

СТАТИСТИЧЕСКАЯ ОБРАБОТКА ЭКСПЕРИМЕНТА

В главе XV рассмотрены основные вопросы статистической обработки результатов эксперимента: определение наиболее достоверного значения измеряемой величины и погрешности этого значения по нескольким измерениям, оценка достоверности различия двух близких величин, установление достоверной функциональной зависимости между двумя величинами и аппроксимация этой зависимости.

Глава носит вспомогательный характер. Материал в ней изложен в справочной форме, без доказательств. Обоснование и более подробное изложение приведенных методов имеется, например, в [7, 26, 43].

1. Ошибки эксперимента. Численные методы часто применяют при математическом моделировании физических и других процессов. Результаты расчетов в этом случае сравнивают с экспериментальными данными и по степени их согласованности судят о качестве выбранной математической модели. Чтобы обоснованно сделать заключение о соответствии или несоответствии, вычислитель должен знать, что такое погрешность эксперимента и как с ней обращаются, а также уметь в случае необходимости провести статистическую обработку первичных данных эксперимента.

Кроме того, задача статистической обработки эксперимента представляет самостоятельный интерес, поскольку она очень важна в тех приложениях, когда или требуется особенно высокая точность (например, уравнивание триангуляционных сетей в геодезии), или разброс отдельных измерений превосходит исследуемый эффект (что нередко встречается в физике элементарных частиц, химии сложных соединений, испытании сельскохозяйственных сортов, медицине и т. д.).

Обычно, чем точнее эксперимент, тем более сложной аппаратуры он требует и дороже обходится. Однако хорошо продуманная математическая обработка результатов в ряде случаев позволяет выявить и частично исключить ошибки измерений; это может оказаться не менее эффективным, чем использование более дорогой и точной аппаратуры. В этой главе будет рассмотрена статистическая обработка, позволяющая существенно уменьшить и аккуратно оценить случайную ошибку измерений.

Ошибки эксперимента условно разбивают на систематические, случайные и грубые; рассмотрим их подробнее.

Систематические ошибки — это те, которые не меняются при многократном повторении данного эксперимента. Примерами таких ошибок являются пренебрежение выталкивающим действием воздуха при точном взвешивании или измерение тока гальванометром, нуль которого неправильно установлен. Различают три вида систематических ошибок.

а) Ошибки известной природы, величину которых можно определить; их называют *поправками*. Так, при точном взвешивании рассчитывают поправку на выталкивающее действие воздуха и прибавляют ее к измеренной величине. Внесение поправок позволяет существенно уменьшить (или даже практически исключить) ошибки такого рода.

Заметим, что иногда расчет поправок бывает самостоятельной сложной математической задачей. Например, некорректно поставленная задача (14.2) о восстановлении переданного радиосигнала по принятому является, по существу, нахождением поправки на искажение принимающей аппаратуры.

б) Ошибки известного происхождения, но неизвестной величины. К ним относится погрешность измерительных приборов, определяемая их классом точности. Для таких ошибок обычно известна только верхняя граница, а как поправки их учесть нельзя.

в) Ошибки, о существовании которых мы не знаем; например, используется прибор со скрытым дефектом или изношенный, фактическая точность которого существенно хуже, чем обозначено в техническом паспорте.

Для выявления систематических ошибок всех видов обычно заранее отлаживают аппаратуру на эталонных объектах с хорошо известными свойствами.

Случайные ошибки вызываются большим числом факторов, которые при повторении одного и того же эксперимента могут действовать по-разному, причем учесть их влияние практически невозможно. Например, при измерении длины предмета линейка может быть неточно приложена, взгляд наблюдателя может падать не перпендикулярно шкале и т. д.

При многократном повторении эксперимента результат вследствие случайной ошибки будет различным. Однако такое повторение и соответствующая статистическая обработка позволяют, во-первых, определить величину случайной ошибки и, во-вторых, уменьшить ее. Повторяя измерение достаточное число раз, можно уменьшить случайную ошибку до требуемой величины (целесообразно уменьшать ее до величины 50—100% от систематической ошибки).

Грубые ошибки — это результат невнимательности наблюдателя, который может записать одну цифру вместо другой. При

единичном измерении грубую ошибку не всегда можно опознать. Но если измерение повторено несколько раз, то при статистической обработке выясняют вероятные пределы случайной ошибки. Измерение, существенно выходящее за полученные пределы, считается грубо ошибочным и не учитывается при окончательной обработке результатов.

Таким образом, если измерение повторено достаточно много раз, то можно практически исключить грубые и случайные ошибки, так что точность ответа будет определяться только систематической ошибкой. Однако во многих приложениях это требуемое число раз оказывается неприемлемо большим, а при реально осуществимом числе повторений случайная ошибка может быть определяющей.

2. Величина и доверительный интервал. Пусть измерение проводят несколько раз, причем условия эксперимента поддерживают, насколько возможно, неизменными. Поскольку строго соблюдать неизменность условий невозможно, результаты отдельных измерений x_i будут несколько различаться. Их можно рассматривать как значения случайной величины ξ , распределенной по некоторому закону, заранее нам неизвестному.

Очевидно, *математическое ожидание $M\xi$ равно точному значению измеряемой величины x* (строго говоря, точному значению плюс систематическая ошибка).

Обработка измерений основана на центральной предельной теореме теории вероятностей: *если ξ есть случайная величина, распределенная по любому закону, то*

$$\eta_n = \frac{1}{n} \sum_{i=1}^n \xi_i$$

есть также случайная величина, причем

$$M\eta_n = M\xi, \quad D\eta_n = \frac{1}{n} D\xi, \quad (1)$$

а закон распределения величины η_n стремится к нормальному (гауссову) при $n \rightarrow \infty$. Поэтому среднееарифметическое нескольких независимых измерений

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \approx M\xi \quad (2)$$

является приближенным значением измеряемой величины, причем с тем большей надежностью, чем больше число измерений n .

