

Глава 6

Распространение ошибок

Содержание

6.1.	Физическая задача: измерение сечения	76
6.2.	Что называется ошибкой	76
6.3.	Одна измеряемая величина	77
6.4.	Где линейное приближение перестаёт работать	78
6.5.	Разброс оценки стандартного отклонения	79
6.6.	Несколько измеряемых величин	83
6.7.	Ковариационная матрица	87
6.8.	Задачи	88
	Литература	89

Аннотация

Физический результат обычно вычисляется по нескольким величинам, которые измеряет детектор. Вместе с результатом требуется найти его неопределённость. В этой главе мы получим линейный закон распространения ошибок, проверим его на нелинейной функции и запишем в матричном виде с учётом корреляций. Отдельный раздел посвящён разбросу оценок дисперсии и стандартного отклонения.

6.1. Физическая задача: измерение сечения

Эксперимент зарегистрировал $N = 1000$ сигнальных событий. Эффективность отбора равна

$$\varepsilon = 0.80 \pm 0.02, \quad (6.1)$$

а интегральная светимость составляет

$$L = (100 \pm 2) \text{ фб}^{-1}. \quad (6.2)$$

В упрощённой модели фон отсутствует, число событий имеет распределение Пуассона, а эффективность и светимость измерены независимо от N и друг от друга. Оценка сечения процесса равна

$$\hat{\sigma}_{\text{proc}} = \frac{N}{\varepsilon L} = 12.5 \text{ фб}. \quad (6.3)$$

Для пуассоновского счёта стандартное отклонение числа событий можно оценить как $\sigma_N = \sqrt{N}$. Относительная неопределённость сечения в линейном приближении тогда равна

$$\begin{aligned} \left(\frac{\sigma_{\hat{\sigma}_{\text{proc}}}}{\hat{\sigma}_{\text{proc}}} \right)^2 &\simeq \left(\frac{\sigma_N}{N} \right)^2 + \left(\frac{\sigma_\varepsilon}{\varepsilon} \right)^2 + \left(\frac{\sigma_L}{L} \right)^2 \\ &= \left(\frac{\sqrt{1000}}{1000} \right)^2 + \left(\frac{0.02}{0.80} \right)^2 + \left(\frac{2}{100} \right)^2 = 0.045^2. \end{aligned} \quad (6.4)$$

Результат можно записать как

$$\hat{\sigma}_{\text{proc}} = (12.5 \pm 0.6) \text{ фб}. \quad (6.5)$$

В одной строке под корнем встретились статистика событий, эффективность детектора и светимость ускорителя. Формулу мы пока использовали как готовую. Ниже выведем её и увидим, где в ней должны стоять ковариации. В реальном анализе N получают после вычитания фона, а эффективность и нормировки могут зависеть от общих калибровок. Тогда три квадрата уже не исчерпывают задачу.

6.2. Что называется ошибкой

В названии главы используется привычное физикам выражение «распространение ошибок». В строгой метрологической терминологии ошибка — это разность между измеренным и опорным значениями, а неопределённость описывает разброс возможных результатов. В формулах этой главы речь идёт о стандартной неопределённости, которую мы обозначаем через σ_X .

Стандартное отклонение описывает масштаб случайного разброса. Смещение метода, неопределённость калибровки и физические границы параметра требуют отдельного учёта. Общие правила представления неопределённостей измерения собраны в руководстве JCGM [1].

6.3. Одна измеряемая величина

Пусть детектор измеряет случайную величину X , а результат вычисляется по формуле

$$Y = f(X). \quad (6.6)$$

Обозначим

$$E[X] = \mu_X, \quad \text{Var}(X) = \sigma_X^2. \quad (6.7)$$

6.3.1. Линейное распространение ошибки

На масштабе σ_X функцию f можно заменить её касательной в точке μ_X , если кривизна функции на этом интервале мала.

Вывод

6.3.1.1. Формула для одной переменной

Первый порядок разложения Тейлора даёт

$$f(X) \approx f(\mu_X) + f'(\mu_X)(X - \mu_X). \quad (6.8)$$

В этом приближении

$$E[Y] \approx f(\mu_X), \quad Y - E[Y] \approx f'(\mu_X)(X - \mu_X). \quad (6.9)$$

Квадрат производной можно вынести из математического ожидания:

$$\begin{aligned} \text{Var}(Y) &\approx E[f'(\mu_X)^2(X - \mu_X)^2] \\ &= f'(\mu_X)^2 \sigma_X^2. \end{aligned} \quad (6.10)$$

Поэтому

$$\sigma_Y \approx |f'(\mu_X)| \sigma_X. \quad (6.11)$$

Производная в уравнении 6.11 показывает, во сколько раз небольшое отклонение входной величины растягивается или сжимается при переходе к Y . Для линейной функции формула точна. Для нелинейной её точность определяется размером области, занятой распределением X .

