА

1. *Румшиский Л.З.* Математическая обработка результатов эксперимента. – М.: Наука, 1971.
2. *Дудникова Н.И., Мищенко С.С., Чен Б.Б.* Обработка результатов физического эксперимента: Методическое пособие /Отв. ред. В.М.Лелевкин. – Бишкек: КРСУ, 1999. – 38 с.

ПРИЛОЖЕНИЯ

36

В.С.Энгельшт, Б.Б.Чен

СТАТИСТИЧЕСКАЯ ОБРАБОТКА
РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Методическое пособие

Редактор И.С.Вязьмина
Технический редактор О.А.Матвеева
Корректор О.А.Старцева
Компьютерная верстка Е.Г.Шевёлкина

Подписано к печати 5.06.2002. Формат 60х84 $^1/_{16}$.
Офсетная печать. Объем 3,0 п.л.
Тираж 100 экз. Заказ 148.

Издательство Кыргызско-Российского Славянского университета
720000, Бишкек, Киевская, 44

Отпечатано в типографии КРСУ
720000, Бишкек, Шопокова, 68

КЫРГЫЗСКО-РОССИЙСКИЙ СЛАВЯНСКИЙ УНИВЕРСИТЕТ

В.С.ЭНГЕЛЬШТ, Б.Б.ЧЕН

СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Бишкек 2002

Таблица 1

Распределение Стьюдента.

Значения $P\{|t|<t_P\}=2\int_0^{t_p} S(t;k)dt$ для различных t_p

k	t_p				*k*	t_p			
	2,0	*2,5*	*3,0*	*3,5*		*2,0*	*2,5*	*3,0*	*3,5*
1	0,7048	0,7578	0,7952	0,8228	12	0,9314	0,9720	0,9890	0,9956
2	0,8164	0,8764	0,9046	0,9276	13	0,9392	0,9737	0,9898	0,9960
3	0,8606	0,9122	0,9424	0,9606	14	0,9848	0,9740	0,9904	0,9964
4	0,8838	0,9332	0,9600	0,9752	15	0,9360	0,9754	0,9910	0,9968
5	0,8980	0,9454	0,9700	0,9828	16	0,9372	0,9764	0,9916	0,9970
6	0,9076	0,9534	0,9760	0,9872	17	0,9382	0,9770	0,9920	0,9972
7	0,9144	0,9590	0,9800	0,9900	18	0,9392	0,9776	0,9924	0,9974
8	0,9194	0,9630	0,9830	0,9920	19	0,9400	0,9782	0,9926	0,9976
9	0,9234	0,9662	0,9850	0,9932	20	0,9408	0,9788	0,9930	0,9978
10	0,9266	0,9686	0,9866	0,9942	∞	0,9545	0,9876	0,9973	0,9995
11	0,9292	0,9704	0,9880	0,9950					

Таблица 2

Распределение Стьюдента $P\{|t| < t_P\} = 2\int_0^{t_P} S(t;k)dt$

k	*P*											
	0,10	*0,20*	*0,30*	*0,40*	*0,50*	*0,60*	*0,70*	*0,80*	*0,90*	*0,95*	*0,98*	*0,99*
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977
15	0,126	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898

Продолжение таблицы 2

k	*P*											
	0,10	*0,20*	*0,30*	*0,40*	*0,50*	*0,60*	*0,70*	*0,80*	*0,90*	*0,95*	*0,98*	*0,99*
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,707
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750
∞	0,12566	0,25335	0,38532	0,52440	0,67449	0,84162	1,03643	1,28155	1,64485	1,95996	2,32634	2,57582

Таблица 3

Интегральная функция χ^2 – распределения Пирсона. Значения $\chi^2_{k;P}$ для различных k и P

k	*P*												
	0,01	*0,02*	*0,05*	*0,10*	*0,20*	*0,30*	*0,50*	*0,70*	*0,80*	*0,90*	*0,95*	*0,98*	*0,99*
1	0,000157	0,000628	0,00393	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,0201	0,0404	0,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,210
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,070	13,388	15,086
6	0,872	1,134	1,635	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	1,646	2,032	2,733	3,490	4,594	5,527	7,344	9,524	11,030	13,362	15,507	18,168	20,090
9	2,088	2,532	3,325	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	2,558	3,059	3,940	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	9,034	11,340	14,011	15,812	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	25,472	27,688
14	4,660	5,368	6,571	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	5,812	6,614	7,962	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32,000

Продолжение таблицы 3

k	*P*												
	0,01	*0,02*	*0,05*	*0,10*	*0,20*	*0,30*	*0,50*	*0,70*	*0,80*	*0,90*	*0,95*	*0,98*	*0,99*
17	6,408	7,255	8,672	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,390	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	15,352	18,338	21,689	23,900	27,204	30,144	33,687	36,191
20	8,260	9,237	10,851	12,444	14,578	16,266	19,337	22,775	25,038	28,412	31,410	35,020	37,566
21	8,897	9,915	11,591	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	9,542	10,600	12,338	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,695	40,289
23	10,196	11,293	13,091	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,270	42,980
25	11,524	12,697	14,611	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	12,879	14,125	16,151	18,114	20,703	22,710	26,336	30,319	32,912	36,741	40,113	44,140	46,963
28	13,565	14,847	16,928	18,939	21,588	23,647	27,336	31,391	34,027	37,916	41,337	45,419	48,278
29	14,256	15,574	17,708	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588
30	14,953	16,306	18,493	20,599	23,364	25,508	29,336	33,530	36,250	40,256	43,773	47,962	50,892

Таблица 4

Интегральная функция нормированного нормального распределения. Значения z для различных *Ф(z)*