Однако равенство $\bar{x} \approx M\xi$ не является точным, и нельзя даже строго указать предел его ошибки; в принципе \bar{x} может сколь угодно сильно отличаться от $M\xi$, хотя вероятность такого собы-

тия ничтожно мала. Ошибка приближенного равенства (2) носит вероятностный характер и описывается *доверительным интервалом* β , т. е. границей, которую с *доверительной вероятностью* p_0 не превышает разность $|\bar{x} - M\xi|$. Символически это записывают следующим образом:

$$p \{ |\bar{x} - M\xi| \leq \beta \} = p_0. \quad (3)$$

Доверительный интервал зависит от закона распределения ξ (а тем самым — от постановки эксперимента), от числа измерений n , а также от выбранной доверительной вероятности p_0 . Из (3) видно, что чем ближе p_0 к единице, тем шире оказывается доверительный интервал.

Доверительную вероятность p_0 выбирают, исходя из практических соображений, связанных с применениями полученных результатов. Например, если мы делаем игрушечный воздушный змей, то вероятность благополучного полета $p_0 = 0,8$ нас устроит, а если конструируем самолет, то даже вероятность $p_0 = 0,999$ недостаточна. Во многих физических измерениях $p_0 = 0,95 \div 0,99$ считается достаточной.

Замечание 1. Пусть требуется найти величину z , но измерять удобнее величину x , связанную с ней известным соотношением $z = f(x)$; например, нас интересует джоулево тепло, а измерять легче ток. При этом следует помнить, что

$$M\xi = \int f(x) \rho(x) dx \neq f(M\xi) = f\left(\int x \rho(x) dx\right);$$

так, среднее значение переменного тока равно нулю, а средний джоулев нагрев отличен от нуля. Поэтому, если мы вычислим сначала \bar{x} , а затем положим $\bar{z} = f(\bar{x})$, это будет грубая ошибка. Следует по каждому измерению x_i вычислять $z_i = f(x_i)$ и далее обрабатывать полученные значения z_i .

Ширина доверительного интервала. Если известна плотность распределения $\rho_n(y)$ величины η_n , то доверительный интервал можно определить из (3), разрешая уравнение

$$p_0 = \int_{M\eta - \beta}^{M\eta + \beta} \rho_n(y) dy, \quad M\eta_n = M\xi, \quad (4)$$

относительно β . Выше отмечалось, что при $n \rightarrow \infty$ распределение η_n стремится к нормальному *):

$$\rho_n(y) \approx \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left[-\frac{(y - M\eta)^2}{2\sigma_n^2}\right], \quad \sigma_n = \sqrt{D\eta_n}; \quad (5)$$

*) В самом худшем случае, когда ξ есть равномерно распределенная случайная величина, распределение η_n близко к нормальному при $n \sim 30$, а интеграл в (3) близок к интегралу от нормального распределения при существенно меньших n .

здесь $D\eta_n$ — дисперсия распределения, а величину σ_n называют *стандартным отклонением* или просто *стандартом* *).

Подставляя (5) в (4) и полагая $\beta = t\sigma_n$, т. е. измеряя доверительный интервал в долях стандарта, получим соотношение

$$p_0 = \sqrt{\frac{2}{\pi}} \int_0^t e^{-\tau^2/2} d\tau. \quad (6)$$

Интеграл ошибок, стоящий в правой части (6), табулирован, так что из этого соотношения можно определить доверительный интервал $t(p_0)$. Зависимость $t(p_0)$ дается в таблице 23 строкой, соответствующей $n = \infty$.

Из таблицы 23 видно, что доверительный интервал $\beta = 3\sigma_n$ соответствует доверительной вероятности $p_0 = 0,997$, так что отклонение \bar{x} от $M\xi$ более чем на $3\sigma_n$ маловероятно. Но отклонение более чем на σ_n довольно вероятно, поскольку ширине $\beta = \sigma_n$ соответствует $p_0 = 0,7$.

Таким образом, если известна дисперсия $D\xi$, то нетрудно определить стандарт $\sigma_n = \sqrt{D\xi/n}$ и, тем самым, абсолютную ширину доверительного интервала β . В этом случае даже при выполнении одного измерения можно оценить случайную ошибку **, а увеличение числа измерений позволяет уменьшать доверительный интервал, поскольку $\sigma_n \sim n^{-1/2}$.

Критерий Стьюдента. Чаще всего дисперсия $D\xi$ неизвестна, поэтому выполнить оценку ошибки указанным выше способом обычно не удастся. При этом точность однократного измерения неизвестна. Однако, если измерение повторено несколько раз, можно приближенно найти дисперсию:

$$D\xi \approx \frac{1}{n} \sum_{i=1}^n (x_i - M\xi)^2 \approx \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (7)$$

Точность этого выражения невелика по двум причинам: во-первых, число членов суммы обычно мало; во-вторых, использование замены $M\xi \approx \bar{x}$ вносит ошибку $O(1/n)$, значительную при малых n . Более хорошее приближение дает так называемая *несмещенная оценка дисперсии*:

$$D\xi \approx s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (8)$$

*) Для произвольного закона распределения $\sqrt{D\eta}$ называют среднеквадратичным отклонением.

***) Однако при $n=1$ считать распределение $\rho_1(y) = \rho(x)$ нормальным и пользоваться формулой (6), вообще говоря, нельзя. Этот вопрос будет рассмотрен ниже,

где величину s называют *стандартом выборки*. Далее будем пользоваться только оценкой (8).

Оценка (8) также является приближенной, поэтому нельзя пользоваться формулой (6), заменяя в ней σ_n на s/\sqrt{n} . Надо вносить в нее поправку, тем большую, чем меньше n . Если распределение η_n считать нормальным при любых n , то связь доверительного интервала со стандартом выборки устанавливается критерием Стьюдента:

$$\beta = t(p_0, n) s_n, \quad s_n = \frac{s}{\sqrt{n}}, \quad (9)$$

где коэффициенты *) Стьюдента $t(p_0, n)$ представлены в таблице 23.

