

Когда гаусс не возникает

Содержание

5.1.	Физическая задача: ионизационные потери в трековом детекторе	61
5.2.	Что именно утверждает центральная предельная теорема	62
5.3.	Распределение Коши	62
5.4.	Энергетические потери в тонком слое вещества	63
5.5.	Коррелированные слагаемые	66
5.6.	Смеси распределений	67
5.7.	Отношение случайных величин	69
5.8.	Причины отклонения от гауссова приближения	70
5.9.	Задачи	71
	Литература	73

Аннотация

В прошлой главе мы увидели, как сумма независимых случайных вкладов приходит к распределению Гаусса. Теперь разберём случаи, в которых этот вывод применять нельзя или до гауссова предела приходится ждать слишком долго.

Начнём с распределений Коши и Ландау, у которых нет конечной дисперсии. Затем посмотрим, как корреляции меняют закон убывания ошибки, чем смесь распределений отличается от суммы и почему отношение двух гауссовых величин может иметь длинные хвосты.

5.1. Физическая задача: ионизационные потери в трековом детекторе

Заряженная частица проходит через трековый детектор и оставляет сигнал в n последовательных слоях. По измеренным зарядам

$$Q_1, Q_2, \dots, Q_n \quad (5.1)$$

нужно оценить её удельные потери энергии dE/dx . Первая идея состоит в том, чтобы взять среднее:

$$\bar{Q}_n = \frac{1}{n} \sum_{i=1}^n Q_i. \quad (5.2)$$

Если Q_i независимы и имеют конечную дисперсию σ_Q^2 , то

$$\text{Var}(\bar{Q}_n) = \frac{\sigma_Q^2}{n}. \quad (5.3)$$

Можно ожидать, что с ростом числа слоёв распределение среднего станет узким и гауссовым. Однако распределение заряда в одном тонком слое имеет длинный правый хвост. Иногда частица передаёт большую энергию одному электрону, и один Q_i оказывается намного больше остальных. Обычное среднее полностью сохраняет этот вклад.

В газовых трековых детекторах часто используют **усечённое среднее**. Заряды сначала располагают по возрастанию,

$$Q_{(1)} \leq Q_{(2)} \leq \dots \leq Q_{(n)}, \quad (5.4)$$

а затем усредняют только m меньших значений:

$$\bar{Q}_{\text{tr}} = \frac{1}{m} \sum_{i=1}^m Q_{(i)}, \quad m < n. \quad (5.5)$$

Так редкие большие потери не определяют оценку dE/dx . Выбор доли отбрасываемых измерений зависит от детектора и оптимизируется по его разрешению. Причину длинного хвоста мы получим ниже из спектра передачи энергии электрону [1].

Этот пример сразу ставит правильный вопрос. Число слагаемых само по себе ещё не определяет форму суммы. Нужно знать распределение отдельных вкладов и связи между ними.

5.2. Что именно утверждает центральная предельная теорема

Пусть

$$S_n = X_1 + X_2 + \dots + X_n. \quad (5.6)$$

Для независимых одинаково распределённых величин с параметрами

$$\mathbb{E}[X_i] = \mu, \quad 0 < \text{Var}(X_i) = \sigma^2 < \infty \quad (5.7)$$

центральная предельная теорема даёт

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{S_n/n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1). \quad (5.8)$$

В этой формулировке содержатся четыре условия, которые понадобятся проверять.

- Рассматривается сумма случайных величин.
- Слагаемые независимы и имеют одно распределение.
- Их дисперсия конечна.
- Переход к пределу выполняется после центрирования и нормировки суммы.

Для независимых величин с разными распределениями существуют более общие варианты теоремы. Условия Линдеберга и Ляпунова запрещают одному слагаемому сохранять конечную долю полной дисперсии. Для некоторых зависимых последовательностей также доказаны свои предельные теоремы [2]. Но из одной фразы «слагаемых много» ни одна из них не следует.

Общая формула для дисперсии суммы имеет вид

$$\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{cov}(X_i, X_j). \quad (5.9)$$

При независимых слагаемых ковариации равны нулю. Корреляции оставляют вторую сумму и могут полностью изменить зависимость ширины от n . Если дисперсия X_i не существует, нормировка через $\sqrt{n}$ лишается основания.