6.3.2. Поперечный импульс в магнитном поле

Радиус R траектории заряженной частицы в однородном магнитном поле связан с поперечным импульсом соотношением

$$p_T = |q|BR. \quad (6.12)$$

В единицах, обычных для физики частиц,

$$p_T[\text{ГэВ}/c] \approx 0.2998 |q/e| B[\text{Тл}] R[\text{м}]. \quad (6.13)$$

Если q и B считаются известными, неопределённость создаёт измерение радиуса. Поскольку зависимость линейна,

$$\sigma_{p_T} = |q|B\sigma_R, \quad \frac{\sigma_{p_T}}{p_T} = \frac{\sigma_R}{R}. \quad (6.14)$$

Ошибка радиуса переходит в такую же относительную ошибку импульса. Если неопределённость магнитного поля существенна, B становится второй входной величиной. Такой случай разберём в разделе о ковариациях.

6.3.3. Степенная зависимость

Для функции

$$Y = X^a \quad (6.15)$$

производная равна

$$f'(x) = ax^{a-1}. \quad (6.16)$$

В точке $x = \mu_X$ уравнение 6.11 принимает вид

$$\frac{\sigma_Y}{|f(\mu_X)|} \approx |a| \frac{\sigma_X}{|\mu_X|}. \quad (6.17)$$

Несколько часто встречающихся функций собраны в таблице 6.1. Здесь $x_0 = \mu_X$ и $y_0 = f(x_0)$.

Таблица 6.1. Линейное распространение ошибки для часто встречающихся функций

Функция $f(x)$	Линейная неопределённость	Условие на точку разложения
$ax + b$	$\sigma_Y = a \sigma_X$	формула точна
x^a	$\sigma_Y/ y_0 \approx a \sigma_X/ x_0 $	функция определена около x_0
$1/x$	$\sigma_Y/ y_0 \approx \sigma_X/ x_0 $	x_0 далёк от нуля
$\ln x$	$\sigma_Y \approx \sigma_X/x_0$	$x_0 > 0$, вероятность $X \leq 0$ мала
e^x	$\sigma_Y/y_0 \approx \sigma_X$	X безразмерна

Таблица даёт локальные формулы. У отношения, логарифма и других функций с границей одной малости σ_X недостаточно. Распределение X должно с большой вероятностью оставаться в области, где функция гладкая.

6.4. Где линейное приближение перестаёт работать

Касательная описывает функцию в малой окрестности точки разложения. Размер этой окрестности задаёт σ_X . Линейная формула требует проверки в четырёх типичных ситуациях:

- производная заметно меняется на интервале порядка σ_X ;
- $f'(\mu_X)$ близка к нулю и первым ненулевым оказывается квадратичный член;
- рядом находится полюс, порог или физическая граница;

- преобразованное распределение заметно асимметрично.

Возьмём $Y = X^2$ и $\mu_X = 0$. Касательная к параболе в нуле горизонтальна, и линейная формула даёт $\sigma_Y \approx 0$. Сама величина X^2 при этом флуктуирует. Нулевой ответ сообщает о непригодности первого порядка разложения.

6.4.1. Нелинейное преобразование $Y = X + aX^2$

Рассмотрим модель из лекции:

$$X \sim \mathcal{N}(1, \sigma_X^2), \quad Y = X + aX^2. \quad (6.18)$$

Касательная в точке $\mu_X = 1$ даёт

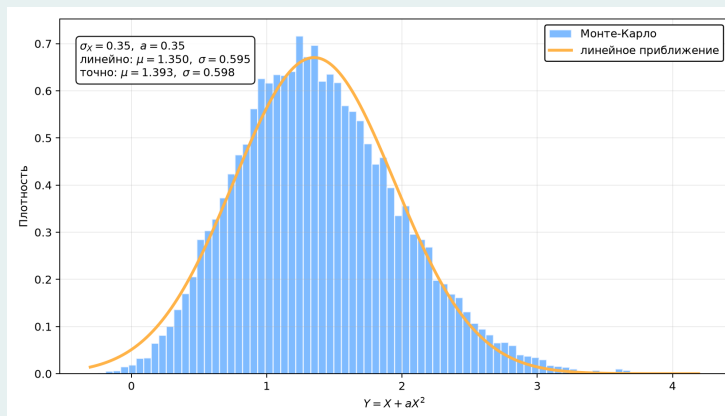
$$\mathbb{E}[Y] \approx 1 + a, \quad \sigma_Y \approx |1 + 2a|\sigma_X. \quad (6.19)$$

Для гауссовой величины моменты можно вычислить без приближения:

$$\mathbb{E}[Y] = 1 + a + a\sigma_X^2, \quad (6.20)$$

$$\text{Var}(Y) = (1 + 2a)^2\sigma_X^2 + 2a^2\sigma_X^4. \quad (6.21)$$

Член $a\sigma_X^2$ сдвигает среднее, а $2a^2\sigma_X^4$ добавляет дисперсию. Оба исчезают в линейном приближении.