Таблица 23

Коэффициенты Стьюдента $t(p_0, n)$

p_0 $n-1$	0,5	0,7	0,8	0,9	0,95	0,98	0,99	0,995	0,998	0,999
1	1,0	2,0	3,1	6,3	13	32	64	127	318	637
2	0,8	1,3	1,9	2,9	4,3	7,0	9,9	14	22	32
3	0,8	1,3	1,6	2,4	3,2	4,5	5,8	7,5	10	13
4	0,7	1,2	1,5	2,1	2,8	3,7	4,6	5,6	7,2	8,6
5	0,7	1,2	1,5	2,0	2,6	3,4	4,0	4,8	5,9	6,9
6	0,7	1,1	1,4	1,9	2,4	3,1	3,7	4,3	5,2	6,0
7	0,7	1,1	1,4	1,9	2,4	3,0	3,5	4,0	4,8	5,4
8	0,7	1,1	1,4	1,9	2,3	2,9	3,4	3,8	4,5	5,0
9	0,7	1,1	1,4	1,8	2,3	2,8	3,3	3,7	4,3	4,8
10	0,7	1,1	1,4	1,8	2,2	2,8	3,2	3,6	4,1	4,6
15	0,7	1,1	1,3	1,8	2,1	2,6	2,9	3,3	3,7	4,1
20	0,7	1,1	1,3	1,7	2,1	2,5	2,8	3,2	3,5	3,8
30	0,7	1,1	1,3	1,7	2,0	2,5	2,8	3,0	3,4	3,7
60	0,7	1,0	1,3	1,7	2,0	2,4	2,7	2,9	3,2	3,5
∞	0,7	1,0	1,3	1,6	2,0	2,3	2,6	2,8	3,1	3,3
Критерий Чебышева	1,4	1,8	2,2	3,2	4,5	7,1	10	14	22	32

Очевидно, при больших n с хорошей точностью выполняется $\sigma_n \approx s_n$. Поэтому при $n \rightarrow \infty$ критерий Стьюдента переходит в формулу (6); выше отмечалось, что этой формуле соответствует строка $n = \infty$ таблицы 23. Однако при малых n доверительный интервал (8) оказывается много шире, чем по критерию (6).

Пример 1. Выбрано $p_0 = 0,99$ и выполнено 3 измерения; по таблице 23 доверительный интервал равен $\beta = 9,9s/\sqrt{3} = 5,7s$.

*) Их называют также *квантилями Стьюдента*.

К сожалению, не все физики и инженеры знакомы с понятием доверительного интервала и критерием Стьюдента. Нередко встречаются экспериментальные работы, в которых при малом числе измерений пользуются критерием (6) или даже считают, что значение s_n является погрешностью величины \bar{x} , и вдобавок оценивают дисперсию по формуле (7).

Для приведенного выше примера при первой ошибке был бы дан ответ $\beta = 1,5s$, при второй — $\beta = 0,6s$, а при третьей — $\beta = 0,7s$, что сильно отличается от правильного значения.

Замечание 2. Зачастую одна и та же величина x измерена в разных лабораториях на разном оборудовании. Тогда следует найти среднее и стандарт по формулам (2) и (8), где суммирование проводится по всем измерениям во всех лабораториях, и определить доверительный интервал по критерию Стьюдента.

Нередко при этом суммарный стандарт s оказывается больше, чем стандарты s_j , определенные по данным отдельных лабораторий. Это естественно. Каждая лаборатория делает при измерениях систематические ошибки, и часть систематических ошибок в разных лабораториях совпадает, а часть — различается. При совместной обработке различающиеся систематические ошибки переходят в разряд случайных, увеличивая стандарт.

Значит, при совместной обработке разнотипных измерений обычно систематическая ошибка значения \bar{x} будет меньше, а случайная — больше. Но случайную ошибку можно сколь угодно уменьшить, увеличивая число измерений. Поэтому такой способ позволяет получить окончательный результат с большей точностью.

Замечание 3. Если в разных лабораториях используется оборудование разного класса точности, то при такой совместной обработке надо суммировать с весами ρ_i :

$$\bar{x} = \frac{1}{R} \sum_{i=1}^n \rho_i x_i, \quad s^2 = \frac{n}{(n-1)R} \sum_{i=1}^n \rho_i (x_i - \bar{x})^2, \quad R = \sum_{i=1}^n \rho_i, \quad (10)$$

где ρ_i относятся, как квадраты точности приборов.

Произвольное распределение. Чаше всего число измерений n невелико и заранее неясно, можно ли считать распределение η_n нормальным и пользоваться приведенными выше критериями.

Для произвольного распределения $\rho(x)$ справедливо неравенство Чебышева

$$p \{ |x - M\xi| \geq \beta \} \leq \frac{D\xi}{\beta^2}.$$

Отсюда можно оценить доверительный интервал:

$$\beta \leq \frac{1}{\sqrt{1-\rho_0}} \sigma_n, \quad \sigma_n = \sqrt{\frac{D\xi}{n}}. \quad (11)$$

Коэффициент $(1 - p_0)^{-1/2}$ в этой оценке приведен в дополнительной строке таблицы 23.

Из таблицы видно, что если в качестве доверительной вероятности принять $p_0 = 0,95$, то для произвольного закона распределения с известной дисперсией доверительный интервал не превышает $5\sigma_n$. Для симметричного одновершинного распределения аналогичные оценки показывают, что доверительный интервал не превышает $3\sigma_n$; напомним, что для нормального распределения он равен $2\sigma_n$ (при выбранном $p_0 = 0,95$).

Разумеется, если вместо σ_n используют найденное по тем же измерениям значение s_n , то надо строить критерий, аналогичный критерию Стьюдента. Оценки при этом будут существенно хуже приведенных.

Проверка нормальности распределения. Из сравнения критериев (6) и (11) видно, что даже при невысокой доверительной вероятности $p_0 \leq 0,9$ оценки доверительного интервала при произвольном распределении вдвое хуже, чем при нормальном. Чем ближе p_0 к единице, тем хуже соотношение этих оценок. Поэтому целесообразно проверять, существенно ли отличается распределение $\rho(x)$ от нормального.

Распространенный способ проверки — исследование так называемых *центральных моментов* распределения:

$$m_k = \int_{-\infty}^{+\infty} (x - M\xi)^k \rho(x) dx, \quad k = 1, 2, \dots \quad (12)$$

Два первых момента, по определению, равны $m_1 = 0$, $m_2 = D\xi = \sigma^2$. Для нормального распределения два следующих момента равны $m_3 = 0$, $m_4 = 3\sigma^4$. Обычно ограничиваются этими моментами. Вычисляют их фактические значения по проведенным измерениям и проверяют, согласуются ли они со значениями, соответствующими нормальному распределению.

