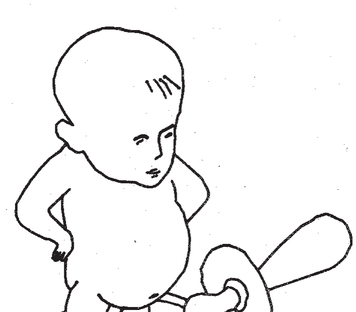
НАЧАЛА ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ



Государственный комитет Российской Федерации
по высшему образованию

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Физический факультет
Кафедра общей физики

ИЗМЕРИТЕЛЬНЫЙ ПРАКТИКУМ

Б.А.Князев, В.С.Черкасский

НАЧАЛА ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

электронный учебник
и программа обработки данных
для начинающих

Учебное пособие

Новосибирск
1996

УДК 53.08+378.147-388
ББК В3 с 13

Князев Б.А., Черкасский В.С. Начала обработки экспериментальных данных. Электронный учебник и программа обработки данных для начинающих: Учебное пособие // Новосиб. ун-т. Новосибирск, 1996. 93 с.

Предназначено для студентов естественно-научных специальностей, выполняющих лабораторные работы в учебных практикумах. Для его чтения достаточно знаний математики в объеме средней школы, но оно может быть полезно и тем, кто уже изучил математическую статистику, поскольку исходным моментом в нем является не математика, а эксперимент. Во второй части пособия подробно описан реальный эксперимент — от появления идеи и проблем постановки эксперимента до получения результатов и обработки данных, что позволяет получить менее формализованное представление о применении математической статистики. Пособие дополнено обучающей программой, которая позволяет как углубить и уточнить знания, полученные в методическом пособии, так и проводить собственно обработку результатов лабораторных работ. Приведен список литературы для желающих углубить свои знания в области математической статистики и обработки данных.

Рецензент
кандидат физ.-мат наук Задорожный А.М.

Печатается по решению кафедры общей физики физического факультета НГУ.

Содержание

1 ФИЗИКА И ИЗМЕРЕНИЯ

Физика по своей сути является экспериментальной наукой, а основу научной и инженерной деятельности составляет получение, обработка и интерпретация экспериментальных данных. Полученные в результате экспериментов численные значения могут быть далее использованы (в практике или теории) лишь в том случае, если они достоверны. Ясно, однако, что любая величина может быть измерена лишь с некоторой определяемой разными факторами точностью. Если вы возьмете любую экспериментальную работу, посвященную измерениям какой-либо величины, то обнаружите, что в ней обязательно встречаются словосочетания "среднее значение", "среднеквадратичная ошибка измерений", "дисперсия", "доверительная вероятность" и др.

Все выделенные термины являются терминами математической статистики и характеризуют измеренную величину и точность измерения. Важно, чтобы начиная с самых первых шагов своей экспериментальной деятельности (уже с самой первой учебной лаборатории) студент понимал что такое точность измерений, как правильно спланировать эксперимент, чтобы получить результат с требуемой точностью, какова степень достоверности полученных результатов. Все это - основа культуры каждого физика и инженера.

2 О ДАННОМ ПОСОБИИ

2.1 Как пользоваться данным пособием

Десять лет назад по инициативе проф. И.Н.Мешкова один из авторов прочитал студентам первого курса физического факультета Новосибирского университета, приступающим к работе в измерительном практикуме, отдельную лекцию, посвященную основам методов обработки экспериментальных данных. В этой лекции была сделана попытка изложить основы математической статистики, исходя из качественных соображений, избегая математических выводов и ограничив объем материала и степень сложности изложения уровнем, доступным для хорошо подготовленного выпускника средней школы. При данном подходе все законы статистики можно сформулировать достаточно строго, но как некие постулаты. Впоследствии из этой лекции выросло методическое пособие [1] для студентов первого курса. Это пособие, выдержавшее два издания, оказалось полезным и для других лабораторных практикумов, выполняемых на кафедре общей физики в течение первых трех курсов.

В настоящее время, в связи с модернизацией измерительного практикума и внедрением в практикумы персональных компьютеров, представляется необходимым создание новой версии пособия по обработке результатов измерений, выполненного в стиле электронного учебника, и сопровождающегося специально разработанной программой статистической обработки данных, адаптированной к уровню знаний и потребностям студентов младших курсов. Текст учебника и программа STAT содержатся на дискете и ориентированы на использование персональных компьютеров IBM PC, кроме того, теоретический материал издан в виде брошюры, которая может использоваться независимо. Мы попытались сохранить положительные качества предыдущей книги и в то же время использовать для обучения новые возможности, обусловленные широким распространением персональных компьютеров. Наиболее важной особенностью пособия, отличающей его от всех существующих, является демонстрация применения всех рассматриваемых методов для обработки реально выполненных экспериментов. Эти примеры вынесены в печатном варианте пособия в *Приложения* и в *Дополнение*, тогда как в электронном варианте читатель может обращаться к ним в любом месте.

Используя электронный учебник, вы можете в диалоговом режиме ознакомиться с основными терминами и идеями математической статистики и обработки данных, получить наглядную графическую информацию о функциях распределения, доверительной вероятности, о роли того или иного параметра, а также другую информацию. Меню программы написано на английском языке, однако соответствующие справки на русском языке всегда помогут найти русский эквивалент термина. Возможности программы не ограничиваются обучением и демонстрациями — она позволяет в пределах, заданных содержанием книги, проводить обработку результатов эксперимента. Соответствующие массивы данных могут вводиться либо непосредственно с клавиатуры, либо импортироваться извне. Эти возможности учебника пригодятся вам при обработке результатов лабораторных работ.

Мы надеемся, что эти книга и программа помогут вам сделать первые шаги в понимании некоторых основных идей математической статистики и обработки экспериментальных данных. Не следует думать, что после изучения этого комплекта вы будете знать о статистике все. В действительности вы узнаете скорее лишь набор некоторых рецептов относительно обработки данных. Наше изложение иногда умышленно упрощено, хотя все определения и понятия мы старались ввести максимально строго. Математические основания теории станут ясны вам полностью лишь после изучения курса теории вероятностей и математической статистики. Тем не менее, объем приведенного здесь материала позволит вам правильно обрабатывать данные, полученные в лабораторных практикумах, а использование программы существенно облегчит рутинные расчеты.

2.2 Примеры статобработки данных

Для того чтобы облегчить еще неопытному читателю восприятие большого объема нового материала и приблизить изложение к реальности, мы решили сопроводить все изложение реальными экспериментальными данными, полученными в Институте ядерной физики СО РАН на установке ГОЛ-3, а также данными из опубликованной в научном журнале работы[2]. Эта работа выбрана не потому, что она является "образцом" метрологической работы, а потому, что один из авторов настоящего пособия является соавтором этой публикации и имеет в лабораторном журнале полный набор исходных экспериментальных данных, которые никогда в полном объеме не приводятся ни в одной публикации. Это обстоятельство позволило приоткрыть завесу над "кухней экспериментатора" и показать как много остается "за кадром" даже небольшой трехстраничной работы. Мы воспроизводим эту работу полностью в виде дополнения к настоящему учебнику, а процедуру последовательной обработки данных, использованную в этой работе, описываем в его соответствующих разделах. Отметим, что в программе STAT вызов соответствующего примера на экран осуществляется с помощью меню только по желанию работающего.

2.3 Рекомендуемые литература и программное обеспечение

Для получения дополнительной информации о методах обработки экспериментальных данных мы рекомендуем вполне доступные для начинающих книги [3]–[5]. Для более глубокого знакомства с предметом следует обращаться к книгам [6]–[12] и др., а также к учебникам матстатистики (например [13]). Следует также знать, что существуют различные "фирменные" программы статистической обработки (GRAPHER, STATGRAPH и др.), для практического использования которых требуется достаточно высокая квалификация.

3 ХАРАКТЕРИСТИКИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

3.1 Что является целью измерений?

Естественный ответ на этот вопрос — "получение истинного значения физической величины" — вызывает следующий вопрос, а что такое "истинное значение"? Если исключить фундаментальные физические постоянные типа скорости света ($c = (2.99792458 \pm 0.00000001) \cdot 10^{10}$см/с [1], постоянной Планка

[1]Здесь и везде далее в тексте будет использоваться американский стандарт написания десятичных чисел — в качестве разделителя целой и дробной частей будет использоваться точка, как это

$\hbar = (1.054589 \pm 0.000005) \cdot 10^{-27}$эрг $\cdot$ с, заряда электрона $e = (1.602189 \pm 0.000006) \cdot 10^{-19}$К и др.), которые, вероятно, имеют точные значения (приведенный разброс связан с ограниченной точностью измерений), все остальные величины, в принципе, могут быть определены только как некоторое *математическое ожидание* $M(X)$.

Действительно, такая величина, как скорость радиоактивного распада атомов определенного сорта, определяется только как среднее значение вследствие случайности времени распада каждого отдельного атома, приводящей к флуктуациям числа распадов в единицу времени. Проводя измерения числа распадов в течение ограниченного времени измерения, мы можем вычислить только некоторое *выборочное среднее*, отличающееся от математического ожидания.

Другие измеряемые величины, такие как сопротивление провода, скорость звука в газе, показатель преломления стекла, также не могут иметь "истинного значения" вследствие флуктуаций во времени и пространстве плотности, температуры и других характеристик сред. Более того, даже такой простой вопрос, как измерение длины стержня, приводит к целой последовательности вопросов:

1. Постоянна ли длина стержня в процессе измерения?
2. Каково качество обработки торцов стержня, а если торец не прямой, то что считать за конец?
3. Если с некоторой точностью (какой?) мы можем считать торец гладким, то параллелен ли ему второй торец? И т.д.

Таким образом, факторы, связанные со статистическими флуктуациями, а также с природным разбросом или распределением измеряемой величины по измеряемому объекту, ведут к разбросу численных значений измеряемой величины, получаемых в разных сериях измерений.

Но даже если максимально зафиксировать параметра среды так, чтобы флуктуации измеряемого параметра стали очень малы, имеются другие источники разброса, связанные с неточностями самих приборов, внешними помехами (вибрация, электрические наводки и т.п.), ошибками экспериментатора. Эти ошибки могут носить как статистический, так и систематический характер. В результате измеряемые в серии опытов значения физической величины лежат в определенном интервале. В этом случае говорят о *распределении* измеряемой величины. Рассмотрим все эти вопросы подробнее.

принято практически во всех компьютерных программах.