Интерактив. Нелинейное преобразование $Y = X + aX^2$

Меняйте ширину исходного гаусса и коэффициент a . Гистограмма показывает распределение Y , а гладкая кривая — линейный прогноз.



На рисунке 6.1 показан режим, в котором кривизна уже существенна на масштабе входного распределения.

Здесь расхождение видно по положению максимума, ширине и длинному правому хвосту. Запись $y \pm \sigma_Y$ сохраняет два числа, но уже не описывает форму распределения. В такой ситуации используют преобразование полного распределения или численное моделирование.

6.5. Разброс оценки стандартного отклонения

До сих пор μ и σ были известными параметрами распределения. В эксперименте их оценивают по выборке

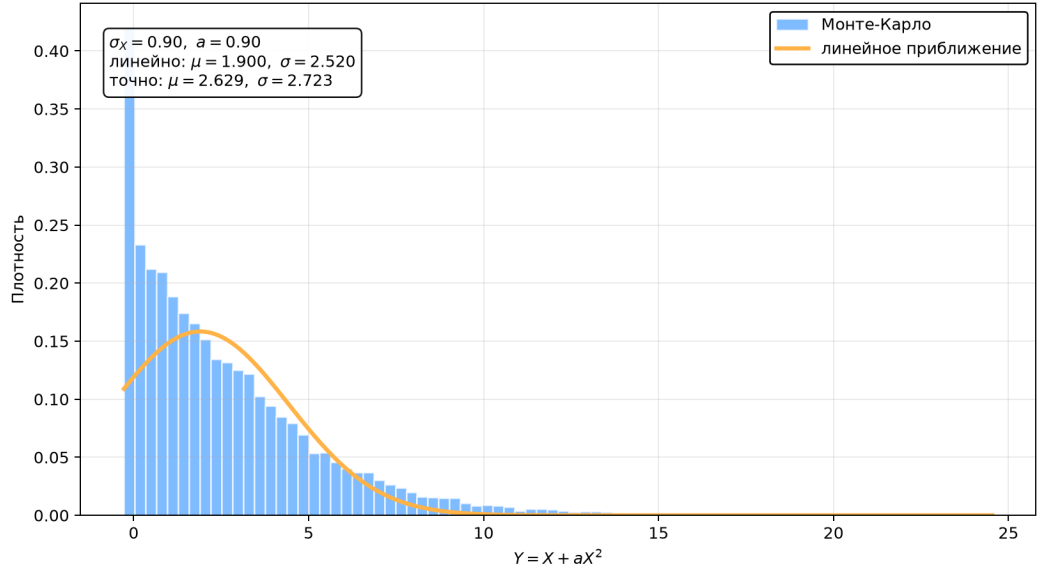


Рис. 6.1. Распределение $Y = X + aX^2$ при $\sigma_X = 0.9$ и $a = 0.9$. Синяя гистограмма получена методом Монте-Карло, оранжевая кривая показывает гауссово распределение с параметрами линейного приближения.

$$X_1, \dots, X_N \sim \mathcal{N}(\mu, \sigma^2), \quad (6.22)$$

где измерения независимы. Определим оценки среднего, дисперсии и стандартного отклонения:

$$\hat{\mu} = \bar{X} = \frac{1}{N} \sum_{k=1}^N X_k, \quad (6.23)$$

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{k=1}^N (X_k - \bar{X})^2, \quad \hat{\sigma} = \sqrt{\hat{\sigma}^2}. \quad (6.24)$$

Если повторить весь эксперимент, выборка изменится вместе со всеми тремя оценками. Поэтому у $\hat{\sigma}$ есть собственное распределение и собственная дисперсия.

Для выборочного среднего

$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{N}. \quad (6.25)$$

Параметр σ^2 неизвестен, и в оценке дисперсии среднего его заменяют выборочной дисперсией:

$$\widehat{\text{Var}}(\hat{\mu}) = \frac{\hat{\sigma}^2}{N}, \quad \hat{\sigma}_{\hat{\mu}} = \frac{\hat{\sigma}}{\sqrt{N}}. \quad (6.26)$$

Возникает следующий вопрос: насколько устойчиво значение $\hat{\sigma}$, найденное по конечной выборке?