Удобно вычислять не сами моменты, а составленные из них безразмерные комбинации — *асимметрию* $A = \sigma^{-3}m_3$ и *эксцесс* $E = \sigma^{-4}m_4 - 3$; для нормального распределения они обращаются в нуль. Аналогично дисперсии, вычислим их по несмещенным оценкам:

$$A \approx \frac{1}{s^3(n-1)} \sum_{i=1}^n (x_i - \bar{x})^3, \quad E = \frac{1}{s^4(n-1)} \sum_{i=1}^n (x_i - \bar{x})^4 - 3, \quad (13)$$

где s определяется формулой (8). Собственные дисперсии этих величин известны и зависят только от числа измерений:

$$D(A) = \frac{6(n-1)}{(n+1)(n+3)}, \quad D(E) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}, \quad (14)$$

причем собственное распределение A является симметричным. Поэтому, если выполняются соотношения

$$|A| \leq 3\sqrt{D(A)}, \quad |E| \leq 5\sqrt{D(E)}, \quad (15)$$

то по критерию Чебышева (11) отличие A и E от нуля недостоверно, так что можно принять гипотезу о нормальности распределения $\rho(x)$.

Формулы (13) — (15) непосредственно относятся к распределению единичного измерения. На самом деле надо проверить, нормально ли распределение среднеарифметического η_n при выбранном n . Для этого делают большое число измерений $N = rn$, разбивают их на r групп по n измерений в каждой и среднее значение в каждой группе \bar{x} рассматривают как единичное измерение. Тогда проверка выполняется по формулам (13) — (15), где вместо n надо подставить r .

Разумеется, такую тщательную проверку проводят не в каждой измеряемой точке, а лишь во время отработки методики эксперимента.

З а м е ч а н и е 4. Аналогично проверяют любые естественнонаучные гипотезы. Производят большое число экспериментов и выясняют, нет ли среди них событий, маловероятных с точки зрения этой гипотезы. Если найдутся такие события, то гипотезу отвергают, если нет — условно принимают.

Выбор n . За счет увеличения числа измерений n можно неограниченно уменьшать доверительный интервал. Однако систематическая ошибка β_0 при этом не уменьшается, так что суммарная ошибка все равно будет больше β_0 . Поэтому целесообразно выбрать n так, чтобы ширина доверительного интервала составляла 50—100% β_0 . Дальнейшее увеличение числа измерений бессмысленно.

Чтобы найти удовлетворяющее этому требованию n , надо отдельные точки измерить достаточное число раз, вычислить стандарт s , убедиться в нормальности распределения η_n и на основании критерия Стьюдента (9) подобрать такое n , чтобы выполнялось неравенство

$$t(p_0, n) \leq \frac{\sqrt{n}}{s} \beta_0, \quad (16)$$

где коэффициенты Стьюдента $t(p_0, n)$ даются таблицей 23.

Из таблицы 23 видно, что при $n = 2$ доверительный интервал чересчур велик, так что следует производить не менее 3—4 измерений. При дальнейшем увеличении n коэффициенты Стьюдента убывают слабо и доверительный интервал s_n сужается почти пропорционально $n^{-1/2}$, т. е. довольно медленно. Поэтому обычно считают нецелесообразным брать $n > 5 - 10$, так как возрастающая трудоемкость эксперимента не оправдывается достигаемой точностью.

Пример 2. Отношение систематической ошибки к стандарту выборки оказалось $\beta_0/s = 0,8$, и принята доверительная вероятность $p_0 = 0,95$. Возьмем соответствующий столбец таблицы 23 и будем перебирать по очереди $n = 2, 3, \dots$, пока не получим $t(p_0, n)/\sqrt{n} \leq 0,8$; этому условию удовлетворяет $n = 9$.

Обнаружение грубых ошибок. Отличить грубую ошибку от случайной не всегда легко. Если число измерений мало, то широк доверительный интервал и даже значительные отклонения от среднего в него укладываются. Если же n велико, то возрастает вероятность того, что хотя бы одно измерение сильно отклонится от среднего случайно, т. е. на законном основании.

Пусть сделано n измерений и вычислены среднее \bar{x} и стандарт s . Чтобы с вероятностью p_1 ни одно из этих измерений не отличалось от $M\xi$ более чем на некоторое δ , каждое измерение должно оставаться в указанных пределах с вероятностью $\sqrt[n]{p_1}$, т. е. должно выполняться условие

$$p \{ |x_i - M\xi| \leq \delta \} = \sqrt[n]{p_1}. \quad (17)$$

Предполагая, что ξ имеет нормальное распределение, сравнивая (17) с критерием Стьюдента (9) и учитывая, что величина s вычислена по всей выборке, а применяется к отклонению единичного измерения, получим

$$\delta = t\left(\sqrt[n]{p_1}, n\right) s. \quad (18)$$

Вместо неизвестной величины $M\xi$ мы вынуждены подставлять в (17) величину \bar{x} , имеющую доверительный интервал $\beta = t(p_0, n) s/\sqrt{n}$. Сравним неравенства

$$|x_i - M\xi| \leq \delta, \quad |\bar{x} - M\xi| \leq \beta;$$

поскольку они носят вероятностный характер, то к ним надо применять не неравенство треугольника, а суммирование квадратов, что дает

$$|x_i - \bar{x}| \leq \sqrt{\delta^2 + \beta^2}. \quad (19)$$

Подставляя сюда найденные δ и β , можно сделать следующий вывод:

Если для всех измеренных величин выполняется оценка

$$|x_i - \bar{x}| \leq s \left[t^2\left(\sqrt[n]{p_1}, n\right) + \frac{1}{n} t^2(p_0, n) \right]^{1/2}, \quad (20)$$

то нет оснований считать одну из них грубо ошибочной. Если какое-либо измерение не укладывается в пределы (20), то его можно считать грубо ошибочным и отбрасывать.

Общепринятых критериев для выбора вероятности p_1 нет; естественно полагать $p_1 = p_0$.

Пример 3. Пусть проведено $n = 10$ измерений и выбрано $p_1 = p_0 = 0,9$. Тогда $p_1^{1/n} = 0,99$ и вычисления по формуле (20) при помощи таблицы 23 дают $|x_i - \bar{x}| \leq 3,3s$. Если при той же вероятности p_0 взять $n = 100$, то получим условие $|x_i - \bar{x}| \leq 4,8s$.