5.3. Распределение Коши

Стандартное распределение Коши имеет плотность

$$f_X(x) = \frac{1}{\pi(1+x^2)}. \quad (5.10)$$

Его кумулятивная функция распределения равна

$$F_X(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x. \quad (5.11)$$

Поэтому величину с распределением Коши можно получить из равномерной $U \sim U(0, 1)$ преобразованием

$$X = F_X^{-1}(U) = \tan\left[\pi\left(U - \frac{1}{2}\right)\right]. \quad (5.12)$$

Хвост плотности убывает как $1/x^2$. Из-за этого

$$E[|X|] = \infty, \quad E[X^2] = \infty. \quad (5.13)$$

Математическое ожидание распределения Коши не существует. Симметричное главное значение интеграла равно нулю, но математическим ожиданием оно не является.

5.3.1. Среднее Коши не сужается

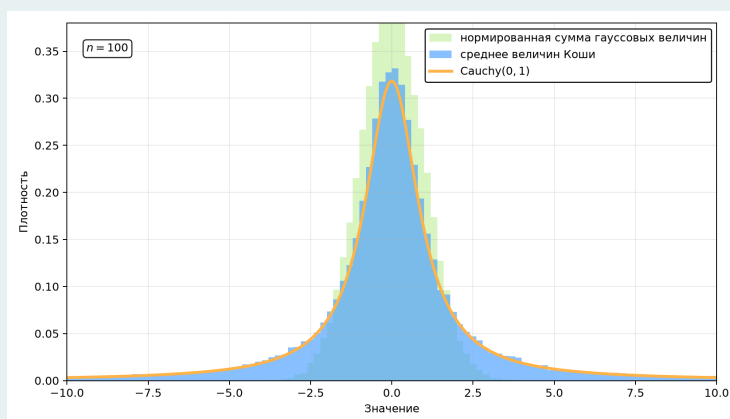
Пусть $X_1, \dots, X_n$ — независимые стандартные величины Коши. В главе 3 мы нашли характеристическую функцию их суммы:

$$\varphi_{S_n}(t) = e^{-n|t|}, \quad S_n = \sum_{i=1}^n X_i. \quad (5.14)$$

Это характеристическая функция распределения Коши с масштабом n . Деление суммы на n возвращает исходный масштаб:

$$\frac{S_n}{n} \sim \text{Cauchy}(0, 1). \quad (5.15)$$

Распределение среднего при $n = 1$ и $n = 200$ совпадает. Для гауссовых слагаемых ширина среднего уменьшилась бы в $\sqrt{200}$ раз.



Интерактив. Среднее распределения Коши

Меняя число слагаемых, можно сравнить среднее величин Коши с нормированной суммой гауссовых величин.



Распределение Коши встречается и в физике. Нерелятивистская линия Брейта—Вигнера имеет ту же форму и служит простой моделью резонансного пика. Для релятивистского резонанса форма зависит от выбранной кинематической переменной и фазового пространства.

Из поведения среднего не следует, что накопление данных для распределения Коши бесполезно. Параметр положения можно оценивать по медиане или методом максимального правдоподобия. Эти оценки используют форму распределения и не требуют существования $E[X]$.

5.4. Энергетические потери в тонком слое вещества

Вернёмся к задаче, с которой началась глава. Заряженная частица теряет энергию во множестве актов ионизации и возбуждения. Полную потерю энергии в слое

запишем как

$$\Delta = \omega_1 + \omega_2 + \dots + \omega_N, \quad (5.16)$$

где ω_a — энергия, переданная в одном столкновении. Формально перед нами сумма. Форма распределения определяется хвостом отдельных ω_a .

5.4.1. Редкие большие передачи энергии

Для быстрой тяжёлой заряженной частицы дифференциальное сечение передачи энергии электрону в существенной для хвоста области ведёт себя как

$$\frac{d\sigma}{d\omega} \propto \frac{1}{\omega^2}, \quad \omega \leq W_{\max}. \quad (5.17)$$

Здесь $W_{\max}$ — максимальная передаваемая энергия, заданная кинематикой столкновения. Вероятность передачи больше порога W пропорциональна

$$P(\omega > W) \propto \int_W^{W_{\max}} \frac{d\omega}{\omega^2} = \frac{1}{W} - \frac{1}{W_{\max}}. \quad (5.18)$$

При $W \ll W_{\max}$ получаем

$$P(\omega > W) \propto \frac{1}{W}. \quad (5.19)$$

Хвост убывает медленно. Один δ -электрон может унести заметную часть энергии, доступной в слое, и дать заряд намного больше наиболее вероятного. Именно эти события образуют правый хвост распределения Q_i в трековом детекторе.