6.5.1. Почему в выборочной дисперсии стоит $N - 1$

Введём отклонения от истинного среднего

$$Y_k = X_k - \mu. \quad (6.27)$$

Для независимой выборки с конечной дисперсией

$$\mathbb{E}[Y_k] = 0, \quad \mathbb{E}[Y_k^2] = \sigma^2. \quad (6.28)$$

Гауссовость для доказательства несмещённости выборочной дисперсии не нужна.

Вывод

6.5.1.1. Несмещённость выборочной дисперсии

Обозначим две суммы

$$A = \sum_{k=1}^N Y_k^2, \quad B = \sum_{k=1}^N Y_k. \quad (6.29)$$

Разность выборочного и истинного среднего равна

$$\bar{X} - \mu = \frac{B}{N}. \quad (6.30)$$

Сумма квадратов отклонений от выборочного среднего принимает вид

$$\begin{aligned} T &= \sum_{k=1}^N (X_k - \bar{X})^2 \\ &= \sum_{k=1}^N \left(Y_k - \frac{B}{N} \right)^2 \\ &= A - \frac{B^2}{N}. \end{aligned} \quad (6.31)$$

Независимость и центрирование дают

$$\mathbb{E}[A] = N\sigma^2, \quad \mathbb{E}[B^2] = N\sigma^2. \quad (6.32)$$

Отсюда

$$\mathbb{E}[T] = N\sigma^2 - \frac{1}{N}N\sigma^2 = (N-1)\sigma^2. \quad (6.33)$$

После деления на $N-1$

$$\boxed{\mathbb{E}[\hat{\sigma}^2] = \sigma^2.} \quad (6.34)$$

Выборочные отклонения связаны условием

$$\sum_{k=1}^N (X_k - \bar{X}) = 0. \quad (6.35)$$

Одно из N отклонений определяется через остальные. Поэтому сумма их квадратов содержит $N-1$ степеней свободы, и такой же делитель появляется в несмещённой оценке дисперсии.

6.5.2. Дисперсия оценки дисперсии

Теперь понадобится гауссовость. Для центрированной гауссовой величины

$$\mathbb{E}[Y_k^4] = 3\sigma^4. \quad (6.36)$$

Поскольку $\hat{\sigma}^2 = T/(N-1)$, достаточно найти дисперсию T .

Вывод**6.5.2.1. Вычисление $\text{Var}(\hat{\sigma}^2)$**

Из уравнения 6.31

$$T^2 = A^2 - \frac{2}{N}AB^2 + \frac{1}{N^2}B^4. \quad (6.37)$$

Первый момент вычисляется раскрытием квадрата суммы:

$$\begin{aligned} E[A^2] &= \sum_{k=1}^N E[Y_k^4] + 2 \sum_{i < j} E[Y_i^2 Y_j^2] \\ &= 3N\sigma^4 + N(N-1)\sigma^4 \\ &= N(N+2)\sigma^4. \end{aligned} \quad (6.38)$$

Для смешанного момента выделим одно слагаемое, $B = Y_i + C_i$, где $C_i = \sum_{j \neq i} Y_j$. Величины Y_i и C_i независимы, $E[C_i] = 0$ и $E[C_i^2] = (N-1)\sigma^2$. Поэтому

$$\begin{aligned} E[Y_i^2 B^2] &= E[Y_i^2 (Y_i + C_i)^2] \\ &= 3\sigma^4 + (N-1)\sigma^4 \\ &= (N+2)\sigma^4, \end{aligned} \quad (6.39)$$

и после суммирования по i

$$E[AB^2] = N(N+2)\sigma^4. \quad (6.40)$$

Сумма B также гауссова, причём $\text{Var}(B) = N\sigma^2$. Её четвёртый момент равен

$$E[B^4] = 3N^2\sigma^4. \quad (6.41)$$

Подстановка трёх средних в уравнение 6.37 даёт

$$E[T^2] = (N^2 - 1)\sigma^4. \quad (6.42)$$

Вместе с уравнением 6.33 это приводит к

$$\text{Var}(T) = 2(N-1)\sigma^4. \quad (6.43)$$

Искомая дисперсия оценки равна

$$\boxed{\text{Var}(\hat{\sigma}^2) = \frac{2\sigma^4}{N-1}.} \quad (6.44)$$

Уравнение 6.44 является точным для независимой гауссовой выборки. Для другого исходного распределения в ответ входит его четвёртый центральный момент.