3. Сравнение величин. Сначала рассмотрим задачу сравнения величины x , измеряемой в эксперименте, с константой a . Величину x можно определить лишь приближенно, вычисляя среднее \bar{x} по n измерениям. Надо узнать, выполняется ли соотношение $M\xi \geq a$. В этом случае ставят две задачи, прямую и обратную:

а) по известной величине \bar{x} найти константу a , которую $M\xi$ превосходит с заданной вероятностью p_0 ;

б) найти вероятность p_0 того, что $M\xi \geq a$, где a — заданная константа.

Очевидно, если $\bar{x} < a$, то вероятность того, что $M\xi \geq a$, меньше $1/2$. Этот случай не представляет интереса, и далее будем считать, что $\bar{x} \geq a$ и $p_0 \geq 1/2$.

Задача сводится к задачам, разобранным в п. 2. Пусть по n измерениям определены \bar{x} и его стандарт s_n :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s_n^2 = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (21)$$

Число измерений будем считать не очень малым, так что \bar{x} есть случайная величина с нормальным распределением. Тогда из критерия Стьюдента (9) при учете симметрии нормального распределения следует, что для произвольно выбранной вероятности p_1 выполняется условие

$$p \{M\xi - \bar{x} \geq -t(p_1, n) s_n\} = \frac{1}{2} (1 + p_1). \quad (22)$$

Полагая $p_0 = 1/2 (1 + p_1)$, перепишем это выражение в следующем виде:

$$p \{M\xi \geq a\} = p_0, \quad a = \bar{x} - t(2p_0 - 1, n) s_n, \quad (23)$$

где $t(p_1, n)$ — заданные в таблице 23 коэффициенты Стьюдента. Тем самым, прямая задача решена: найдена константа a , которую с вероятностью p_0 превышает $M\xi$.

Обратная задача решается при помощи прямой. Перепишем формулы (23) следующим образом:

$$t(2p_0 - 1, n) = \frac{\bar{x} - a}{s_n}, \quad p \{M\xi \geq a\} = p_0. \quad (24)$$

Это значит, что надо вычислить t по известным значениям a , \bar{x} , s_n , выбрать в таблице 23 строку с данным n и найти по величине t соответствующее значение p_1 . Оно определяет искомую вероятность $p_0 = 1/2 (1 + p_1)$.

Две случайные величины. Часто требуется установить влияние некоторого фактора на исследуемую величину — например, увеличивает ли (и насколько) прочность металла определенная присадка. Для этого надо измерить прочность исходного металла x и прочность легированного металла y и сравнить эти две величины, т. е. найти $z = y - x$.

Сравниваемые величины являются случайными; так, свойства металла определенной марки меняются от плавки к плавке, поскольку сырье и режим плавки не строго одинаковы. Обозначим эти величины через ξ и η . Величина исследуемого эффекта равна $M\zeta = M\eta - M\xi$, и требуется определить, выполняется ли условие $M\zeta \geq a$.

Таким образом, задача свелась к сравнению случайной величины ζ с константой a , разобранному выше. Прямая и обратная задачи сравнения в этом случае формулируются следующим образом:

а) по результатам измерений x_i и y_j найти константу a , которую $M\zeta$ превосходит с заданной вероятностью p_0 (т. е. оценить величину исследуемого эффекта);

б) определить вероятность p_0 того, что $M\zeta \geq a$, где a — желательная величина эффекта; при $a = 0$ это означает, что надо определить вероятность, с которой $M\eta \geq M\xi$.

Для решения этих задач надо вычислить \bar{z} и дисперсию этой величины. Рассмотрим два способа их нахождения.

Независимые измерения. Измерим величину x в n экспериментах, а величину y — в m экспериментах, независимых от первых n экспериментов. Вычислим средние значения по обычным формулам:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{m} \sum_{j=1}^m y_j. \quad (25)$$

Эти средние сами являются случайными величинами, причем их стандарты (не путать со стандартами единичных измерений!) приближенно определяются несмещенными оценками:

$$s_{nx}^2 = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2, \quad s_{my}^2 = \frac{1}{m(m-1)} \sum_{j=1}^m (y_j - \bar{y})^2. \quad (26)$$

Поскольку эксперименты независимы, то случайные величины \bar{x} и \bar{y} также независимы, так что при вычислении \bar{z} их математические ожидания вычитаются, а дисперсии складываются:

$$\bar{z} \approx M\zeta = M\eta - M\xi \approx \bar{y} - \bar{x}, \quad (27)$$

$$D\bar{z} \approx s_z^2 = s_{nx}^2 + s_{my}^2. \quad (28a)$$

Несколько более точная оценка дисперсии такова:

$$s_z^2 = \frac{1}{(n+m-2)} \left(\frac{1}{n} + \frac{1}{m} \right) \left[\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{j=1}^m (y_j - \bar{y})^2 \right]. \quad (286)$$

Таким образом, \bar{z} и ее дисперсия найдены, и дальнейшие вычисления производятся по формулам (23) или (24).

Согласованные измерения. Более высокую точность дает другой способ обработки, когда в каждом из n экспериментов одновременно измеряют x и y . Например, после выпуска половины плавки в оставшийся в печи металл добавляют присадку, а затем сравнивают образцы металла из каждой половины плавки.

При этом, по существу, в каждом эксперименте измеряют сразу значение $z_i = y_i - x_i$ одной случайной величины ξ , которую надо сравнить с константой a . Обработка измерений тогда производится по формулам (21)—(24), где вместо x надо всюду подставить z .

Дисперсия при согласованных измерениях будет меньше, чем при независимых, поскольку она обусловлена только частью случайных факторов: те факторы, которые согласованно меняют ξ и η , не влияют на разброс их разности. Поэтому такой способ позволяет получить более достоверные выводы.

Пример. Любопытной иллюстрацией сравнения величин является определение победителя в тех видах спорта, где судейство ведется «на глазок» — гимнастика, фигурное катание и т. д.

Таблица 24
Судейские оценки в баллах

Судья \ Всадница	Судья					Среднее	Место
	1	2	3	4	5		
Линзенхофф	254	257	232	245	241	245,8	I
Петушкова	230	243	220	257	235	237,0	II

В таблице 24 приведен протокол соревнований по выездке на Олимпийских играх 1972 г. Видно, что разброс судейских оценок велик, причем ни одну оценку нельзя признать грубо ошибочной и откинуть. На первый взгляд кажется, что достоверность определения победителя невелика.