3.2 Ошибки измерений

В табл.1 приведены результаты измерения роста студентов 1-го курса физического факультета, проведенные в 1993 г. на первой лекции по методам физического эксперимента, а в табл.2 приведены результаты измерения напряжения (в вольтах) на ячейке Керра в разных экспериментах при определении постоянной Керра воды [2].

Таблица 1. Результаты измерений роста студентов 1-го курса

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	177	178	171	189	189	189	176	183	176	175	176	179
i	13	14	15	16	17	18	19	20	21	22	23	24
x_i	181	171	154	164	164	165	173	173	184	188	188	176
i	25	26	27	28	29	30	31	32	33	34	35	36
x_i	174	174	176	177	177	178	176	173	182	177	159	178
i	37	38	39	40	41	42	43	44	45	46	47	48
x_i	186	173	180	171	166	177	173	177	192	175	181	183
i	49	50	51	52	53	54	55	56	57	58	59	60
x_i	159	166	181	160	179	175	185	186	170	172	168	182
i	61	62	63	64	65	66	67	68	69	70	71	72
x_i	184	166	175	171	179	178	187	176	152	169	180	192
i	73	74	75	76	77	78	79	80	81	82	83	84
x_i	180	178	184	187	175	183	179	171	171	177	159	1799
i	85	86	87	88	89	90	91	92	93	94	95	96
x_i	159	181	185	177	175	178	169	170	175	175	183	173
i	97	98	99	100	101	102	103	104	105	106	107	108
x_i	178	183	177	186	190	185	170	178	169	165	182	186
i	109	110	111	112	113	114	115	116	117	118	119	120
x_i	184	176	169	183	168	181	179	180	183	174	175	177
i	121	122	123	124								
x_i	176	169	176	178								

Отклонение измеренного значения x определяемой величины от истинного ее значения $M(X)$ (здесь и далее под "истинным значением" понимается математическое ожидание величины) называется "погрешностью" или "ошибкой" (error) i-го измерения Δx_i. Термин "ошибка" в статистике не несет в себе эмоциональной негативной окраски, а имеет смысл разброса значений, присущего данному явлению или измерительному устройству (на это указывает и некоторое смысловое различие между словами error и mistake в английском языке). По-видимому, наиболее адекватным (по смыслу) переводом на русский язык было бы слово "отклонение". Ошибки измерений возникают вследствие того, что измерительные при-

Таблица 2. Результаты обработки керрограмм для He-Cd лазер $\lambda = 4416$ Å

Номер эксп.	φ					
	$\pi/2$ (макс.)	π (мин.)	$3\pi/2$ (макс.)	2π (мин.)	$5\pi/2$ (макс.)	3π (мин.)
1	51.5	69.7	86.5	98.7	110.3	123.1
2	52.0	69.1	84.8	98.1	111.0	123.4
3	49.4	68.3	84.2	98.0	110.5	122.2
4	50.0	68.3	82.5	97.0	110.0	122.0
5	50.0	69.2	83.2	98.7	111.1	123.1
6	48.6	68.3	83.8	97.0	111.1	122.6
7	47.8	67.7	83.5	97.0	110.0	122.7
8	51.2	70.4	84.8	98.5	111.4	122.1
9	47.8	67.8	82.7	97.5	110.7	122.1
10	50.5	69.9	85.0	98.8	111.5	123.1
11	45.5	69.9	86.0	98.2	110.8	123.8
12	50.0	69.8	84.0	98.5	110.8	123.5
$\overline{U}$	49.52	69.03	84.25	98.0	110.8	122.8
S	0.56	0.28	0.37	0.2	0.15	0.19
$\delta\overline{U}$	1.23	0.61	0.81	0.44	0.33	0.42

боры имеют ограниченную точность, не всегда можно учесть влияние неконтролируемых экспериментальных условий. Иногда сама природа изучаемого явления служит основной причиной разброса измерений. Ошибки измерений принято разделять на два типа — систематические и случайные.

3.2.1 Систематические ошибки

Систематическими называют такие ошибки, которые вызываются факторами, действующими одинаковым образом при многократном повторении одних и тех же измерений. При измерении роста студентов на лекции систематической ошибкой было измерение роста в обуви. При измерении напряжения на ячейке Керра и обработке соответствующих результатов на первом этапе обработки не была учтена... задержка, связанная с временем пролета электронов в фотоумножителе! Это время составляло 12.5 наносекунд (всего-то $1.25 \cdot 10^{-8}$ секунды!) и, тем не менее, вносило значительную ошибку в результат (подробнее см. *Дополнение*).

При измерении глубины эрозии пластинки графита под действием плазменного пучка на установке ГОЛ-3 причиной систематической ошибки могло быть неправильное определение нулевого уровня поверхности, наличие на этой поверхности царапин и т.п.

Систематические ошибки, в принципе, могут быть устранены или учтены, хотя нахождение источников этих ошибок в конкретном опыте чрезвычайно сложное дело. Любой эксперимент не гарантирован от наличия неучтенной систематической ошибки.

3.2.2 Случайные ошибки

Случайные ошибки (отклонения) всегда присутствуют в эксперименте. При отсутствии систематических ошибок они служат причиной разброса результатов повторных измерений как между собой, так и относительно истинного значения измеряемой величины. Природа случайных ошибок может быть различной. Например, при измерении длины стрежня, которая точно определена (с учетом сказанного выше), к случайным ошибкам приводит вибрация стержня, флуктуации нулевого положения измерительного прибора и т.п. В опытах по измерению скорости радиоактивного распада ядер сама измеряемая величина определена лишь статистически (как некоторое среднее значение), и флуктуации измеренного числа распадов в равные промежутки времени будут наблюдаться даже при идеально точной аппаратуре. Увеличивая число измерений и используя формулы теории ошибок, можно получить достаточно точную оценку случайной ошибки.

Случайные ошибки последовательных измерений как правило независимы и характеризуются тем или иным *законом распределения* (см.ниже). Они обладают свойством концентрации, т.е. малые по абсолютной величине случайные ошибки встречаются чаще, чем большие.

Отметим еще раз, что разделение ошибок на случайные и систематические в достаточной мере условно. В зависимости от обстоятельств один и тот же фактор может приводить либо к случайным, либо к систематическим ошибкам. Например, в экспериментах по измерению постоянной Керра вариации толщины водяного конденсатора можно учесть в величине случайной ошибки, используя разброс толщины изолятора, приведенный в заводском сертификате. Если же измерить реальный профиль использованных изоляторов, то можно вычислить и вычесть из найденного среднего связанную с этим систематическую ошибку, уменьшив за счет этого погрешность измерений.

3.2.3 Промахи

Следует особо выделить такой вид ошибки, как грубый просчет — промах. Под промахом понимается ошибка измерения, сделанная вследствие недосмотра экспериментатора, или вызванная неисправностями аппаратуры. Так, например, неправильно записанный отсчет, замыкание электрической цепи и т.п. являются промахами, которых следует по возможности избегать. Как правило, грубые ошибки легко обнаруживаются. Такие измерения следует отбрасывать, хотя при этом желательно определить причину данного промаха.

3.3 Статистическая обработка результатов измерений

Систематические ошибки, присущие системе измерений, должны быть обнаружены и ликвидированы (или учтены при обработке данных). Оставшиеся необнаруженными систематические ошибки вносят в результаты измерений неизвестный сдвиг относительно истинного значения. Поскольку практически невозможно выявить все систематические ошибки, даже очень точные измерения разных авторов могут не совпадать. В примере с водяным конденсатором разброс значений постоянной Керра у разных авторов можно связать, в частности, с различием в методах очистки воды и наличием неконтролируемых примесей. Можно без преувеличения сказать, что мастерство экспериментатора состоит в искусстве обнаружения систематических ошибок.

Если вам удалось снизить до достаточно низкого уровня систематические ошибки, то точность измерений определяется теперь случайными погрешностями. Если же точность измерений определяется случайными ошибками, то она может быть подвергнута статистическому анализу. Обсудим далее, как следует оценивать точность значения некоторой средней величины X, полученной экспериментально.

3.4 Генеральная совокупность и выборка

Генеральной совокупностью называют полный набор всех возможных значений, которые может принимать *случайная величина* при бесконечном числе испытаний. Генеральную совокупность можно описать с помощью *распределения вероятности* появления в отдельном испытании некоторой конкретной величины x_i . Набор n значений величин x_i, полученный из генеральной совокупности в результате конечного числа испытаний N, называют *выборкой* объема N. Цель *статистической обработки* набора величин x_i заключается в попытке описать характеристики генеральной совокупности как можно точнее по отдельной выборке.

3.5 Кривая распределения результатов

Проведем серию из N измерений и получим некоторый набор (см.табл. 1) значений $x_1, \ldots, x_i, \ldots, x_n$. Это — так называемая выборка. По оси абсцисс (рис.1) будем откладывать полученные в отдельных измерениях значения x_i величины x. Разобъем ось x на равные интервалы Δx и подсчитаем число изме-

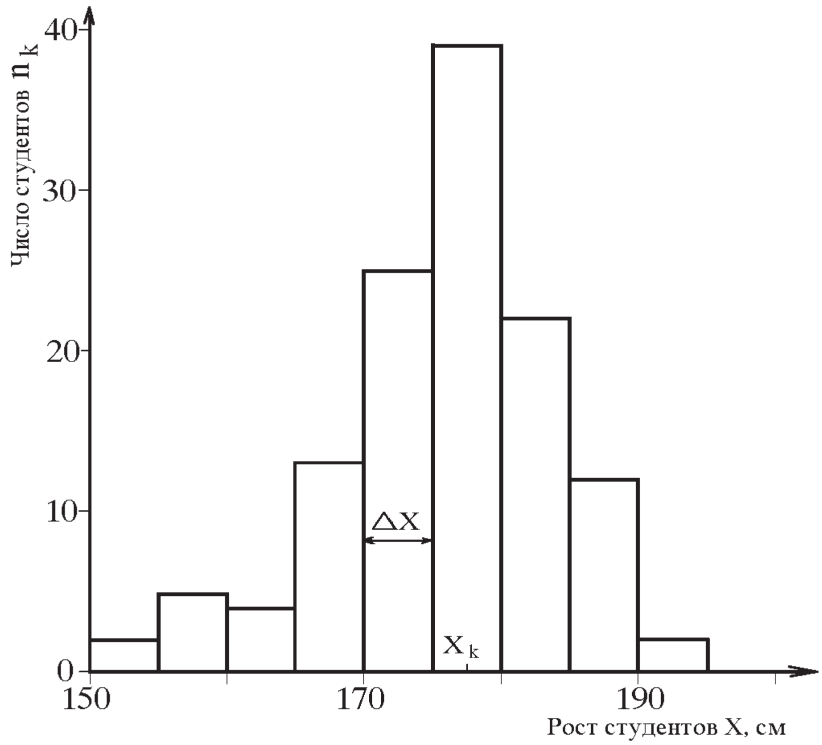


Рис. 1. Гистограмма распределения роста студентов 1-го курса физического факультета

рений Δn_k, в результате которых получены значения x, лежащие в интервале $x_k \pm 1/2\Delta x$ (здесь x_k — координата центра интервала на оси x). На каждом интервале построим прямоугольник высотой Δn_k и шириной Δx (точки, лежащие точно на границе интервала, будем всегда относить, например, в левый столбик). Диаграмму, полученную таким образом, называют гистограммой. Гистограммы могут быть построены как для непрерывных величин (например, скорость частиц), так и для дискретных (например, число радиоактивных распадов в секунду). Сумма высот прямоугольников гистограммы равна полному числу экспериментов в данной серии.