5.4.2. Распределение Ландау

В идеализированном пределе тонкого слоя распределение потери энергии имеет ландау-подобную форму

$$p(\Delta) = \frac{1}{\xi} L\left(\frac{\Delta - \Delta_0}{\xi}\right). \quad (5.20)$$

Параметр Δ_0 задаёт положение распределения, а ξ — масштаб флуктуаций. При использовании здесь определения стандартной функции Ландау Δ_0 не совпадает с наиболее вероятным значением. Для частицы с зарядом ze в веществе

$$\xi = \frac{K}{2} \frac{Z}{A} \frac{z^2}{\beta^2} \rho x. \quad (5.21)$$

Здесь Z и A — атомный номер и молярная масса вещества, ρx — массовая толщина слоя, $\beta = v/c$, а $K = 4\pi N_A r_e^2 m_e c^2 \simeq 0.307 \text{ МэВ см}^2 \text{ моль}^{-1}$.

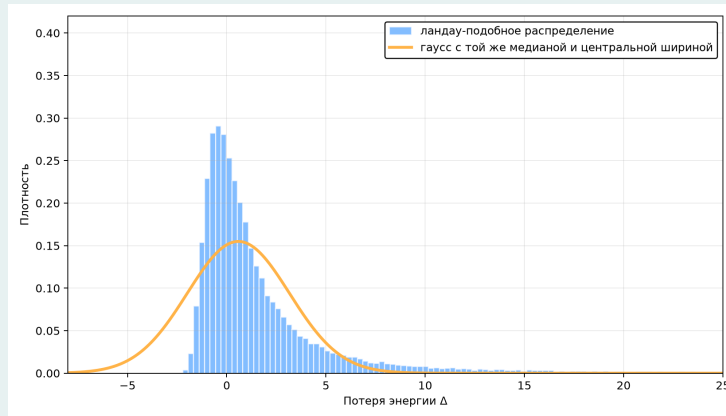
Стандартную функцию Ландау можно определить интегралом

$$L(\lambda) = \frac{1}{\pi} \int_0^\infty \exp[-u \ln u - \lambda u] \sin(\pi u) du. \quad (5.22)$$

При больших положительных λ

$$L(\lambda) \sim \frac{1}{\lambda^2}. \quad (5.23)$$

Наиболее вероятное значение конечно, а среднее и дисперсия идеального распределения Ландау не определены. Подробное описание флуктуаций ионизационных потерь в тонких кремниевых детекторах дано в [1].



Интерактив. Длинный хвост распределения Ландау

Интерактив сопоставляет ландау-подобное распределение с гауссовой кривой, имеющей ту же медиану и ту же центральную ширину.

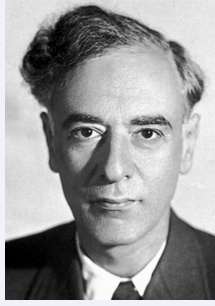


В идеализированной модели верхнюю границу передачи энергии удаляют. Тогда хвост $p(\omega) \propto 1/\omega^2$ даёт логарифмически расходящийся первый момент и линейно расходящийся второй:

$$\int_{-\infty}^{\infty} \omega \frac{d\omega}{\omega^2} = \int_{-\infty}^{\infty} \frac{d\omega}{\omega}, \quad \int_{-\infty}^{\infty} \omega^2 \frac{d\omega}{\omega^2} = \int_{-\infty}^{\infty} d\omega. \quad (5.24)$$

Обычная ЦПТ к такой модели неприменима. В реальном столкновении $W_{\max}$ конечно. Распределение Ландау, столь любимое экспериментаторами для описания потерь энергии в тонких детекторах — это приближение. Если хотите точнее, используйте распределение Вавилова. При увеличении толщины слоя последнее переходит к гауссову распределению.

Теперь понятна и конструкция из начала главы. Усечённое среднее подавляет редкие большие Q_i , созданные δ -электронами. Цена этой операции — потеря части измерений и необходимость отдельно определить распределение и смещение оценки $\bar{Q}_{\text{tr}}$.