Рассчитаем, насколько правильно определен победитель, т. е. какова вероятность события $M\xi > 0$. Поскольку оценки обоим всадникам выставлялись одними и теми же судьями, можно воспользоваться способом согласованных измерений. По таблице 24 вычисляем $\bar{z} = +8,8$ и $s_{nz} = 5,9$; подставляя в формулу (24) эти значения и $a = 0$, получим $t(p_1, n) = 1,5$. Выбирая в таблице 23

строку $n = 5$, находим, что этому значению t соответствует $p_1 = = 2p_0 - 1 = 0,8$. Отсюда $p_0 = 0,9$, т. е. с вероятностью 90% золотая медаль присуждена правильно.

Сравнение по способу независимых измерений даст несколько худшую оценку, поскольку оно не использует информацию о том, что оценки выставляли одни и те же судьи.

Сравнение дисперсий. Пусть требуется сравнить две методики эксперимента. Очевидно, точнее та методика, у которой дисперсия σ^2 единичного измерения меньше (разумеется, если при этом не увеличивается систематическая ошибка). Значит, надо установить, выполняется ли неравенство $\sigma_x > \sigma_y$.

О дисперсиях единичных измерений судят по стандартам выборок

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{m-1} \sum_{j=1}^m (y_j - \bar{y})^2, \quad (29)$$

вычисленным соответственно по n и m измерениям. Эти стандарты сами являются случайными величинами. Однако сравнивать их на основании критерия Стьюдента нельзя, поскольку распределение s не гауссово. Нетрудно видеть, что оно является асимметричным: значения $s < 0$ невозможны, а сколь угодно большие $s > 0$ возможны.

Дисперсии сравнивают по критерию Фишера. Если

$$\frac{s_x^2}{s_y^2} > F(n, m; p_0), \quad (30)$$

то с вероятностью p_0 первая дисперсия больше второй. Коэффициенты Фишера *) для случаев $n = m$, $n = \infty$, $m = \infty$ приведены в таблице 25. При малых n , m эти коэффициенты довольно велики; поэтому различие дисперсий можно установить только в том случае, если это различие велико или велико число экспериментов.

З а м е ч а н и е. Критерий Фишера позволяет также найти отношение дисперсий. Если выполнено неравенство

$$\frac{s_x^2}{s_y^2} > aF(n, m; p_0), \quad (31)$$

то с вероятностью p_0 первая дисперсия в a раз больше второй.

Методы, изложенные в пп. 2 и 3, применимы не только к измерениям непрерывных величин, но и для суждения об очень большой партии объектов (генеральной совокупности) по небольшой случайной выборке из n объектов. Эти формулы и критерии применяются в статистике, социологии, выборочной оценке больших партий товара и т. д. В статистике и социологии законы распределения величин нередко сильно отличаются от нормального, и выяснение закона распределения играет там большую роль.

*) Их называют также *квантилями* Фишера.

Таблица 25

Коэффициенты Фишера $F(n, m; p_0)$

p_0	0,8	0,95	0,975	0,990	0,995	0,999
$n-1=m-1$	Случай $n = m$					
1	9,5	161	648	4052	16 211	$4,1 \times 10^5$
2	4,0	19	39	99	199	999
3	2,9	9,3	15	29	47	141
4	2,5	6,4	9,6	16	23	53
5	2,2	5,0	7,2	11	15	30
6	2,1	4,3	5,8	8,5	11	20
12	1,7	2,7	3,3	4,2	4,9	7,0
24	1,4	2,0	2,3	2,7	3,0	3,7
∞	1,0	1,0	1,0	1,0	1,0	1,0
$m-1$	Случай $n = \infty$					
1	16	254	1018	6366	25 465	$6,4 \times 10^5$
2	4,5	19	40	100	200	1000
3	3,0	8,5	14	26	42	124
4	2,4	5,6	8,3	13	19	44
5	2,1	4,4	6,0	9,0	12	24
6	2,0	3,7	4,9	6,9	8,9	16
12	1,5	2,3	2,7	3,4	3,9	5,4
24	1,3	1,7	1,9	2,2	2,4	3,0
∞	1,0	1,0	1,0	1,0	1,0	1,0
$n-1$	Случай $m = \infty$					
1	1,6	3,8	5,0	6,6	7,9	11
2	1,6	3,0	3,7	4,6	5,3	6,9
3	1,6	2,6	3,1	3,8	4,3	5,4
4	1,5	2,4	2,8	3,3	3,7	4,6
5	1,5	2,2	2,6	3,0	3,4	4,1
6	1,4	2,1	2,4	2,8	3,1	3,7
12	1,3	1,8	1,9	2,2	2,4	2,7
24	1,2	1,5	1,6	1,8	1,9	2,1
∞	1,0	1,0	1,0	1,0	1,0	1,0

4. Нахождение стохастической зависимости. Пусть требуется исследовать зависимость $z(x)$, причем обе величины z и x измеряются в одних и тех же экспериментах. Для этого проводят серию экспериментов при разных значениях x , стараясь сохранить прочие условия эксперимента неизменными.

Измерение каждой величины содержит случайные ошибки (систематические ошибки здесь рассматривать не будем); следовательно, эти величины являются случайными. Закономерная

связь случайных величин называется *стохастической* *). Будем рассматривать две задачи:

а) установить, существует ли (с определенной вероятностью) зависимость z от x или величина z от x не зависит;

б) если зависимость существует, описать ее количественно.

Первую задачу называют *дисперсионным анализом*, а если рассматривается функция многих переменных $z(x, y, \dots)$ — то *многофакторным дисперсионным анализом*. Вторую задачу называют *анализом регрессии*. Если случайные ошибки велики, то они могут маскировать искомую зависимость и выявить ее бывает нелегко.

Без ограничения общности можно считать, что величина x измеряется точно. В самом деле, если z от x не зависит, то ошибка δx ни на что не влияет. Если же зависимость существует, то ошибка δx эквивалентна дополнительной ошибке зависимой переменной $\delta z = (dz/dx) \delta x$.

Таким образом, достаточно рассмотреть случайную величину $\zeta(x)$, зависящую от x как от параметра. Математическое ожидание этой величины $M\zeta(x) \equiv z(x)$ зависит от x ; эта зависимость является искомой и называется *законом регрессии*.

Дисперсионный анализ. Проведем при каждом значении x_i небольшую серию измерений и определим z_{ij} ($1 \leq j \leq n_i$). Рассмотрим два способа обработки этих данных, позволяющих исследовать, имеется ли *значимая* (т. е. с принятой доверительной вероятностью) зависимость z от x .