Если мы имеем две выборки с разным полным числом событий, то амплитудные значения гистограмм, естественно, будут различны, и их трудно будет сравнивать. Поэтому более удобно представлять гистограмму в несколько ином виде. Отложим по оси ординат не Δn_k, а величину

$$\frac{1}{N}\frac{\Delta n_k}{\Delta x_k},$$

где $N = \sum n_k$. В этом случае произведение высоты на Δx, т.е. площадь каждого столбика, имеет размерность вероятности попадания результата отдельного измерения в данный интервал Δx, а суммарная площадь под всей гистограммой, как и

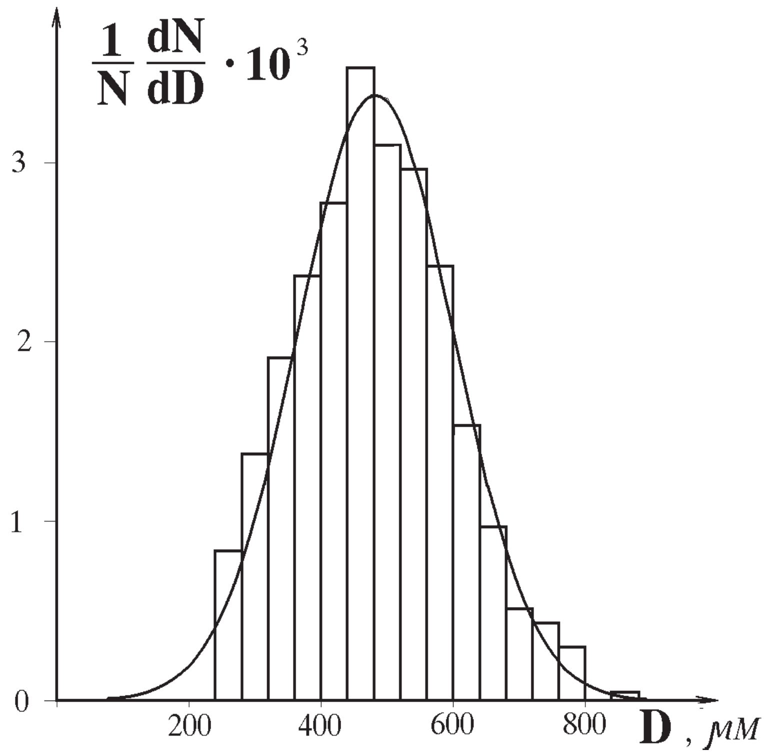


Рис. 2. Гистограмма распределения глубины испарения графитовой пластины в эксперименте на ГОЛ-3

должно быть в этом случае, равна единице:

$$\sum \frac{1}{N}\frac{\Delta n_k}{\Delta x_k}\cdot \Delta x_k = 1. \tag{1}$$

Теперь нетрудно сравнивать гистограммы с сильно отличающимся числом измерений.

На рис.2 представлена такая гистограмма для результатов измерений глубины испарения графитовой пластины под действием высокотемпературного плазменного потока (см. *Приложение* **??**) при разбиении промежутка на оси x на 25 интервалов.

Если число измерений N достаточно велико, то ширину интервала можно сделать очень малой (при этом в каждом интервале будет еще достаточно много отсчетов). Тогда в пределе вместо гистограммы мы получим график типа показанного на рис. 3, на котором по оси ординат отложена величина $f(x)$, имеющая размерность $[x^{-1}]$ и пропорциональная доле числа отсчетов n_k/n, попадающей в каждый интервал. Такой график называют кривой распределения, а функцию $f(x)$ называют *плотностью вероятности*.

3.6 Плотность вероятности

Смысл плотности вероятности заключается в том, что произведение $f(x)dx$ (см. рис. 3) дает долю полного числа отсчетов n, приходящуюся на интервал от x до $x+dx$ или, иными словами, вероятность того, что результат любого очередного измерения x_i будет иметь значение, лежащее в указанном интервале. Очевидно, что вероятность получить при измерении хоть какое-нибудь значение равна единице, и функция $f(x)$ должна удовлетворять соотношению

$$\int\limits_{-\infty}^{+\infty} f(x)dx = 1. \qquad (2)$$

Отсюда видно, что размерность функции плотности вероятности $f(x)$ есть $[x]^{-1}$. Понятно, что бесконечные пределы интегрирования здесь имеют формальный смысл, а реально измеряемые величины должны быть ограничены некоторым разумным диапазоном.

В случае дискретной величины вместо f(x)dx используют вероятность p_i получить в результате измерения значение x_i. Естественно, что

$$\sum_{i=1}^{N} p_i = 1, \qquad (3)$$

где N— число интервалов дискретного распределения. Ниже в этом разделе мы будем говорить о непрерывных распределениях, предоставляя читателю самостоятельно преобразовать результаты на случай дискретных распределений.

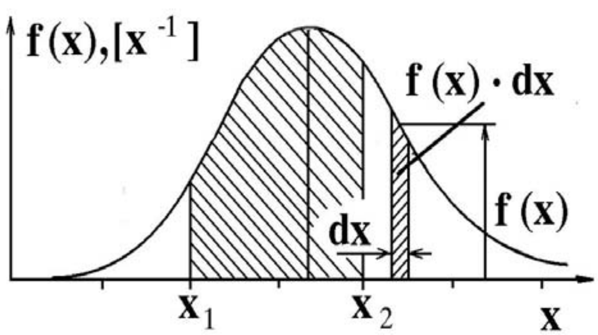


Рис. 3. Пример кривой распределения

3.7 Функция распределения, математическое ожидание, дисперсия и другие моменты

Вероятность попадания измеряемой величины (в данном измерении) в интервал от $-\infty$ до x называют *функцией распределения* или интегральной функцией распределения:

$$F(x) = \int_{-\infty}^{x} f(z)dz \tag{4}$$

Понятно, что плхотность вероятности $f(x)$ равна просто производной функции распределения $F(x)$ в точке x.

Если проинтегрировать плотность вероятности в пределах от x_1 до x_2 (рис. 3), то полученная величина будет представлять вероятность $\mathcal{P}$ того, что результат отдельного измерения будет лежать в интервале x_1, x_2:

$$\mathcal{P}(x_1 < x < x_2) \equiv \int_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1). \tag{5}$$

Эта вероятность $\mathcal{P}$, равная разности интегральных функций распределения в соответствующих точках, равна заштрихованной площади под кривой. При расширении пределов она стремится к единице. Смысл введения такой величины заключается в том, что при заданной функции распределения результат любого отдельного измерения с достоверностью $\mathcal{P}$ даст величину, лежащую внутри указанного интервала. Это определение будет нам полезно при определении понятия доверительного интервала и доверительной вероятности.

Если функция распределения известна, то *математическое ожидание* измеряемой величины, которое было бы получено при бесконечном числе измерений, вычисляется с помощью выражения

$$M(x) = \overline{x} = \int_{-\infty}^{+\infty} xf(x)dx, \tag{6}$$

а математическое ожидание функции $y(x)$ от непрерывной случайной величины x [14] есть

$$M(y(x)) = \overline{y(x)} = \int_{-\infty}^{+\infty} y(x)f(x)dx. \tag{7}$$

Здесь следует отметить, что в физической литературе очень часто функцией распределения называют плотность вероятности $f(x)$, подразумевая, что читатель понимает о чем идет речь.

Если в выражении (6) под интегралом стоит n-я степень x, то такая величина называется n-м моментом функции распределения. Функция распределения однозначно определяется всей совокупностью (вообще говоря, бесконечной) своих моментов. Наиболее часто используемые в теории ошибок и в физике распределения, как правило, определяются небольшим числом своих моментов. Помимо математического ожидания, важнейшим из таких моментов является второй момент или *дисперсия*[2], которая определяется выражением

$$\sigma^2 = D(x) = M((x - \overline{x})^2) = \int_{-\infty}^{+\infty} (x - \overline{x})^2 f(x) dx. \tag{8}$$

Поскольку дисперсия имеет размерность квадрата случайной величины, а это не всегда удобно, то вводится *среднее квадратическое отклонение* σ, которое представляет собой положительный квадратный корень из дисперсии

$$\sigma = \sqrt{\sigma^2} \tag{9}$$

3.8 Пример. Определение среднего роста человека

Предположим, перед нами стоит задача определить средний рост людей на Земле. Генеральной совокупностью в данном случае является набор чисел, соответствующих росту всех живущих на данный момент времени жителей земли, от младенцев до стариков. Если бы мы были в состоянии получить сведения о всех людях, то могли бы построить функцию плотности вероятности $f(x)$, имеющую размерность $1/$. Далее, используя выражение (6), можно было бы найти искомую величину X. К сожалению, проводя измерения, мы можем получить сведения только об ограниченном числе людей. Совокупность полученных данных как раз и представляет собой выборку, из которой мы будем пытаться найти "истинное значение", т.е. X.

Совершенно очевидно, что эта выборка должна быть "представительной" (репрезентативной), т.е. люди должны быть действительно *случайно* выбраны из всех человеческих групп. Нельзя для этой цели, например, выбрать роту солдат, ясельную группу детского сада или студентов 1-го курса. В выборке должны быть

[2] В литературе используются различные буквы и сочетания для обозначения дисперсии. В большинстве случаев ее обозначают так, как показано здесь.