6.5.3. Разброс оценки стандартного отклонения

Оценка $\hat{\sigma}$ связана с $\hat{\sigma}^2$ функцией квадратного корня. Применим к ней формулу 6.11 в точке σ^2 :

$$f(y) = \sqrt{y}, \quad f'(\sigma^2) = \frac{1}{2\sigma}. \quad (6.45)$$

Получаем

$$\text{Var}(\hat{\sigma}) \simeq \frac{1}{4\sigma^2} \text{Var}(\hat{\sigma}^2) = \frac{\sigma^2}{2(N-1)}. \quad (6.46)$$

Относительный разброс оценки стандартного отклонения равен

$$\frac{\sqrt{\text{Var}(\hat{\sigma})}}{\sigma} \simeq \frac{1}{\sqrt{2(N-1)}}. \quad (6.47)$$

Здесь снова использовано линейное приближение, теперь уже для квадратного корня. При конечном N оценка $\hat{\sigma}$ немного смещена вниз; с ростом выборки смещение исчезает, и формула становится точнее.

Для относительного разброса около 10%

$$\frac{1}{\sqrt{2(N-1)}} \simeq 0.1, \quad N \simeq 51. \quad (6.48)$$

Пятьдесят измерений нужны лишь для того, чтобы сама оценка ширины флуктуировала примерно на десять процентов. Ширину распределения обычно измерить труднее, чем его среднее.

6.5.4. Десять гауссовских измерений

k	X_k	$X_k - \bar{X}$	$(X_k - \bar{X})^2$
1	1.250	0.377	0.142
2	1.008	0.134	0.018
3	1.850	0.976	0.953
4	1.533	0.659	0.435
5	-0.080	-0.953	0.909
6	2.278	1.404	1.972
7	3.534	2.661	7.080
8	-0.146	-1.020	1.040
9	-0.648	-1.521	2.314
10	-1.844	-2.717	7.383

Истинные параметры
 $\mu = 0.800$, $\sigma = 1.500$, $N = 10$

Оценки по выборке

$\bar{X} = 0.874$

$\hat{\sigma}^2 = 2.472$

$\hat{\sigma} = 1.572$

Неопределённость среднего

$\hat{\sigma}/\sqrt{N} = 0.497$

$\text{Var}(\bar{X}) = 0.247$

Неопределённость $\hat{\sigma}$

$\hat{\sigma}/\sqrt{2(N-1)} = 0.371$

$\text{Var}(\hat{\sigma}) = 0.137$

Интерактив. Десять гауссовских измерений

Меняйте параметры распределения и начальное состояние генератора. Таблица показывает выборку, оценки $\hat{\mu}$, $\hat{\sigma}^2$, $\hat{\sigma}$ и их расчётные неопределённости.



Десять чисел выглядят вполне убедительно, пока рядом не написаны истинные параметры. Новая выборка даёт другие $\hat{\mu}$ и $\hat{\sigma}$. Формулы выше описывают разброс этих результатов в серии повторений.

6.6. Несколько измеряемых величин

В задаче о сечении использовались три входные величины. Теперь выведем формулу, которая учитывает их совместные флуктуации.

6.6.1. Ковариация и ошибка суммы

Ковариация X и Y равна

$$\text{cov}(X, Y) = \text{E}[(X - \mu_X)(Y - \mu_Y)]. \quad (6.49)$$

Для суммы $S = X + Y$

$$\text{Var}(S) = \sigma_X^2 + \sigma_Y^2 + 2 \text{cov}(X, Y). \quad (6.50)$$

Коэффициент корреляции убирает размерность ковариации:

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \quad -1 \leq \rho_{XY} \leq 1. \quad (6.51)$$

Для суммы и разности

$$\sigma_{X \pm Y}^2 = \sigma_X^2 + \sigma_Y^2 \pm 2\rho_{XY}\sigma_X\sigma_Y. \quad (6.52)$$

Положительная корреляция увеличивает ошибку суммы и уменьшает ошибку разности. При отрицательной корреляции знаки меняются. Складывать ошибки в квадратуре можно при нулевой ковариации.

Независимость влечёт $\rho_{XY} = 0$. Обратное утверждение для произвольного распределения неверно: при нулевой корреляции нелинейная зависимость может сохраняться.

6.6.2. Оценка корреляции по выборке

Пусть получены N независимых пар (x_k, y_k) из одного совместного распределения. Выборочная ковариация равна

$$\widehat{\text{cov}}(X, Y) = \frac{1}{N-1} \sum_{k=1}^N (x_k - \bar{x})(y_k - \bar{y}), \quad (6.53)$$

а выборочный коэффициент корреляции —

$$\hat{\rho} = \frac{\widehat{\text{cov}}(X, Y)}{s_X s_Y}. \quad (6.54)$$

При небольшом N значение $\hat{\rho}$ заметно меняется от выборки к выборке. В детекторных задачах ковариации получают также из моделирования, калибровочных данных и вариаций систематических параметров.