БИОГРАФИЯ

Лев Давидович Ландау

1908–1968

Лев Ландау - выдающийся советский физик-теоретик. Он разработал квантовую теорию диамагнетизма, теорию фазовых переходов второго рода, теорию сверхтекучести жидкого гелия и теорию ферми-жидкости. Вместе с Виталием Гинзбургом он построил феноменологическую теорию сверхпроводимости. За работы по теории конденсированного состояния, прежде всего жидкого гелия, Ландау получил Нобелевскую премию по физике 1962 года.

Для этой главы особенно важна его теория флуктуаций энергетических потерь быстрой заряженной частицы в тонком слое вещества. Полученное распределение сохраняет след редких больших передач энергии и имеет длинный правый хвост.

Лев Ландау и Евгений Лифшиц создали многотомный «Курс теоретической физики», воспитавший множество поколений учёных в нашей стране и во всём мире.

Источник: Nobel Prize; фото архив Нобелевского фонда, общественное достояние

5.5. Коррелированные слагаемые

Ковариационная сумма в уравнении 5.9 исчезает только для некоррелированных слагаемых. Рассмотрим простой случай:

$$\text{Var}(X_i) = \sigma^2, \quad \text{corr}(X_i, X_j) = \rho, \quad i \neq j. \quad (5.25)$$

Для среднего

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (5.26)$$

получаем

$$\begin{aligned} \text{Var}(\bar{X}_n) &= \frac{1}{n^2} \left[n\sigma^2 + 2 \frac{n(n-1)}{2} \rho\sigma^2 \right] \\ &= \frac{\sigma^2}{n} [1 + (n-1)\rho]. \end{aligned} \quad (5.27)$$

При $\rho = 0$ остаётся σ^2/n . При фиксированном $\rho > 0$

$$\text{Var}(\bar{X}_n) \xrightarrow{n \rightarrow \infty} \rho\sigma^2. \quad (5.28)$$

Матрица с одинаковой корреляцией всех пар допустима при $-1/(n-1) \leq \rho \leq 1$. Поэтому фиксированное отрицательное ρ нельзя сохранять при неограниченном росте n . Положительная корреляция оставляет конечную ширину среднего.

5.5.1. Общий сдвиг как систематическая ошибка

Пусть каждое измерение содержит одну общую и одну независимую составляющую:

$$X_i = \mu + C + \varepsilon_i. \quad (5.29)$$

В ансамбле повторных калибровок и измерений зададим

$$C \sim \mathcal{N}(0, \sigma_C^2), \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad (5.30)$$

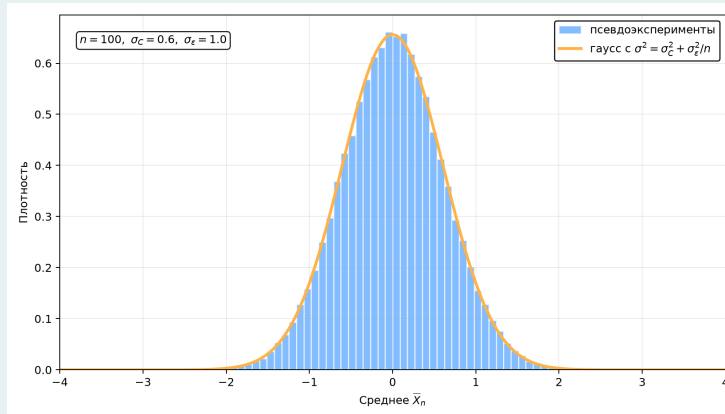
причём C и все ε_i независимы. Тогда

$$\bar{X}_n = \mu + C + \frac{1}{n} \sum_{i=1}^n \varepsilon_i \quad (5.31)$$

и

$$\text{Var}(\bar{X}_n) = \sigma_C^2 + \frac{\sigma_\varepsilon^2}{n}. \quad (5.32)$$

Статистический член уменьшается с числом измерений. Общий сдвиг повторяется во всех X_i и остаётся целиком. Так ведут себя, например, неопределённость энергетической шкалы, нормировки потока или светимости, коэффициента усиления и общий временной сдвиг.



Интерактив. Статистическая и общая ошибки

Меняйте статистическую и общую составляющие ошибки и следите за шириной распределения среднего при росте выборки.