Ф(z)	*z*	*Ф(z)*	*z*	*Ф(z)*	*z*	*Ф(z)*	*z*
0,005	-3,2905	0,25	-0,6745	0,50	+0.0000	0,80	+0,8416
0,005	-2,575	0,30	-0,5244	0,55	+0.1257	0,85	+1,0364
0,01	-2,3267	0,35	-0,3853	0,60	+0.2533	0,90	+1,2816
0,05	-1,6449	0,40	-0,2533	0,65	+0.3853	0,95	+1,6449
0,10	-1,2816	0,45	-0,1257	0,70	+0.5244	0,99	+2,3267
0,15	-1,0364	0,50	-0,0000	0,75	+0.6745	0,995	+2,5750
0,20	-0,8416					0,9995	+3,2905

Таблица 5

Дифференциальная функция нормированного нормального распределения $p(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\cdot t^2}$

t	*0,00*	*0,01*	*0,02*	*0,03*	*0,04*	*0,05*	*0,06*	*0,07*	*0,08*	*0,09*	*t*
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973	0,0
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918	0,1
0,2	3910	3602	3894	3885	3876	3867	3857	3847	3836	3825	0,2
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697	0,3
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538	0,4
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352	0,5
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144	0,6
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920	0,7
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685	0,8
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444	0,9
1,0	0,2420	2396	2371	2347	2323	2299	2275	2215	2227	2203	1,0
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965	1,1
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736	1,2
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518	1,3
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315	1,4
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127	1,5
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957	1,6
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804	11,7
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669	1,8
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551	1,9
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449	2,0

Продолжение таблицы 5

t	*0,00*	*0,01*	*0,02*	*0,03*	*0,04*	*0,05*	*0,06*	*0,07*	*0,08*	*0,09*	*t*
2,1	0440	0431	0422	0413	0404	0396	0338	0379	0371	0363	2,1
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290	2,2
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229	2,3
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180	2,4
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139	2,5
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107	2,6
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081	2,7
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061	2,8
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046	2,9
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034	3,0
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025	3,1
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018	3,2
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013	3,3
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009	3,4
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006	3,5
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004	3,6
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003	3,7
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002	3,8
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001	3,9

Таблица 6

Значения ν_α при различных числах измерения n

n	*q=1-α*				*n*	*q=1-α*			
	0.10	*0.05*	*0.025*	*0.01*		*0.10*	*0.05*	*0.025*	*0.01*
3	1,406	1,412	1,414	1,414	14	2,297	2,461	2,602	2,759
4	1,645	1,689	1,710	1,723	15	2,326	2,493	2,638	2,808
5	1,731	1,869	1,917	1,955	16	2,354	2,532	2,670	2,837
6	1,894	1,996	2,067	2,130	17	2,380	2,551	2,701	2,871
7	1,974	2,093	2,182	2,265	18	2,404	2,557	2,728	2,903
8	2,041	2,172	2,273	2,374	19	2,426	2,600	2,754	2,932
9	2,097	2,237	2,349	2,464	20	2,447	2,623	2,778	2,959
10	2,146	2,294	2,414	2,540	21	2,467	2,644	2,801	2,984
11	2,190	2,383	2,470	2,606	22	2,486	2,664	2,823	3,008
12	2,229	2,387	2,519	2,663	23	2,504	2,683	2,843	3,030
13	2,264	2,426	2,562	2,714	24	2,520	2,701	2,862	3,051
					25	2,537	2,717	2,880	3,071

<table>
<tr><th>Номер интервала i</th><th>Середина интервала X_i</th><th>Частота m_i</th><th>Отклонение от среднего $X_i - \overline{X}$</th><th>Нормированное отклонение от среднеарифметического $t_i = \frac{X_i - \overline{X}}{S_x}$</th><th>Плотность нормированного распределения $p(t_i)$</th><th>Плотность в середине интервалов $p(x_i) = \frac{p(t_i)}{S_x}$</th><th>Теоретическая частота $np_i = n\Delta X_i \cdot p(X_i)$</th><th>Отклонение $\chi_i^2 = \frac{(m_i - nP_i)^2}{nP_i}$</th></tr>
<tr><td>1</td><td>8,912</td><td rowspan="2">1
5 } 6</td><td>-0,00736</td><td>-2,53</td><td>0,0163</td><td>5,8</td><td rowspan="2">1,16
4,52 } 5,68</td><td rowspan="2">0,018</td></tr>
<tr><td>2</td><td>8,914</td><td>-0,00536</td><td>-1,92</td><td>0,0632</td><td>22,6</td></tr>
<tr><td>3</td><td>8,916</td><td>14</td><td>-0,00336</td><td>-1,20</td><td>0,1942</td><td>69,5</td><td>13,9</td><td>0,0007</td></tr>
<tr><td>4</td><td>8,918</td><td>27</td><td>-0,00136</td><td>-0,485</td><td>0,3546</td><td>126,7</td><td>25,34</td><td>0,1086</td></tr>
<tr><td>5</td><td>8,920</td><td>24</td><td>+0,00064</td><td>+0,229</td><td>0,3885</td><td>138,8</td><td>27,76</td><td>0,5092</td></tr>
<tr><td>6</td><td>8,922</td><td>18</td><td>+0,00264</td><td>+0,943</td><td>0,2558</td><td>91,3</td><td>18,26</td><td>0,0037</td></tr>
<tr><td>7</td><td>8,924</td><td rowspan="2">9
2 } 11</td><td>+0,00464</td><td>+1,66</td><td>0,1006</td><td>36</td><td rowspan="2">7.2
1.72 } 8.92</td><td rowspan="2">0,4850</td></tr>
<tr><td>8</td><td>8,926</td><td>-0,00664</td><td>+2,37</td><td>0,0241</td><td>8,6</td></tr>
</table>