6.6.3. Общая формула распространения

Пусть

$$F = f(X_1, \dots, X_n), \quad (6.55)$$

а средние входных величин образуют вектор $\mu = (\mu_1, \dots, \mu_n)^T$. Первый порядок разложения имеет вид

$$F - \mathbb{E}[F] \approx \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_{\mu} (X_i - \mu_i). \quad (6.56)$$

Вывод

6.6.3.1. Дисперсия функции нескольких переменных

Обозначим производные в точке μ через

$$g_i = \left. \frac{\partial f}{\partial x_i} \right|_{\mu}. \quad (6.57)$$

После возведения линейного разложения в квадрат

$$\begin{aligned}\text{Var}(F) &\simeq \mathbb{E} \left[\left(\sum_i g_i (X_i - \mu_i) \right) \left(\sum_j g_j (X_j - \mu_j) \right) \right] \\ &= \sum_{i,j} g_i \text{cov}(X_i, X_j) g_j.\end{aligned}\quad (6.58)$$

Итак,

$$\text{Var}(F) \simeq \sum_{i,j} \frac{\partial f}{\partial x_i} \text{cov}(X_i, X_j) \frac{\partial f}{\partial x_j}.\quad (6.59)$$

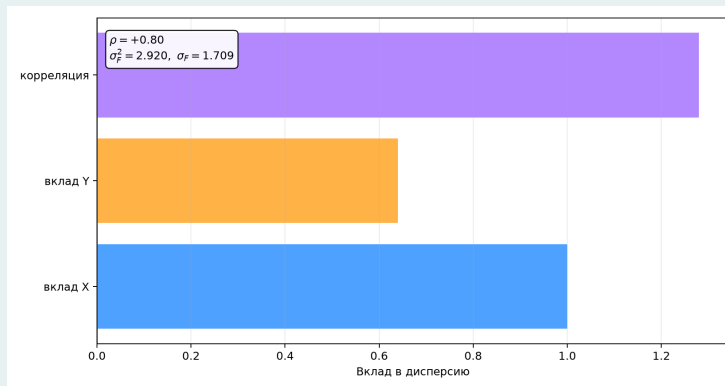
Все производные в уравнении 6.59 вычисляются в одной точке μ . На практике неизвестные средние заменяют полученными центральными значениями.

Для двух переменных формула разворачивается в три слагаемых:

$$\begin{aligned}\sigma_F^2 &\simeq \left(\frac{\partial f}{\partial x} \right)^2 \sigma_X^2 + \left(\frac{\partial f}{\partial y} \right)^2 \sigma_Y^2 \\ &\quad + 2\rho_{XY} \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \sigma_X \sigma_Y.\end{aligned}\quad (6.60)$$

Знак корреляционного вклада определяется одновременно знаком ρ_{XY} и знаками двух производных.

6.6.4. Корреляционный вклад в дисперсию



Интерактив. Корреляционный вклад в дисперсию

Меняйте производные, стандартные отклонения и ρ . Три полосы показывают два диагональных и один корреляционный вклад в σ_F^2 .



На рисунке 6.2 производные имеют одинаковый знак, а корреляция отрицательна. Корреляционный член вычитается из двух положительных вкладов.

6.6.5. Отношение двух величин

Для

$$R = \frac{X}{Y}\quad (6.61)$$

две производные равны

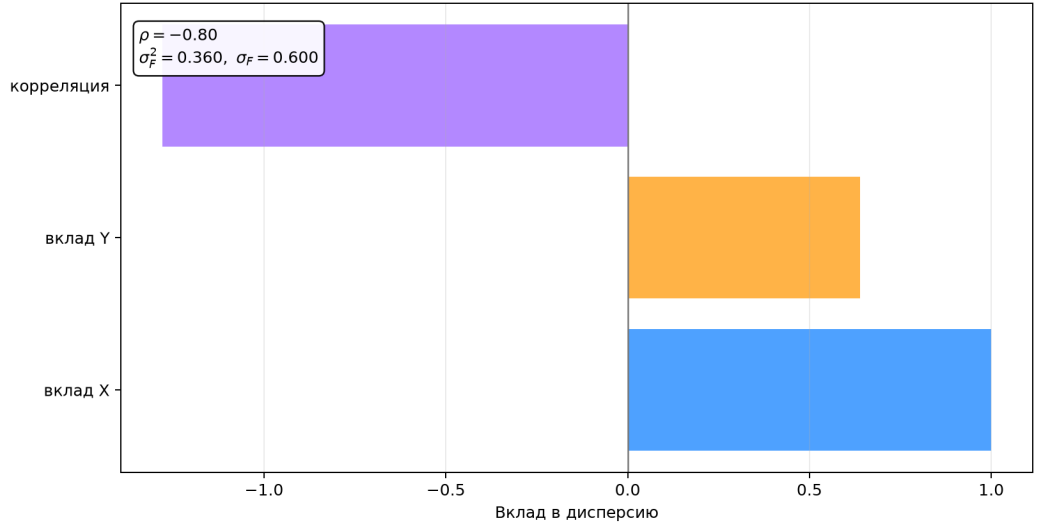


Рис. 6.2. Вклады в дисперсию функции двух переменных при $\partial f/\partial x = \partial f/\partial y = 1$, $\sigma_X = 1$, $\sigma_Y = 0.8$ и $\rho = -0.8$. Отрицательный корреляционный член уменьшает полную дисперсию до 0.36.