представители всех наций (пигмеи и скандинавы), мужчины и женщины, представители всех имущественных сословий и т.п., пропорционально их доле в составе человечества. Этот простой пример показывает, с какой тщательностью нужно подходить к статистике. Вопрос определения функции распределения и оценки его параметров по экспериментальным (выборочным) данным является предметом *математической статистики*, некоторые аспекты которой будут рассмотрены далее.

4 НЕКОТОРЫЕ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И ИХ СВОЙСТВА

В данном разделе мы кратко рассмотрим лишь те функции распределения, с которыми вам придется встретиться в практикуме. Для более детального знакомства с функциями распределения и их свойствами можно использовать рекомендованную литературу.

4.1 Равномерное (прямоугольное) распределение

Если в интервале от x_1 до x_2 вероятность случайной величины $dp = f(x)dx$ не зависит от x, то такое распределение вероятности называется *равномерным* или прямоугольным:

$$f(x) = \begin{cases} 0, & \text{если } x < x_1, \\ \frac{1}{x_2 - x_1}, & \text{если } x_1 \le x \le x_2, \\ 0, & \text{если } x > x_2. \end{cases} \tag{10}$$

Прямоугольному распределению подчиняется, например, светимость звездного неба в единичном интервале телесного угла.

4.2 Распределение Пуассона

Распределение Пуассона - это дискретное распределение, описывающее случайные процессы, в которых вероятность отдельного события мала и постоянна по величине $p \ll 1$ (практически, достаточно, чтобы p было меньше 0.1). А само распределение показывает вероятность того, что за выбранный промежуток времени t произойдет n событий, причем среднее число событий в единицу времени в этом интервале равно μ. Примером процесса, описываемого таким распределением, является число распадов α-частиц за время t для очень большого числа радиоактивных атомов с большим периодом полураспада (т.е. с малой вероятностью

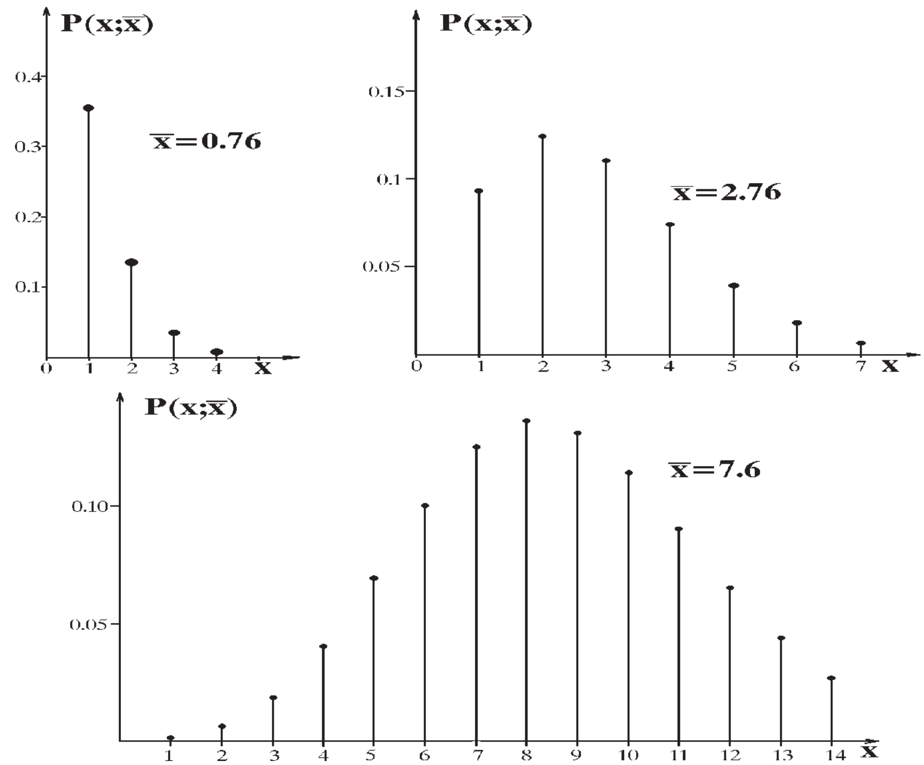


Рис. 4. Распределения Пуассона с различной величиной математического ожидания $\overline{x}$

одного распада). В этом случае, естественно, за время измерения число атомов не должно существенно измениться.

Распределение Пуассона можно записать в виде

$$P_n(t,\mu) = \frac{(\mu t)^n}{n!} e^{-\mu t}. \tag{11}$$

Если заменить произведение μt на параметр μ, то распределение Пуассона запишется в более привычном виде

$$P_n(\mu) = \frac{\mu^n}{n!} e^{-\mu}. \tag{12}$$

Здесь параметр μ может принимать любые положительные значения, тогда как n - только целые. Замечательным свойством распределения Пуассона является то, что математическое ожидание

$$M(n) = \overline{n} = \sum_n n P_n(\mu) = \mu \tag{13}$$

равно дисперсии

$$\sigma^2 = \sum_n (n-\overline{n})^2 P_n(\mu) = \mu. \tag{14}$$

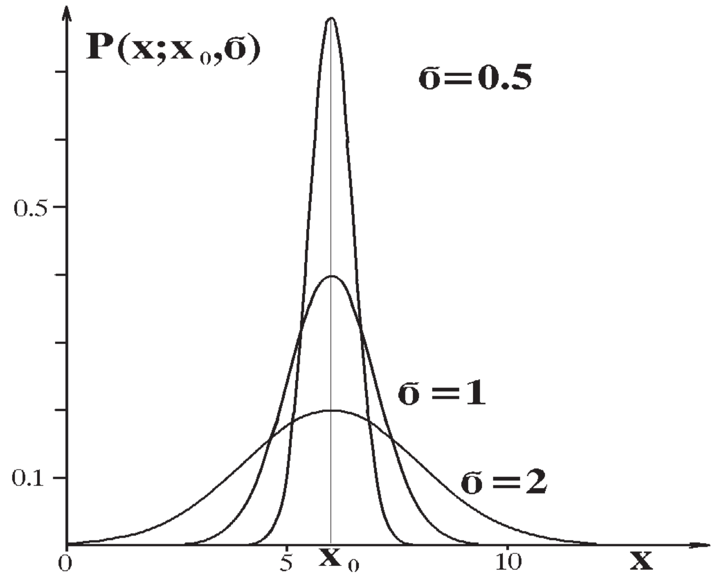


Рис. 5. Плотности вероятностей для нормального распределения (Гаусса) при значениях параметра $\sigma = 0.5; 1; 2$

Распределение Пуассона для трех различных значений среднего приведены на рис. 4. Видно, что при больших $\overline{n}$ распределение становится более симметричным. При дальнейшем возрастании $\overline{n}$ распределение Пуассона хорошо аппроксимируется гауссовым распределением.

4.3 Распределение Гаусса. Функция ошибок

Распределение Гаусса, или *нормальное распределение*(см. рис. 5), это распределение, хорошо описывающее разброс непрерывной случайной величины при большом числе независимых случайных погрешностей. При измерениях непрерывных величин экспериментатор, обычно, *a priori* полагает, что распределение результатов - гауссово. В последующем мы покажем, что имеются объективные критерии проверки этого предположения.

Функция плотности вероятности для распределения Гаусса имеет вид

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}}. \qquad (15)$$

Она является, в отличие от распределения Пуассона, функцией двух параметров - X и σ. Нормальное распределение симметрично относительно X, а его ширина пропорциональна σ (см. рис. 5). Чем точнее измерения, тем плотнее вблизи среднего значения лежат результаты отдельных измерений, т.е. величина σ меньше.

Если перейти к переменной $u = x - X$ и проинтегрировать выражение (15) в пределах от $-u$ до $+u$ (см. рис. 6), то полученная величина будет представлять

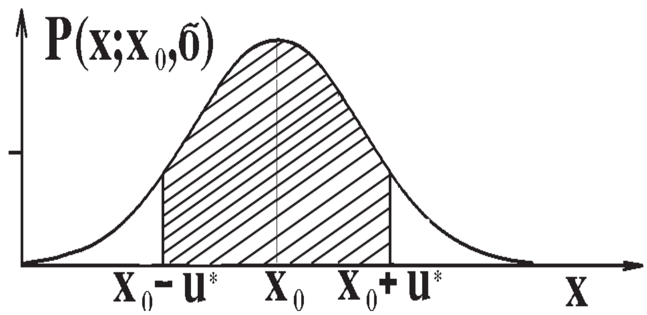


Рис. 6. Доверительный интервал $x_0 - u^* \leq x \leq x_0 + u^*$

вероятность P того, что результат измерения будет лежать в интервале от $X - u^*$ до $X + u^*$ (*доверительный интервал*).

$$P(-u^* < u < u^*) \equiv P(|u| < u^*) = \int_{-u^*}^{u^*} f(u)du = 2\Phi(\frac{u^*}{\sigma}). \qquad (16)$$

Здесь символ $\Phi(\frac{u^*}{\sigma})$ обозначает интеграл вероятностей:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{\theta^2}{2}} d\theta,\ \theta = \frac{(x - X)}{\sigma}. \qquad (17)$$

В табл. 3 (*Приложение* В) приведены его численные значения в зависимости от параметра $t = |u^*|/\sigma$.

Согласно таблице, вероятность того, что случайная величина отклонится от среднего значения не более, чем на σ, равна 0.683, на 2σ - 0.955, на 3σ - 0.997. Отсюда ясно происхождение эмпирического правила — считать измерения, результаты которых выходят за 3σ , промахами и исключать из рассмотрения.