В этой модели $\bar{X}_n$ имеет гауссово распределение при любом n . Однако его ширина стремится к σ_C и не обращается в нуль. Гауссова форма гистограммы сама по себе не подтверждает закон $1/\sqrt{n}$.

5.6. Смеси распределений

Сумма и смесь описывают разные устройства эксперимента. Для суммы два вклада присутствуют в каждом событии. Например, если

$$N_1 \sim \text{Pois}(1000), \quad N_2 \sim \text{Pois}(100) \quad (5.33)$$

независимы, то

$$N_1 + N_2 \sim \text{Pois}(1100). \quad (5.34)$$

В смеси каждое событие принадлежит одному из нескольких классов. Введём скрытую переменную $K \in \{1, 2\}$ и зададим

$$N | K = 1 \sim \text{Pois}(1000), \quad N | K = 2 \sim \text{Pois}(100). \quad (5.35)$$

Если класс события не зарегистрирован, в данных остаётся только N . Даже узкие распределения внутри классов образуют широкую или многопиковую смесь.

5.6.1. Среднее и дисперсия смеси

Пусть X при фиксированном классе K имеет среднее μ_K и дисперсию σ_K^2 . Полное среднее равно

$$\mathbb{E}[X] = \mathbb{E}_K[\mu_K]. \quad (5.36)$$

Формула полной дисперсии разделяет два источника разброса:

$$\text{Var}(X) = \mathbb{E}_K[\sigma_K^2] + \text{Var}_K(\mu_K). \quad (5.37)$$

Первое слагаемое усредняет дисперсии внутри классов. Второе измеряет разброс между их средними.

Для двух равновероятных пуассоновских классов из уравнения 5.35

$$\mathbb{E}[N] = \frac{1000 + 100}{2} = 550. \quad (5.38)$$

Средняя дисперсия внутри классов равна

$$\mathbb{E}_K[\text{Var}(N | K)] = \frac{1000 + 100}{2} = 550, \quad (5.39)$$

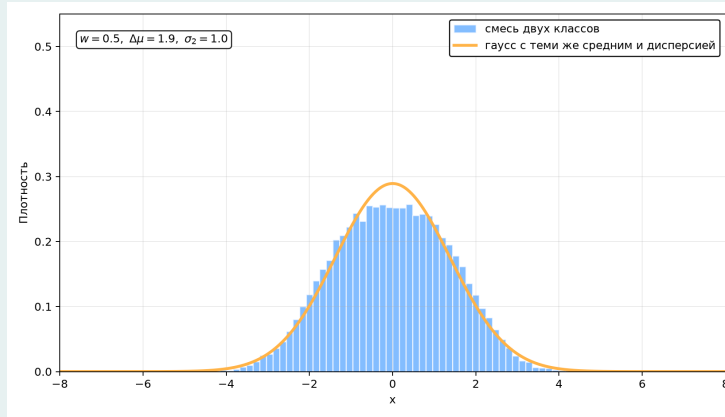
а дисперсия условных средних составляет

$$\text{Var}_K(\mathbb{E}[N | K]) = \frac{(1000 - 550)^2 + (100 - 550)^2}{2} = 202\,500. \quad (5.40)$$

Полная дисперсия

$$\text{Var}(N) = 550 + 202\,500 = 203\,050 \quad (5.41)$$

почти целиком определяется различием двух классов. Пуассоновские флуктуации внутри классов дают меньше трёх десятых процента результата.



Интерактив. Смесь двух гауссовых классов

Интерактив показывает, как веса, ширины и расстояние между средними меняют форму смеси.



5.6.2. Пример: жидкий сцинтиллятор

Сигнал жидкосцинтилляционного детектора можно схематично записать как

$$Q = Q(E, K) + Q_{\text{stat}} + Q_{\text{el}}, \quad (5.42)$$

где Q_{stat} описывает фотонную статистику, Q_{el} — шум электроники, а класс K включает тип частицы, топологию события, положение в детекторе, долю черенковского света и другие детали отклика.