При первом способе вычисляют стандарты выборки единичного измерения по каждой серии отдельно и по всей совокупности измерений:

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2, \quad s^2 = \frac{1}{N - 1} \sum_{i,j} (z_{ij} - \bar{z})^2, \quad (32)$$

где $N = \sum_i n_i$ — полное число измерений, а

$$\bar{z}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} z_{ij}, \quad \bar{z} = \frac{1}{N} \sum_{i,j} z_{ij} \quad (33)$$

являются средними значениями соответственно по каждой серии и по всей совокупности измерений.

Сравним дисперсию совокупности измерений $\sigma^2 \approx s^2$ с дисперсиями отдельных серий $\sigma_i^2 \approx s_i^2$. Если окажется, что при выбранном уровне достоверности p_0 можно считать $\sigma > \sigma_i$ для всех i , то зависимость z от x имеется. Если достоверного превышения

*) С такой связью мы уже встречались в стохастических задачах нахождения корня уравнения (гл. V, § 2, п. 4) и минимума функции (гл. VII, § 1, п. 4).

нет, то зависимость не поддается обнаружению (при данной точности эксперимента и принятом способе обработки).

Дисперсии сравнивают по критерию Фишера (30). Поскольку стандарт s определен по полному числу измерений N , которое обычно достаточно велико, то почти всегда можно пользоваться коэффициентами Фишера $F(\infty, m; p_0)$, приведенными в таблице 25.

Второй способ анализа заключается в сравнении средних \bar{z}_i при разных значениях x_i между собой. Величины \bar{z}_i являются случайными и независимыми, причем их собственные стандарты выборки равны

$$s_{ni} = s_i / \sqrt{n_i}.$$

Поэтому их сравнивают по схеме независимых измерений, описанной в п. 3. Если различия \bar{z}_i значимы, т. е. превышают доверительный интервал, то факт зависимости z от x установлен; если различия всех \bar{z}_i незначимы, то зависимость не поддается обнаружению.

Многофакторный анализ имеет некоторые особенности. Величину $z(x, y)$ целесообразно измерять в узлах прямоугольной сетки (x_i, y_k) , чтобы удобнее было исследовать зависимость от одного аргумента, фиксируя другой аргумент. Проводить серию измерений в каждом узле многомерной сетки слишком трудоемко. Достаточно провести серии измерений в нескольких узлах сетки, чтобы оценить дисперсию единичного измерения; в остальных узлах можно ограничиться однократными измерениями. Дисперсионный анализ при этом проводят по первому способу.

З а м е ч а н и е 1. Если измерений много, то в обоих способах отдельные измерения или серии могут с заметной вероятностью довольно сильно отклониться от своего математического ожидания. Это надо учитывать, выбирая доверительную вероятность p_0 достаточно близкой к 1 (как это делалось в п. 2 при установлении пределов, отделяющих допустимые случайные ошибки от грубых).

А н а л и з р е г р е с с и и. Пусть дисперсионный анализ указал, что зависимость z от x есть. Как ее количественно описать?

Для этого аппроксимируем искомую зависимость некоторой функцией $z(x) \approx f(x, \mathbf{a})$, $\mathbf{a} = \{a_1, a_2, \dots, a_m\}$. Оптимальные значения параметров a_k найдем методом наименьших квадратов, решая задачу

$$\sum_{i=1}^N \omega(x_i) [z_i - f(x_i, \mathbf{a})]^2 = \min, \quad (34)$$

где $\omega(x_i)$ — веса измерений, выбираемые обратно пропорционально квадрату погрешности измерения в данной точке (т. е. $\omega_i \sim (Dz_i)^{-1}$). Эта задача была разобрана в главе II, § 2. Остановимся здесь лишь на тех особенностях, которые вызваны присутствием больших случайных ошибок.

Вид $f(x, a)$ подбирают либо из теоретических соображений о природе зависимости $z(x)$, либо формально, сравнивая график $z(x)$ с графиками известных функций. Если формула подобрана из теоретических соображений и правильно (с точки зрения теории) передает асимптотику $z(x)$, то обычно она позволяет не только неплохо аппроксимировать совокупность экспериментальных данных, но и экстраполировать найденную зависимость на другие диапазоны значений x . Формально подобранная функция $f(x, a)$ может удовлетворительно описывать эксперимент, но редко пригодна для экстраполяции.

Проще всего решить задачу (34), если $f(x, a)$ является алгебраическим многочленом $\sum a_k x^k$. Однако такой формальный выбор функции редко оказывается удовлетворительным. Обычно хорошие формулы зависят от параметров нелинейно (*трансцендентная регрессия*). Трансцендентную регрессию наиболее удобно строить, подбирая такую *выравнивающую* замену переменных $Z(z)$, $X(x)$, чтобы зависимость $Z(X)$ была почти линейной (см. гл. II, § 1, п. 8). Тогда ее нетрудно аппроксимировать алгебраическим многочленом: $Z \approx P(X, a)$.

Выравнивающую замену переменных ищут, используя теоретические соображения и учитывая асимптотику $z(x)$. Дальше будем считать, что такая замена уже сделана.

Замечание 2. При переходе к новым переменным задача метода наименьших квадратов (34) принимает вид

$$\sum_{i=1}^N W_i [Z_i - P(X_i, a)]^2 = \min, \tag{35}$$

где новые веса связаны с исходными соотношениями

$$W_i = \left(\frac{dZ}{dz} \right)_i^{-2} \omega_i. \tag{36}$$

Поэтому, даже если в исходной постановке (34) все измерения имели одинаковую точность, так что $\omega_i \equiv 1$, то для выравнивающих переменных веса не будут одинаковыми.

Корреляционный анализ. Надо проверить, действительно ли замена переменных была выравнивающей, т. е. близка ли зависимость $Z(X)$ к линейной. Это можно сделать, вычислив коэффициент парной корреляции

$$\rho = \frac{M[(\xi - M\xi)(\zeta - M\zeta)]}{\sqrt{D\xi \cdot D\zeta}} \approx \frac{\sum_{i=1}^N W_i (X_i - \bar{X}_i)(Z_i - \bar{Z}_i)}{\sqrt{\sum_{i=1}^N W_i (X_i - \bar{X}_i)^2 \cdot \sum_{i=1}^N W_i (Z_i - \bar{Z}_i)^2}}. \tag{37}$$

Нетрудно показать, что всегда выполняется соотношение $|\rho| \leq 1$.