5 ОЦЕНКА ПОГРЕШНОСТИ ПРЯМЫХ И КОСВЕННЫХ ИЗМЕРЕНИЙ

5.1 Случайная ошибка и ее описание

Как было отмечено ранее, ошибки делятся на случайные и систематические. Есть еще промахи, но о них мы уже говорили. Кроме того, измерения бывают однократные и многократные, прямые и косвенные и т.д. Как уже отмечалось выше, все результаты измерения содержат случайную погрешность, а значит, при обработке таких результатов необходимо использовать аппарат и методы теории вероятности. Поскольку теория вероятности описывает характеристики генеральной совокупности, то ее результаты непосредственно применимы при обработке результатов эксперимента, только если мы имеем дело с генеральной совокупностью. Но мы в реальной практике всегда имеем дело с выборкой - т.е. с конечным числом измерений. Для описания таких процессов существует наука, которая называется *математической статистикой*. Суть ее в том, что она описывает поведение случайной величины и позволяет по выборке судить о генеральной совокупности и о ее параметрах. Но проблема состоит в том, что мы *априори* не знаем функцию распределения ошибок измерения даже в простейшем случае прямого измерения. Мы должны одновременно решать две взаимосвязанные задачи — определять "истинное" значение измеряемой величины и вычислять погрешность. Следует хорошо усвоить, что и измеренное нами значение и наши оценки являются величинами *случайными*, причем относительно характера распределения этих величин мы можем выдвигать более или менее обоснованные гипотезы, которые, вообще говоря, подлежат проверке в том же или другом эксперименте. Имеются как минимум две гипотезы, на которых чаще всего базируется теория ошибок:

1. Погрешность непосредственно измеряемой величины аддитивна, т.е. измеренное значение $x_i = X + \delta x$, где X — "истинное" значение, а δx — погрешность данного измерения (случайная величина). Это допущение настолько очевидно, что зачастую его переносят и на те области, в которых оно несправедливо. Ключевым словом в допущении является *непосредственно измеряемая* — если же речь идет о результате косвенного измерения, например, измерении сопротивления по измерению тока и напряжения, то очевидно, что относительно сопротивления мы не можем сделать такого вывода, поскольку
$$R = \frac{U}{I} = \frac{U_0 + \delta U}{I_0 + \delta I} \neq R_0 + \delta R.$$
(Подробнее этот вопрос обсуждается в *Дополнении*).

2. Предположение о характере функции распределения случайной величины — погрешности измерения. Это прежде всего гипотеза о нулевом математическом ожидании погрешности и гипотеза о конечности дисперсии этого распределения. Как правило, хотя и не всегда, сюда же добавляется предположение о нормальном (гауссовом) виде функции распределения.

Если первая гипотеза представляется почти очевидной, то использование второй чаще всего обосновывается центральной предельной теоремой [13], суть которой состоит в том, что распределение суммы случайных независимых величин, имеющих близкие распределения, стремится при увеличении их числа к распреде-

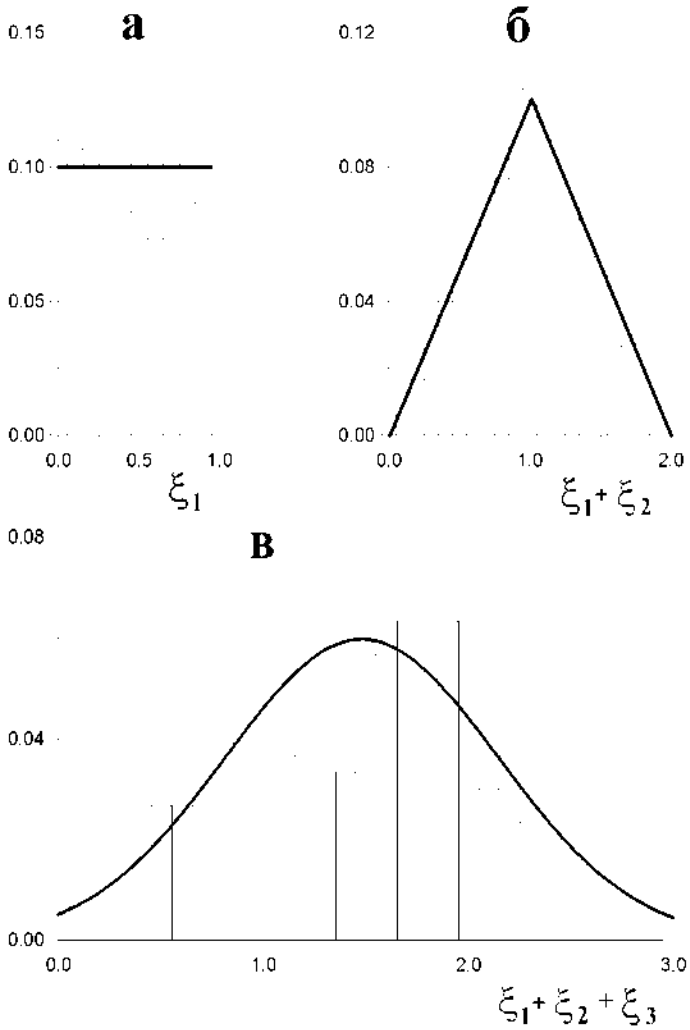


Рис. 7. Гистограммы распределения: а — равномерно распределенной случайной величины; б — суммы двух равномерно распределенных случайных величин; в — суммы трех равномерно распределенных случайных величин

лению Гаусса. Не уточняя более строго эту формулировку и не приводя доказательства, рассмотрим пример, представленный на рис. 7.

Используя генератор случайных чисел, была сгенерирована последовательность из 300 случайных чисел, равномерно распределенных на интервале $[0 \div 1]$. Гистограмма такого экспериментально полученного распределения показана на рис. 7,а. Сплошной горизонтальной линией показана теоретическая зависимость. На рис. 7,б показана гистограмма распределения суммы двух одинаковых равномерных распределений. Для его получения были сгенерированы две последовательности случайных чисел, распределенных равномерно, после чего новая случайная величина была образована суммированием первых двух последовательностей почленно[3]. Можно показать теоретически, что такая сумма должна иметь "треугольное" распределение, что и видно на рисунке. Полученное таким же способом распределение суммы трех равномерно распределенных величин представлено гистограммой на рис. 7,в и весьма похоже на распределение Гаусса. Там же на рисунке показано распределение Гаусса, построенное по первым двум моментам выборки. Как видно из этого примера, распределение всего трех случайных независимых величин весьма близко к распределению Гаусса.

В реальности, при проведении экспериментов таких случайных составляющих у погрешности гораздо больше, чем три, так что при рассмотрении случайных погрешностей, как правило, есть основания считать распределение гауссовым.

5.2 Оценка истинного значения искомой величины $\hat{x}$ и оценка ошибки измерения $\hat{\Delta x}$ по экспериментальным данным

При обработке результатов измерения величины, которая имеет определенное значение, но в результате влияния различных случайных факторов измеряется нами с некоторой случайной ошибкой, возникает задача — используя *конечный* набор эмпирических данных x_i, полученных в выборке из n измерений, найти "наилучшее" значение оценки $\hat{x}$ "точного" значения измеряемой велины X и определить точность наших измерений. В курсе математической статистики доказывается, что наилучшей оценкой величины X, непосредственно измеряемой и удовлетворяющей пунктам 1-2 предыдущего параграфа, является *среднее значение* выборки

$$\hat{x} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{18}$$

[3]То-есть первый член одной последовательности с первым членом второй последовательности, второй со вторым и т.д.

Поскольку мы предположили, что распределение ошибок (отклонений измеренных значений от истинного $X - x_i$) гауссово, то для оценки величины этой погрешности хорошо бы знать дисперсию σ^2 этого распределения. Как было показано ранее (см. параграф 4.3, рис. 6 и уравнение(16)) , вероятность того, что случайная величина x_i отклонится от среднего значения не более, чем на величину среднего квадратического отклонения σ, равна 0.683, на 2σ — 0.955, на 3σ — 0.997. Наилучшей оценкой [4] *среднего квадратического отклонения* является величина

$$\hat{\sigma} = s_n = \sqrt{\frac{1}{(n-1)} \sum (x_i - \overline{x})^2}. \tag{19}$$

5.3 Точность определения величины X. Среднеквадратичная ошибка среднего и распределение Стьюдента

Хотя оценка величины погрешности одного измерения x_i и представляет интерес (см. предыдущий параграф), тем не менее гораздо важнее знать с какой точностью значение $\hat{x}$, найденное нами из некоторой выборки, соответствует истинному значению искомой величины X. Поскольку $\hat{x}$ найдено из ограниченного числа измерений, то повторяющиеся серии измерений давали бы нам новые значения $\overline{x}$. То есть, эмпирические средние тоже являются случайной величиной и их поведение можно описать некоторой функцией распределения относительно величины математического ожидания со своей дисперсией s_x. Теория показывает, что если $\overline{x}$ определено из n измерений, то

$$s_x = \frac{\sigma}{\sqrt{n}}. \tag{20}$$

Это выражение, однако, невозможно использовать, т.к. σ нам неизвестна. Вместо σ в выражении(20) используется оценка среднего квадратического отклонения, а полученную величину s_x мы далее будем называть *стандартной ошибкой* [5].

Выражение для нее

$$s_x = \frac{s_n}{\sqrt{n}} \tag{21}$$

содержит только измеряемые экспериментально величины. Легко видеть, что стандартная ошибка s_x, в отличие от стандартного отклонения s_n, уменьшается с ро-

[4] Иногда при упоминании величины s_n слово “оценка” опускается.

[5] В разных книгах эту величину называют еще *эмпирическим стандартом* или *среднеквадратичной ошибкой среднего*

стом числа измерений приблизительно как $1/\sqrt{n}$, т.е. точность результата возрастает с ростом числа испытаний. Естественно, однако, что нет смысла увеличивать количество измерений после того, как величина s_x станет сравнима с величиной систематической ошибки (например, величиной ошибки измерительного прибора).

Поскольку мы оценили дисперсию отклонения эмпирического среднего $\overline{x}$ от истинного значения величины X, то, используя распределение Гаусса для этой величины, можно определить вероятность того, что это отклонение лежит в диапазоне $\pm u^*$ (см. параграф 4.3). Однако чаще представляет интерес обратная задача: при выбранном уровне вероятности $\mathcal{P}$ найти величину доверительного интервала. В идеализированном случае, когда σ точно известна, можно задать необходимую доверительную вероятность и, используя таблицы для интеграла вероятностей и соотношение

$$2\Phi(t) = \mathcal{P},$$

определить из табл. 3 (*Приложение* B) соответствующее значение верхнего предела

$$t = t(\mathcal{P}).$$

Тогда мы можем утверждать, что с вероятностью $\mathcal{P}$ отклонение измеренного значения $\overline{x}$ от X не превышает следующей величины:

$$|\overline{x} - X| < t(\mathcal{P})\frac{\sigma}{\sqrt{n}}. \tag{22}$$

В реальном случае дисперсия генеральной совокупности неизвестна, и поэтому кроме выборочного среднего $\overline{x}$ используется выборочная оценка дисперсии s_n^2. В теории показывается, что случайная величина

$$t = \frac{\overline{x} - X}{s_n/\sqrt{n}}$$

распределена не по нормальному закону. Закон распределения этой величины называется $t-$распределением или распределением Стьюдента. Плотность вероятности этого распределения показана на рис. 8. Это распределение похоже на нормальное, но имеет несколько меньшую величину вблизи максимума (при $x = X$) и более длинные "хвосты". При $n \to \infty$ оно превращается в гауссово распределение.