Для γ -кванта энергия передаётся нескольким вторичным электронам:

$$\gamma \longrightarrow e_1 + e_2 + \dots + e_m, \quad E_\gamma = \sum_{a=1}^m E_a. \quad (5.43)$$

Если световой выход электрона $L(E)$ нелинеен, средний заряд имеет вид

$$Q_\gamma \simeq \sum_{a=1}^m L(E_a). \quad (5.44)$$

При одном и том же E_γ разные наборы E_a дают разные значения Q_γ . Разброс топологий добавляет к фотонной статистике второй член из формулы полной дисперсии 5.37. Описание сигнала одним пуассоновским или гауссовым распределением этот вклад потеряет.

5.7. Отношение случайных величин

Центральная предельная теорема относится к суммам. Нелинейное преобразование может изменить распределение даже при точно гауссовых исходных величинах.

Пусть X и Y независимы и

$$X, Y \sim \mathcal{N}(0, 1). \quad (5.45)$$

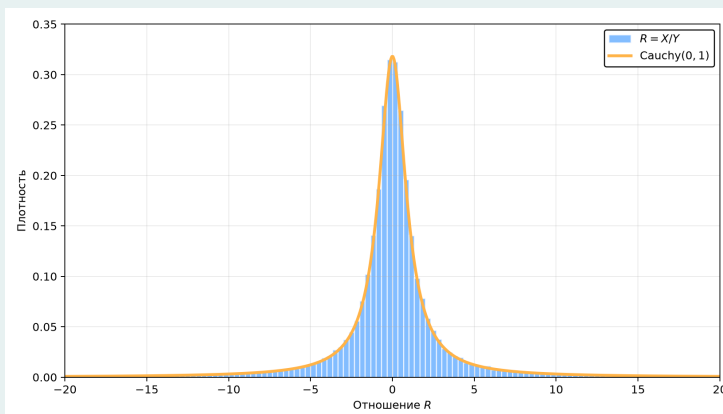
Для отношения $R = X/Y$ совместная плотность после замены $x = ry$ даёт

$$\begin{aligned}
 f_R(r) &= \int_{-\infty}^{+\infty} |y| f_X(ry) f_Y(y) dy \\
 &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} |y| \exp\left[-\frac{(1+r^2)y^2}{2}\right] dy \\
 &= \frac{1}{\pi(1+r^2)}.
 \end{aligned}
 \tag{5.46}$$

Получилось стандартное распределение Коши. Отношения появляются при нормировке сигнала на контрольную область, введении поправки на эффективность, сравнении двух скоростей счёта и построении асимметрии

$$A = \frac{N_1 - N_2}{N_1 + N_2}. \tag{5.47}$$

Линейное распространение ошибок для отношения применимо, когда знаменатель с высокой вероятностью далёк от нуля и его относительная неопределённость мала. Если знаменатель может оказаться около нуля, редкие значения отношения уходят далеко в хвост. Симметричная запись $r \pm \sigma_r$ перестаёт описывать такое распределение.



Интерактив. Отношение гауссовых величин

Меняйте среднее и ширину знаменателя и следите за появлением длинных хвостов у отношения.



5.8. Причины отклонения от гауссова приближения

В таблице 5.1 собраны рассмотренные причины.

Таблица 5.1. Причины отклонения от простого гауссова приближения

Механизм	Какое условие требует проверки	Что видно в данных
Бесконечная дисперсия	Существует ли $\text{Var}(X_i)$	Среднее может не сужаться
Тяжёлый хвост	Насколько быстро наступает предельный режим	Редкие большие значения сохраняются при конечном n

Механизм	Какое условие требует проверки	Что видно в данных
Доминирующий вклад	Мал ли каждый вклад по сравнению со всей суммой	Форму хвоста задаёт один механизм
Корреляции	Независимы ли слагаемые	Ширина не обязана убывать как $1/\sqrt{n}$
Смесь классов	Однородна ли выборка	Возникают дополнительная ширина, асимметрия или несколько пиков
Нелинейное преобразование	Является ли величина суммой	Отношение или асимметрия наследует границы и хвосты знаменателя

Проверять гауссово приближение следует на той случайной величине, которая действительно входит в анализ. Для этого строят псевдоэксперименты из полной модели и сравнивают распределения при нескольких размерах выборки. Центральная часть гистограммы может выглядеть гауссовой раньше хвостов, а именно хвосты определяют вероятности редких отклонений.