Если зависимость $Z(X)$ строго линейная (и не содержит случайных ошибок), то $\rho = +1$ или $\rho = -1$ в зависимости от знака наклона прямой. Чем меньше $|\rho|$, тем менее зависимость $Z(X)$ похожа на линейную. Поэтому, если $|\rho| \approx 1$, а число измерений N достаточно велико, то выравнивающие переменные выбраны удовлетворительно.

Подобные заключения о характере зависимости по коэффициентам корреляции называют *корреляционным анализом*.

При корреляционном анализе не требуется, чтобы в каждой точке x_i проводилась серия измерений. Достаточно в каждой точке сделать одно измерение, но зато взять побольше точек на исследуемой кривой, что часто делают в физических экспериментах.

З а м е ч а н и е 3. Существуют критерии близости $|\rho|$ к 1, позволяющие указать, является ли зависимость $Z(X)$ практически линейной. Мы на них не останавливаемся, поскольку далее будет рассмотрен выбор степени аппроксимирующего многочлена.

З а м е ч а н и е 4. Соотношение $|\rho| \approx 0$ указывает на отсутствие линейной зависимости $Z(X)$, но не означает отсутствия какой-либо зависимости. Так, если $Z = X^2$ на отрезке $-1 \leq X \leq 1$, то $\rho = 0$.

Оптимальная степень многочлена. Подставим в задачу (35) аппроксимирующий многочлен степени m :

$$P(X) = \sum_{k=0}^m a_k X^k. \quad (38)$$

Тогда оптимальные значения параметров a_k удовлетворяют системе линейных уравнений (2.43):

$$\sum_{l=0}^m A_{kl} a_l = B_k, \quad 0 \leq k \leq m, \quad (39)$$

$$A_{kl} = \sum_{i=1}^N W_i X_i^{k+l}, \quad B_k = \sum_{i=1}^N W_i Z_i X_i^k,$$

и найти их нетрудно. Но как выбрать степень многочлена?

Для ответа на этот вопрос вернемся к исходным переменным и вычислим дисперсию аппроксимационной формулы с найденными коэффициентами. Несмещенная оценка этой дисперсии такова*):

$$D_m = D[f(x, a_0, a_1, \dots, a_m)] \approx \approx \frac{N}{N-m-1} \sum_{i=1}^N w_i [z_i - f(x_i, a)]^2 / \sum_{i=1}^N w_i. \quad (40)$$

*) В формулах типа (32) имелся делитель вида $N-1$; он связан с тем, что при вычислении стандарта выборки используется одна величина \bar{z} , определенная по той же выборке. В (40) используется $m+1$ коэффициентов a_k , определенных по выборке, поэтому появляется делитель $N-(m+1)$.

Очевидно, при увеличении степени многочлена m дисперсия (40) будет убывать: чем больше взято коэффициентов, тем точнее можно аппроксимировать экспериментальные точки.

Сравним D_m с дисперсиями s_i^2 единичных измерений (32), определенными по небольшим сериям экспериментов хотя бы в нескольких точках x_i . Если $D_m > s_i^2$ для всех i , то погрешность аппроксимации больше погрешности, с которой измерены значения z_i . Надо увеличивать m до тех пор, пока отличие D_m от s_i^2 хотя бы для одного i не перестанет быть значимым по критерию Фишера (30). Наоборот, если $D_m < s_i^2$, то надо уменьшать m .

Если полученная таким образом оптимальная степень m_0 удовлетворяет условию $m_0 \ll N$, то выравнивающие переменные выбраны удачно; если $m_0 \sim N$, то следует подобрать другую замену переменных.

Замечание 5. Описанный способ нахождения оптимального числа параметров m_0 можно применять при произвольном виде функции $f(x, a)$; но сами коэффициенты в этом случае вычисляются не по формулам (39).

Точность коэффициентов. Коэффициенты a_k определяются по случайным величинам z_i и поэтому сами являются случайными величинами. Какие их значащие цифры достоверны, а какие можно отбросить?

На первую половину вопроса ответить нетрудно. Проведем математический эксперимент. Зная дисперсию единичных измерений s_i^2 , искусственно внесем в величины z_i случайные ошибки δz_i , распределенные по нормальному закону с дисперсиями s_i^2 (это делается методами Монте-Карло), и вычислим соответствующие a_k .

Повторим эту процедуру многократно. Для каждого a_k получим набор случайных значений, по которому вычислим среднее \bar{a}_k и стандарт $s(\bar{a}_k)$. Отсюда по критерию Стьюдента (9) найдем для \bar{a}_k доверительный интервал и, тем самым, выясним, какие значащие цифры коэффициента достоверны.

Однако, вообще говоря, *недоверительные цифры коэффициента a_k нельзя отбрасывать*. Коэффициенты a_k можно округлять только все одновременно, меняя их на взаимно согласованные величины Δa_k . Для такого округления многочлен $P_m(X)$ представляют в виде линейной комбинации:

$$P_m(X) \equiv \sum_{k=0}^m a_k X^k = \sum_{i=0}^m \gamma_i Q_i(X), \quad (41)$$

где $Q_i(X)$ — алгебраические многочлены, ортогональные на системе точек X_i ($1 \leq i \leq N$) с весами W_i^*). Коэффициенты γ_i можно округлять независимо друг от друга в пределах их доверительных интервалов.

*) См. Приложение.

ЗАДАЧИ

1. Для примера, приведенного в таблице 24, определить достоверность победителя способом независимых измерений и сравнить результат с результатом, полученным способом согласованных измерений.

2. Учитывая, что веса в формулах (34) и (35) обратно пропорциональны дисперсиям, обосновать выражение веса (36).

3. Показать, что коэффициент парной корреляции (37) всегда по модулю не превышает 1.

4. Найти коэффициент парной корреляции (37) на отрезке $-1 \leq x \leq 1$ для а) линейной функции $z = ax + b$, б) квадратичной функции $z = ax^2 + b$; предполагается, что $\omega_i \equiv 1$.

5. Установить связь между коэффициентом парной корреляции (37) и наклоном сглаженной кривой (3.28).