Доверительная оценка при выбранном уровне статистической достоверности $\mathcal{P}$ в этом случае принимает вид

$$|\overline{x} - X| < t(\mathcal{P}, \nu)\frac{s_n}{\sqrt{n}}, \tag{23}$$

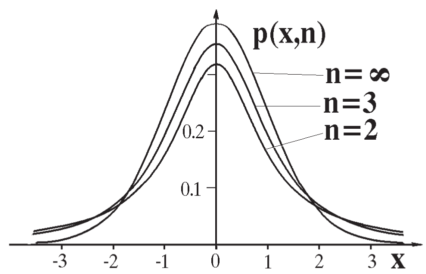


Рис. 8. Распределение Стьюдента для n = 2, 3 и бесконечности

где $\nu = n - 1$, а $s_n/\sqrt{n} \equiv s_x$.

Функция $t(\mathcal{P}, \nu)$, называемая далее коэффициентом Стьюдента, зависит теперь не только от $\mathcal{P}$, но и от числа измерений ($\nu = n - 1$). Таблицы значений этого коэффициента для нескольких уровней доверительной вероятности (называемых еще уровнями надежности) $\mathcal{P}$ и различных n приведены в табл. 4 (*Приложение* C).

5.4 Правила обработки прямого многократного измерения

При проведениии прямого измерения некоторой величины необходимо:

1. Провести многократные измерения при одних и тех же условиях и записать их в таблицу.

2. Рассчитать среднее значение по формуле
$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

3. Вычислить оценку дисперсии
$$\hat{\sigma}^2 = s_n^2 = \frac{\sum (x_i - \overline{x})^2}{(n-1)}.$$

4. Вычислить среднеквадратичную ошибку среднего
$$s_x = \frac{s_n}{\sqrt{n}}.$$

5. Задавшись требуемым уровнем доверительной вероятности $\mathcal{P}$, определить по табл. 4 (*Приложение* C) коэффициент Стьюдента $t(\mathcal{P}, n-1)$ и модуль доверительного интервала
$$\Delta x = s_x \cdot t(\mathcal{P}, \nu).$$

6. Округлив соответствующие результаты, записать ответ в виде

$$X = \overline{x} \pm \Delta x \quad \text{при доверительной вероятности } \mathcal{P}.$$

5.5 Оценка погрешности при косвенных измерениях

В большинстве экспериментов измеряется не непосредственно величина, подлежащая определению, а другая величина или ряд величин, зависящих от интересующей нас величины тем или иным образом. Например, сопротивление R может быть вычислено из закона Ома $R = U/I$ по измеренным значениям U и I. При таких измерениях, которые называются косвенными, необходимо уметь вычислять ошибку результата. Когда интересующая величина y является функцией измеряемых величин $x_1, x_2, \ldots, x_k$

$$y = f(x_1, x_2, \ldots, x_k), \tag{24}$$

а для случайных ошибок, измеряемых непосредственно, справедливы описанные выше допущения, необходимо действовать [15, 16] следующим образом:

1. Оценка математического ожидания $\hat{y}$ функции от нескольких измеренных переменных вычисляется по формуле

$$\hat{y} = \overline{y} = f(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_k})\,. \tag{25}$$

2. Оценка среднего квадратического отклонения вычисляется способом, основанным на дифференциальном исчислении. Если среднеквадратичные ошибки $s_{x_1}, s_{x_2}, \ldots, s_{x_k}$ величин $x_1, x_2, \ldots, x_k$, определенные по данным измерений малы по сравнению с измеряемыми величинами, то

$$s_y = \sqrt{(\frac{\partial f}{\partial x_1} s_{x_1})^2 + (\frac{\partial f}{\partial x_2} s_{x_2})^2 + \ldots + (\frac{\partial f}{\partial x_k} s_{x_k})^2}, \tag{26}$$

где $\partial f/\partial x_i$ — частная производная функции f, вычисленная при значении переменных, соответствующих средним значениям
$\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k$.

3. Для определения доверительного интервала используется та же процедура, что и при определнии доверительного интервала для непосредственно измеряемой величины (см. параграф 5.4).

Сравнивая слагаемые в подкоренном выражении, легко определить ошибки какой из измеряемых величин вносят основной вклад в суммарную ошибку. Предварительный анализ выражения для s_y полезен для того, чтобы при планировании

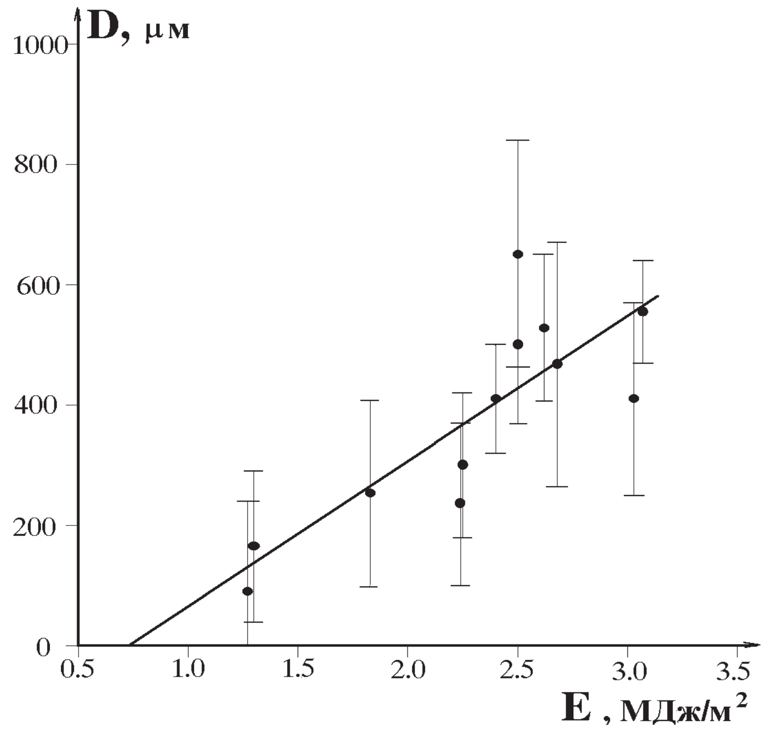


Рис. 9. Пример построения линейной зависимости методом наименьших квадратов

постановки эксперимента попытаться добиться повышенной точности измерения величины, вносящей основной вклад в ошибку, а также понять, какие измерения следует производить тщательно, а на какие не следует тратить больших усилий.

6 МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

6.1 Основные формулы метода наименьших квадратов

В предыдущих разделах мы рассмотрели вопрос о статистической обработке результатов измерения некоторой величины, которую мы обозначим здесь буквой y, при фиксированных условиях эксперимента (например, измерение сопротивления отрезка провода при фиксированных температуре, давлении). Точность полученного значения $\bar{y}$ характеризуется величиной стандартной ошибки, определяемой в результате многократных измерений.

Изменение условий эксперимента может приводить к изменению y. Например, зависимость сопротивления от температуры или показанная на рис. 9 зависимость средней глубины испарения графитовой пластинки от средней плотности энергии в пучке вблизи порога испарения (см. *Приложение* **??**) описывается линейной функцией вида

$$y = a + bx, \tag{27}$$

где a и b — постоянные, которые необходимо определить[6]. И хотя на рисунке величины, зависимость между которыми мы хотим определить, называются D и E соответственно, мы для простоты в этом параграфе будем говорить о зависимости $y(x)$. Пусть при нескольких значениях x_i измерены средние значения величины y_i и ее погрешность (эти погрешности показаны на рисунке вертикальными линиями с горизонтальными ограничителями). Значения y_i могут быть представлены точками на графике (см. рис. 9). Теперь можно задать вопрос, как, используя эти данные, определить параметры a и b прямой, которая наилучшим образом проходила бы через экспериментальные точки.

Имеется стандартный метод решения такой задачи, называемый методом наименьших квадратов, который является частным примером применения метода максимального правдоподобия (см., например, [6]). Зависимость $y(x)$ может быть любой, но здесь мы рассмотрим частный случай линейной зависимости (27), считая, что ошибка измерений x_i очень мала ($s_{x_i} \approx 0$), а s_{y_i} одинаковы для всех точек. Разумно предположить, что наилучшим приближением к "истинной" кривой будет такая прямая, для которой сумма

$$S = \sum_{i=1}^{n} (y_i - bx_i - a)^2 \tag{28}$$

была бы минимальной. Здесь выражение в скобках — расстояние i-й экспериментальной точки от прямой с параметрами a и b.

Известно, что минимум (или максимум) функции можно найти приравняв нулю производную по варьируемому параметру, т.е. искомые значения параметров должны удовлетворять соотношениям:

$$\frac{\partial S}{\partial b} = \sum_{i=1}^{n} (-2x_i)(y_i - bx_i - a) = 0, \tag{29}$$

$$\frac{\partial S}{\partial a} = \sum_{i=1}^{n} (-2)(y_i - bx_i - a) = 0. \tag{30}$$

Просуммировав и обратив внимание, что $\sum x_i/n$ и $\sum y_i/n$ есть просто $\overline{x}$ и $\overline{y}$ - координаты "центра тяжести" экспериментальных точек, получим

$$b \sum {x_i}^2 + a \sum x_i = \sum x_i y_i, \tag{31}$$

[6]Следует обратить внимание на то, что когда говорят о линейной аппроксимации, то речь идет о линейности по коэффициентам a и b. Это может быть, например, зависимость вида $y = a + b\sin(x)$ или $y = a + b\sqrt(x)$ как в *Дополнении*.

$$b\overline{x} + a = \overline{y}. \tag{32}$$

Из (32) следует, что “наилучшая” кривая проходит через центр тяжести $(\overline{x}, \overline{y})$, т.е.

$$a = \overline{y} - b\overline{x}, \tag{33}$$

а значение b вычисляется по формуле:

$$b = \frac{\sum x_i y_i - \overline{y} \sum x_i}{\sum x_i^2 - \overline{x} \sum x_i} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2}. \tag{34}$$

Таким образом, мы получим аналитическое выражение прямой, наилучшим образом аппроксимирующей экспериментальные точки. Поскольку число точек конечно, прямая, полученная из другой серии измерений, будет идти несколько по другому. Следуя идеологии, изложенной нами ранее, можно, используя значения s_{y_i} для точек y_i, вычислить среднеквадратичную ошибку для параметров b и a и представить окончательный ответ в виде

$$y = (b \pm \Delta b)x + (a \pm \Delta a), \tag{35}$$

$$\text{где } (\Delta b)^2 = \frac{\sum (y_i - bx_i - a)^2}{(n-2)(\sum x_i^2 - n\overline{x}^2)},$$

$$(\Delta a)^2 = \left(\frac{\sum x_i^2}{n}\right) \cdot (\Delta b)^2.$$

Подробнее о способе вычисления Δb и Δa см.[3].