$$\frac{\partial R}{\partial X} = \frac{1}{Y}, \quad \frac{\partial R}{\partial Y} = -\frac{X}{Y^2}. \quad (6.62)$$

Линейная относительная неопределённость составляет

$$\left(\frac{\sigma_R}{R}\right)^2 \simeq \left(\frac{\sigma_X}{X}\right)^2 + \left(\frac{\sigma_Y}{Y}\right)^2 - 2\rho_{XY} \frac{\sigma_X}{X} \frac{\sigma_Y}{Y}. \quad (6.63)$$

Производные по числителю и знаменателю имеют разные знаки. Поэтому положительная корреляция уменьшает ошибку отношения. Формула требует, чтобы знаменатель с большой вероятностью оставался вдали от нуля. В главе 5 мы видели крайний случай: отношение двух независимых стандартных гауссовых величин имеет распределение Коши.

6.6.6. Возвращение к сечению

Для функции

$$\hat{\sigma}_{\text{proc}}(N, \varepsilon, L) = \frac{N}{\varepsilon L} \quad (6.64)$$

производные удобно записать через само сечение:

$$\frac{\partial \hat{\sigma}_{\text{proc}}}{\partial N} = \frac{\hat{\sigma}_{\text{proc}}}{N}, \quad \frac{\partial \hat{\sigma}_{\text{proc}}}{\partial \varepsilon} = -\frac{\hat{\sigma}_{\text{proc}}}{\varepsilon}, \quad \frac{\partial \hat{\sigma}_{\text{proc}}}{\partial L} = -\frac{\hat{\sigma}_{\text{proc}}}{L}. \quad (6.65)$$

Общая относительная дисперсия равна

$$\begin{aligned} \left(\frac{\sigma_{\hat{\sigma}_{\text{proc}}}}{\hat{\sigma}_{\text{proc}}}\right)^2 &\simeq \frac{\sigma_N^2}{N^2} + \frac{\sigma_\varepsilon^2}{\varepsilon^2} + \frac{\sigma_L^2}{L^2} \\ &\quad - 2\frac{\text{cov}(N, \varepsilon)}{N\varepsilon} - 2\frac{\text{cov}(N, L)}{NL} + 2\frac{\text{cov}(\varepsilon, L)}{\varepsilon L}. \end{aligned} \quad (6.66)$$

При нулевых ковариациях остаётся формула, использованная в начале главы. Знаки трёх ковариационных членов следуют из знаков производных, поэтому механическое сложение всех вкладов в квадратуре может дать неверный ответ.

6.7. Ковариационная матрица

Для случайного вектора

$$\mathbf{X} = (X_1, \dots, X_n)^T \quad (6.67)$$

определим матрицу

$$(\mathbf{V}_X)_{ij} = \text{cov}(X_i, X_j). \quad (6.68)$$

В двумерном случае

$$\mathbf{V}_X = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}. \quad (6.69)$$

Ковариационная матрица симметрична. На диагонали стоят дисперсии, а для любого вектора $\mathbf{a}$

$$\mathbf{a}^T \mathbf{V}_X \mathbf{a} = \text{Var}(\mathbf{a}^T \mathbf{X}) \geq 0. \quad (6.70)$$

Последнее равенство объясняет требование положительной полуопределённости: никакая линейная комбинация измерений не может иметь отрицательную дисперсию. Матрица с нарушенным условием не описывает совместное распределение случайных величин.

6.7.1. Матричная запись распространения ошибки

Соберём производные скалярной функции в градиент

$$\mathbf{g} = \nabla f|_{\mu} = \begin{pmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix}_{\mu}. \quad (6.71)$$

Уравнение 6.59 записывается одной строкой:

$$\sigma_F^2 \simeq \mathbf{g}^T \mathbf{V}_X \mathbf{g}. \quad (6.72)$$

Пусть выходных величин несколько,

$$\mathbf{Y} = \mathbf{f}(\mathbf{X}), \quad (6.73)$$

а матрица Якоби имеет элементы

$$J_{ai} = \left. \frac{\partial f_a}{\partial x_i} \right|_{\mu}. \quad (6.74)$$

Тогда

$$\mathbf{V}_Y \simeq \mathbf{J} \mathbf{V}_X \mathbf{J}^T. \quad (6.75)$$

Для линейного преобразования матрица $\mathbf{J}$ постоянна и уравнение 6.75 является точным. Для нелинейной функции это первый порядок разложения около μ .