Итоги главы

- Конечная дисперсия и отсутствие доминирующего слагаемого входят в условия гауссова предела. У среднего величин Коши ширина не уменьшается с ростом выборки.
- Длинный правый хвост ионизационных потерь создают редкие δ -электроны. В тонком слое используется ландау-подобное описание; кинематическая граница передачи энергии приводит к более точному распределению Вавилова.
- Положительные корреляции оставляют конечный вклад в дисперсию среднего. Общая систематическая ошибка может сохранить гауссову форму и одновременно нарушить закон $1/\sqrt{n}$.
- Смесь классов получает дополнительную дисперсию от различия условных средних. Этот вклад описывается формулой полной дисперсии.
- Отношение двух независимых стандартных гауссовых величин имеет распределение Коши. Для нелинейной величины распределение нужно выводить отдельно.

5.9. Задачи

Задача 1. Среднее распределения Коши

Для независимых $X_i \sim \text{Cauchy}(0, 1)$ покажите, что дисперсия не существует, а

$$\bar{X}_n = \frac{1}{n} \sum_i X_i$$

имеет то же распределение Коши. Почему это не противоречит ЦПТ?

Задача 2. Тяжёлый хвост с конечной дисперсией

Пусть

$$p(x) \sim \frac{C}{|x|^{1+\alpha}}.$$

1. При каких α существуют среднее и дисперсия?
2. Почему при α немного больше 2 сходимость к гауссу может быть медленной?

Задача 3. Один доминирующий вклад

Пусть

$$S = X_0 + \epsilon_1 + \dots + \epsilon_n,$$

где $\epsilon_i \sim N(0, \sigma^2)$, а X_0 негауссов и

$$\text{Var}(X_0) \gg n\sigma^2.$$

Объясните форму распределения S .

Задача 4. Общая систематическая ошибка

Пусть

$$X_i = \mu + C + \epsilon_i,$$

где $C \sim N(0, \sigma_C^2)$ и $\epsilon_i \sim N(0, \sigma_\epsilon^2)$. Найдите $E[\bar{X}_n]$, $\text{Var}(\bar{X}_n)$ и предел дисперсии при $n \rightarrow \infty$.

Задача 5. Усреднение коррелированных измерений

Пусть n величин имеют одинаковую дисперсию σ^2 и одинаковый коэффициент корреляции ρ для любой пары.

1. Получите дисперсию их среднего.
2. Укажите допустимый диапазон ρ при фиксированном n .
3. Рассмотрите случаи $\rho = 0$ и $\rho > 0$ при увеличении n .

Задача 6. Смесь двух гауссов

Пусть

$$X|A \sim N(\mu_A, \sigma_A^2), \quad X|B \sim N(\mu_B, \sigma_B^2),$$

а вероятность класса A равна w . Найдите полное среднее и дисперсию. Когда смесь плохо похожа на один гаусс?

Задача 7. Скрытая топология в жидком сцинтилляторе

Моноэнергетический γ -квант с энергией

$$E_\gamma = 1 \text{ МэВ}$$

может передать энергию вторичным электронам двумя способами:

$$\theta = A : \quad T_1 = 1 \text{ МэВ},$$

$$\theta = B : \quad T_1 = T_2 = 0.5 \text{ МэВ}.$$

Из-за нелинейности светового выхода средние сигналы различаются:

$$Q|\theta = A \sim \text{Pois}(1000), \quad Q|\theta = B \sim \text{Pois}(980).$$

Считайте классы равновероятными.

1. Найдите полное среднее $E[Q]$.
2. Вычислите среднюю внутриклассовую дисперсию

$$E_\theta[\text{Var}(Q|\theta)].$$

3. Вычислите топологический вклад

$$\text{Var}_\theta(E[Q|\theta]).$$

4. Найдите полную дисперсию и стандартное отклонение Q .
5. Сравните результат с пуассоновской моделью

$$Q \sim \text{Pois}(E[Q]).$$

Какой вклад не описывается этой моделью?

Задача 8. Отношение гауссовых величин

Для независимых $X, Y \sim N(0, 1)$ выведите плотность отношения $R = X/Y$ и покажите, что это распределение Коши.

Литература

- [1] Bichsel, H. Straggling in Thin Silicon Detectors. *Reviews of Modern Physics*, **60**, 663–699. 1988. doi:10.1103/RevModPhys.60.663.
- [2] Feller, W. *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons. 1971.