6.2 Метод наименьших квадратов и проблема “нуля”

При обработке данных по методу наименьших квадратов возникает много “подводных” камней, которые мы здесь не разбираем. Тем не менее, один вопрос стоит обсудить. Речь идет об аппроксимации экспериментальных данных прямой, проходящей через начало координат.

Зачастую из теории ясно, что аппроксимируемая зависимость должна проходить через начало координат, т.е. $a = 0$. Это, например, имеет место при выполненнии лабораторной работы по определению скорости звука в стержнях (зависимость времени прохождения звуковой волны вдоль стержня от его длины). Тем не менее, при обработке таких данных нельзя использовать зависимость вида $y = bx$, поскольку это эквивалентно добавлению несуществующих экспериментальных точек к реально измеренному набору данных. Более того, аппроксимация экспериментальных данных зависимостью вида (27) позволяет определить систематическую погрешность, которая присутствует в измерениях. Если величина a не равна нулю, и “нуль” не попадает внутрь диапазона погрешности $\pm\Delta a$, то следует сделать вывод, что либо принятая модель (т.е. что $y = bx$) неверна, либо точность проведенных экспериментов (например, неучтенная систематическая погрешность) больше статистической и невозможно с требуемой точностью определить величину b.

Для более аккуратного определения величины b необходимо, как правило, расширить диапазон изменения x_i в эксперименте, если же это не улучшает ситуацию, то необходимо проверить, нет ли других источников систематической ошибки (см., например, описание экспериментов в *Дополнении*).

7 УЧЕТ ВСЕХ ВИДОВ ПОГРЕШНОСТЕЙ

7.1 Взвешивание результатов

В ряде случаев необходимо объединить результаты измерений, выполненных в разное время, различными методами, на разных установках и, вообще говоря, с разной точностью. При этом встает вопрос, как находить правильное значение интересующей величины и соответствующей ей случайной ошибки. В таком случае приходится проводить, так называемое взвешивание результатов. В простейшем случае объединение (взвешивание) серий измерений, проведенных на одной установке с одинаковой точностью, можно понять на следующем примере[3]. Пусть 10 измерений величины x проведены утром и среднее их значений

$$\overline{x_1} = \frac{1}{10}(x_1 + x_2 + \cdots + x_{10}),$$

а 5 значений $x_{11}, \cdots x_{15}$ получено в вечернем эксперименте и их среднее

$$\overline{x_2} = \frac{1}{5}(x_{11} + x_{12} + \cdots + x_{15}).$$

Очевидно, наилучшим значением среднего для 15 измерений будет *взвешенное среднее*:

$$\overline{x}_* = \frac{1}{15}(x_1 + x_2 + \cdots + x_{15}),$$

что не совпадает с простым средним величин $\overline{x}_1$ и $\overline{x}_2$. Если использовать эти величины, то выражение для $\overline{x}_*$ будет иметь вид

$$\overline{x}_* = \frac{10\overline{x}_1 + 5\overline{x}_2}{15}.$$

Числа 10 и 5 называются *статистическими весами* величин $\overline{x}_1$ и $\overline{x}_2$. Следовательно, больший вклад в окончательный результат вносит в данном случае серия измерений с большей статистикой (большим количеством измерений).

В общем случае иногда приходится суммировать результаты, полученные с разной точностью. Пусть мы имеем N измерений величины x_i, имеющие стандартные отклонения s_i. Тогда статвес g_i каждого измерения можно определить как величину, обратную дисперсии (заметим, что, как будет видно из последующих выражений, существенны не абсолютные, а относительные значения статвесов):

$$g_i = \frac{1}{{\sigma_i}^2}. \tag{36}$$

Это выражение выглядит очень естественно: больший вклад во взвешенный результат дают измерения, выполненные с большей точностью.

В теории показывается, что средневзвешенное значение определяется соотношением:

$$\overline{x}_* = \frac{\sum g_i \overline{x}_i}{\sum g_i}, \tag{37}$$

а *оценка дисперсии* среднего взвешенного выражением:

$$s^2_{x_*} = \frac{\sum g_i (\overline{x}_i - \overline{x}_*)^2}{(N-1)\sum g_i} \ . \tag{38}$$

7.2 Учет систематической ошибки

До сих пор мы говорили об оценке случайных ошибок. Что касается систематической ошибки, то на нее надо внести поправку (если это возможно) и тем самым вовсе ее устранить, либо, по крайней мере, сделать пренебрежимо малой по сравнению со случайной. Если по каким-либо причинам оставшиеся систематические ошибки сравнимы со случайными, и они в одинаковой степени определяют точность результата, встает вопрос о сложении ошибок.

В настоящее время нет общепринятого мнения, как определить суммарную ошибку измерений в этом случае. Иногда рекомендуют обращаться с систематической ошибкой как со случайной, определяя квадрат общей ошибки как сумму квадратов систематической и случайной ошибок. При этом остается неясным, каким доверительным интервалом надлежит пользоваться и как определить соответствующую ему доверительную вероятность. На наш взгляд, представляется полезным в окончательном результате приводить отдельно величины случайной и возможной систематической ошибок. Это позволяет более корректно сравнивать результаты, полученные разными авторами на различных установках.

В ряде случаев случайные ошибки много меньше систематических. Например, если при считывании показаний прибора в серии измерений разброс измеренных значений много меньше погрешности самого прибора. В этом случае бессмысленно увеличивать число измерений для нахождения случайной ошибки, достаточно двух-трех измерений для исключения возможного промаха. Оценку систематической ошибки косвенных измерений можно при этом проводить по следующей формуле:

$$\Delta y = \sum_{i=1}^{k} \left|\frac{\partial f}{\partial x_i}\right| \cdot |\Delta x_i|, \tag{39}$$

где Δx_i — оцененные систематические ошибки измеряемых величин. Выражение (39) дает максимально возможную в данном случае оценку ошибки. Если вы приводите ошибку, вычисленную по этой формуле, то в пояснении к работе следует отметить, что приводится оцененная максимальная систематическая ошибка.

7.3 Ошибки и здравый смысл

В некоторых работах степень точности экспериментального результата полностью определяет ценность работы. Высокая точность часто требуется в так называемых метрологических работах, то есть в работах, целью которых является определение констант: величин мировых постоянных, плотностей веществ, длин волн в оптических спектрах атомов и т.п. В то же время, во многих других случаях (особенно в поисковых работах) столь высокой точности не требуется. Часто достаточно иметь результат с ошибкой 25%, а иногда и 100% от измеряемой величины.

Не тратьте слишком много усилий на повышение точности измерений и набор статистики, ели Вам этого не нужно!

Найденная величина ошибки показывает, в какой мере значим конечный результат измерений. Ошибку следует вычислять до второй значащей цифры, а при записи результата округлять в большую сторону, оставляя одну значащую цифру. В записи результата необходимо также сделать округление так, чтобы порядок последней цифры результата соответствовал порядку приведенной ошибки. Например, если вычисленное среднее значение оказалось равным 45.0876, а стандартная ошибка среднего (с учетом коэффициента Стьюдента) - 0.00376, то следует округлить ошибку до 0.004, а результат записать в виде 45.088 ± 0.004.

8 ЗАКЛЮЧЕНИЕ

Приведенное здесь элементарное изложение методов обработки экспериментальных данных далеко от полноты и предназначено для самого первоначального знакомства с теорией ошибок. Тем не менее мы надеемся, что с его помощью вы сможете обработать данные, полученные в лаборатории, правильно представить результаты и оценить их точность. Возможно полученный опыт позволит вам более осознанно воспринять позднее теорию математической статистики.

Формулы "для лентяев"

Функция распределения (вероятность попадания случайной величины в интервал от $-\infty$ до x) или интегральная функция распределения	$F(x) = \int\limits_{-\infty}^{x} f(z)dz$
Вероятность $\mathcal{P}$ того, что результат отдельного измерения будет лежать в интервале x_1, x_2	$\mathcal{P}(x_1 < x < x_2) \equiv \int\limits_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1)$
Математическое ожидание случайной величины	$M(x) = \overline{x} = \int\limits_{-\infty}^{+\infty} xf(x)dx$
Математическое ожидание функции $y(x)$ от непрерывной случайной величины x	$M(y(x)) = \overline{y(x)} = \int\limits_{-\infty}^{+\infty} y(x)f(x)dx$
Второй момент или дисперсия	$\sigma^2 = D(x) = M((x-\overline{x})^2) = \int\limits_{-\infty}^{+\infty} (x-\overline{x})^2 f(x)dx$
Среднее квадратическое отклонение	$\sigma = \sqrt{\sigma^2}$
Плотность вероятности *равномерного (прямоугольного)* распределение	$f(x) = \begin{cases} 0, & \text{если } x < x_1 \\ \frac{1}{x_2 - x_1}, & \text{если } x_1 \le x \le x_2 \\ 0, & \text{если } x > x_2 \end{cases}$
Распределение *Пуассона*	$P_n(\mu) = \frac{\mu^n}{n!} e^{-\mu}$
Распределение *Гаусса*	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}}$
Среднее значение выборки (оценка математического ожидания) случайной величины	$\hat{x} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Оценка математического ожидания $\hat{y}$ функции от нескольких случайных величин (измеренных переменных)	$\hat{y} = \overline{y} = f(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_k})$
Наилучшая оценка *среднего квадратического отклонения* случайной величины	$\hat{\sigma} = s_n = \sqrt{\frac{1}{(n-1)} \sum (x_i - \overline{x})^2}$
Оценка *среднего квадратического отклонения* функции нескольких переменных	$s_y = \sqrt{(\frac{\partial f}{\partial x_1} s_{x_1})^2 + (\frac{\partial f}{\partial x_2} s_{x_2})^2 + \ldots + (\frac{\partial f}{\partial x_k} s_{x_k})^2}$
Стандартная ошибка или *среднеквадратичная ошибка среднего*	$s_x = \frac{s_n}{\sqrt{n}}$
Доверительная оценка при выбранном уровне статистической достоверности $\mathcal{P}$	$\lvert\overline{x} - X\rvert < t(\mathcal{P}, \nu)\frac{s_n}{\sqrt{n}}$, где $\nu = n - 1$, а $t(\mathcal{P}, \nu)$ – функция Стьюдента

Аппроксимация экспериментальных данных прямой $y = a + bx$

$$y = (b \pm \Delta b)x + (a \pm \Delta a),$$

$$\text{где } a = \overline{y} - b\overline{x}, \quad b = \frac{\sum x_i y_i - \overline{y}\sum x_i}{\sum x_i^2 - \overline{x}\sum x_i} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\overline{x^2} - \overline{x}^2}.$$

$$(\Delta a)^2 = \left(\frac{\sum x_i^2}{n}\right)\cdot(\Delta b)^2, \quad (\Delta b)^2 = \frac{\sum (y_i - bx_i - c)^2}{(n-2)(\sum x_i^2 - n\overline{x}^2)}.$$

ПРИЛОЖЕНИЯ

А СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

Итак, мы знаем, что разброс экспериментальных данных может подчиняться различным статистическим распределениям. При этом вне зависимости от вида распределения среднее значение и стандартное отклонение определяются с помощью выражений (18) и (19), а среднеквадратичная ошибка среднего (стандартная ошибка)— выражением (21). Таким образом, прочитав предыдущие главы, вы получили основы знаний, достаточных в большинстве случаев для первичной обработки данных. В некоторых случаях, однако, появляется необходимость более глубокой статистической обработки данных и прежде всего необходимо ответить на вопрос:

Какой функции распределения отвечает распределение экспериментальных точек?