6.7.2. Сумма и разность

Возьмём

$$\mathbf{X} = \begin{pmatrix} X \\ Y \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} S \\ D \end{pmatrix} = \begin{pmatrix} X + Y \\ X - Y \end{pmatrix}. \quad (6.76)$$

Матрица Якоби равна

$$\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (6.77)$$

Для входной матрицы

$$\mathbf{V}_X = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} \quad (6.78)$$

получаем

$$\mathbf{V}_U = \mathbf{J}\mathbf{V}_X\mathbf{J}^T = \begin{pmatrix} \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY} & \sigma_X^2 - \sigma_Y^2 \\ \sigma_X^2 - \sigma_Y^2 & \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY} \end{pmatrix}. \quad (6.79)$$

Диагональные элементы воспроизводят ошибки суммы и разности. Внедиагональный элемент даёт

$$\text{cov}(S, D) = \sigma_X^2 - \sigma_Y^2. \quad (6.80)$$

При $\sigma_X = \sigma_Y$ сумма и разность некоррелированы даже при $\sigma_{XY} \neq 0$. В следующей главе мы увидим геометрический смысл этого результата на двумерном гауссовом распределении.

Итоги главы

- При малой входной неопределённости функция заменяется касательной, и для одной переменной $\sigma_Y \simeq |f'(\mu_X)|\sigma_X$.
- Выборочная дисперсия с делителем $N - 1$ является несмещённой. Для гауссовой выборки её дисперсия равна $2\sigma^4/(N - 1)$.
- Корреляции входят в ошибку результата со знаками, заданными производными. Квадратурная сумма применима при нулевых ковариациях.
- Ковариационная матрица преобразуется по закону $\mathbf{V}_Y \simeq \mathbf{J}\mathbf{V}_X\mathbf{J}^T$; для линейной функции это равенство точное.

6.8. Задачи

Задача 1. Радиус и площадь круга

Измеренный радиус круга равен

$$R = (10.0 \pm 0.3) \text{ см.}$$

1. Найдите площадь $S = \pi R^2$ и её ошибку в линейном приближении.
2. Сравните относительные ошибки S и R .
3. Оцените поправку к математическому ожиданию площади, возникающую из-за квадратичности функции.

Задача 2. Сумма и разность коррелированных величин

Пусть

$$\sigma_X = 2, \quad \sigma_Y = 3, \quad \rho_{XY} = 0.8.$$

Вычислите ошибки:

$$S = X + Y, \quad D = X - Y.$$

Повторите вычисление для $\rho_{XY} = -0.8$ и объясните изменение результата.

Задача 3. Ошибка отношения

Эффективность определяется отношением

$$\varepsilon = \frac{N_{\text{selected}}}{N_{\text{total}}}.$$

1. Выведите линейную формулу для σ_ε , предполагая известные σ_{selected} , σ_{total} и корреляцию ρ .
2. Объясните, почему предположение о независимости числителя и знаменателя обычно неверно.
3. Сравните ответы для $\rho = 0$ и $\rho > 0$.

Задача 4. Преобразование ковариационной матрицы

Пусть

$$\mathbf{X} = \begin{pmatrix} X \\ Y \end{pmatrix}, \quad \mathbf{V}_X = \begin{pmatrix} 4 & 1.2 \\ 1.2 & 1 \end{pmatrix}.$$

Для новых величин

$$S = X + Y, \quad D = X - Y$$

найдите:

1. матрицу Якоби;
2. ковариационную матрицу $\mathbf{V}_{SD}$;
3. коэффициент корреляции между S и D .

Задача 5. Проверка линейного приближения

Пусть

$$X \sim N(1, \sigma_X^2), \quad Y = X + aX^2.$$

1. Получите линейную оценку σ_Y .
2. Выведите точные $E[Y]$ и $\text{Var}(Y)$.
3. Сгенерируйте не менее 10^5 значений X и сравните выборочные моменты Y с точным и линейным результатами.
4. На сетке $0 \leq a \leq 1$ и $0.05 \leq \sigma_X \leq 1$ найдите область, в которой линейная оценка стандартного отклонения отличается от точной менее чем на 5%.

Литература

[1] Joint Committee for Guides in Metrology. Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement. 2008. doi:10.59161/JCGM100-2008E.