Какова достоверность предположения (гипотезы) о том, что экспериментальные данные действительно соответствуют предполагаемой нами функции распределения?

Пусть в результате эксперимента с общим числом измерений N мы получили набор данных, который представлен в виде гистограммы на рис. 10. По вертикальной оси на рисунке отложено число измерений n_i^* $(i = 1, \ldots, k)$, результат которых x попал в соответствующий интервал ("бин") Δx_i на оси абсцисс. Эти интервалы не обязательно одинаковы по величине.

Предположим, что у нас есть основания полагать, что плотность вероятности предполагаемого распределения описывается функцией вида $f(x)$, показанной на рисунке сплошной линией (гипотеза!). Для проверки данной гипотезы следует количественно сравнить ожидаемое и реально наблюдаемое в каждом бине число событий.

Прежде всего отметим, что при достаточно большом числе бинов вероятность попадания результата в каждый бин $p_i \ll 1$, а $n_i \ll N$. Мы знаем, что в этом случае распределение числа событий (число экспериментов, давших результат, попавший в данный бин) подчиняется статистике Пуассона. Если среднее число событий в бине (в нашем случае это бин k) очень мало (~ 1), то функция распределение Пуассона для данного бина $P_k(n_k, \overline{n_k})$ (11) имеет несимметричный вид, показанный на рисунке слева. Для достаточно больших n (практически при $n_i \geq 5$) распределение начинает по форме приближаться к гауссову с $\sigma_i \sim \sqrt{n_i}$. Тогда

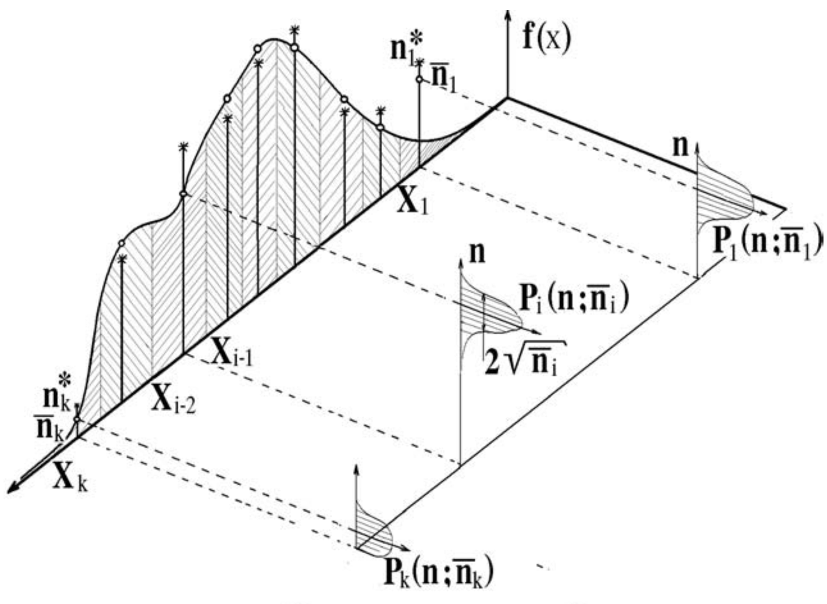


Рис. 10. Проверка гипотез и распределение χ^2

можно ожидать, что разность между ожидаемым $\overline{n_i}$ и наблюдаемым n_i^* числом событий будет для каждого бина в среднем порядка стандартного отклонения:

$$\overline{n_i} - n_i^* \sim \sqrt{n_i^*}.$$

Для того чтобы распределение могло аппроксимироваться гауссовым, ширины всех бинов, включая крайние, следует выбирать такими, чтобы в каждом было не менее пяти событий (сравните распределения в первом и последнем бинах).

Из приведенных рассуждений понятно, что критерием справедливости нашей гипотезы о виде распределения может служить неравенство, называемое *критерием* χ^2:

$$\chi^2 = \sum_{i=1}^{k} \frac{(\overline{n_i} - N_i^*)^2}{\overline{n}_i} \leq k, \tag{40}$$

где k полное число бинов[7].

Если мы проделали описанную процедуру и получили $\chi^2 > k$, то можно усомниться в том, что результаты описываются распределением $f(x)$.

Ясно, что описанный критерий достаточно грубый. При более детальном исследовании мы обнаружим, что имеется распределение вероятности появления того или иного значения величины χ^2, называемое *распределением* χ^2. Следовательно, и здесь мы можем ввести понятие доверительной вероятности и доверительного

[7]Вообще говоря, следует говорить не о числе бинов, а о числе степеней свободы d, которое равно числу бинов минус число дополнительных условий.

интервала для χ^2. Для более детального знакомства с этим вопросом мы рекомендуем книги [4], [7].

B ИНТЕГРАЛ ОШИБОК

Значения $\mathcal{P}(|x| < x^*) = 2\Phi(t)$ для t, выраженного в долях среднеквадратичной ошибки $t = \Delta x/\sigma$

Таблица 3. Интеграл ошибок (интеграл вероятностей)

t	P	t	P	t	P
0	0	1.2	0.77	2.6	0.990
0.05	0.04	1.3	0.80	2.7	0.993
0.1	0.08	1.4	0.84	2.8	0.995
0.15	0.12	1.5	0.87	2.9	0.996
0.2	0.16	1.6	0.89	3.0	0.997
0.3	0.24	1.7	0.91	3.1	0.9981
0.4	0.31	1.8	0.93	3.2	0.9986
0.5	0.38	1.9	0.94	3.3	0.9990
0.6	0.45	2.0	0.95	3.4	0.9993
0.7	0.51	2.1	0.964	3.5	0.9995
0.8	0.57	2.2	0.972	3.6	0.9997
0.9	0.63	2.3	0.978	3.7	0.9998
1.0	0.68	2.4	0.984	3.8	0.99986
1.1	0.73	2.5	0.988	3.9	0.99990

C КОЭФФИЦИЕНТ СТЬЮДЕНТА

Таблица 4. Величины t для различных значений доверительного уровня $\mathcal{P}$

n-1	P=68.3%	P= 95%	P=99%	P=99.73%
(1)	(1.8)	(12.7)	(64)	(235)
2	1.32	4.70	9.9	19.2
3	1.20	3.18	5.8	9.2
4	1.15	2.78	4.6	6.6
5	1.11	2.57	4.0	5.5
6	1.09	2.45	3.7	4.9
7	1.08	2.37	3.5	4.5
8	1.07	2.31	3.4	4.3
9	1.06	2.26	3.2	4.1
10	1.05	2.23	3.2	4.0
15	1.03	2.13	3.0	3.6
20	1.03	2.09	2.8	3.4
30	1.02	2.04	2.8	3.3
50	1.01	2.01	2.7	3.2
100	1.00	1.98	2.6	3.1
200	1.00	1.97	2.6	3.0
предел	1.00	1.96	2.58	3.0

Список литературы

1. Росляков Г.В., Князев Б.А. *Методы обработки экспериментальных данных*. Новосибирск: НГУ, 1985. 5
2. Князев Б.А., Кругляков Э.П., Воробьев В.В., Капитонов В.А. *Постоянная Керра воды* ЖПМТФ. 1976. (1):157–160. 7, 9
3. Сквайрс Дж. *Практическая физика*. М.: Мир, 1971. 31, 32
4. Кунце Х.-И. *Методы физических измерений*. М.: Мир, 1989. 39
5. Зайдель А.Н. *Ошибки измерений физических величин*. Л.: Наука, 1974.
6. Худсон Д. *Статистика для физиков*. М.: Мир, 1967. 30
7. Тейлор Дж. *Введение в теорию ошибок*. М.: Мир, 1985. 39
8. Румшисский Л.З. *Математическая обработка результатов эксперимента*. М.: Наука, 1971.
9. Вильямс А., Кэмпион П.Дж., Барнс Д.Е. *Практическое руководство по представлению результатов измерений*. М.: Атомиздат, 1979.
10. Агекян Т.А. *Основы теории ошибок для астрономов и физиков*. М.: Наука, 1968.
11. Маркин Н.С. *Основы теории обработки результатов измерений*. М.: Изд. стандартов, 1991.
12. Секей Г. *Парадоксы в теории вероятностей и в математической статистике*. М.: Мир, 1990.
13. Мантуров О.В. *Курс высшей математики*. М.: Выс.школа, 1991. 7, 22
14. Корн Т., Корн Г. *Справочник по математике для научных работников и инженеров*. М.: Наука, 1968. 16
15. Сирая Т.Н., Грановский В.А. *Методы обработки экспериментальных данных при измерениях*. Л.: Энергоатомиздат, 1990. 28
16. Дойников А.С., Брянский Л.Н. *Краткий справочник метролога*. М.: Стандарты, 1991. 28

Б.А.Князев, В.С.Черкасский
НАЧАЛА ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ
ДАННЫХ

Подписано в печать Формат $60 \times 84/16$
Офсетная печать. Уч.-изд. л. 5.5
Заказ ь Тираж 500 экз. Цена 500р.

Редакционно-издательский отдел Новосибирского университета;
участок оперативной полиграфии НГУ; 630090, Новосибирск 90,
ул. Пирогова, 2.