4. Определение эффективных оценок с помощью неравенства Рао-Крамера-Фреше. Нижняя граница дисперсии оценки генеральных средней и дисперсии нормально распределенной генеральной совокупности.

Неравенство Рао-Крамера-Фреше

Пусть $\varphi(x,\theta)$ — плотность вероятностей случайной величины X, если X — непрерывная случайная величина, и $\varphi(x_i,\theta) = P(X = x_i, \theta)$ — вероятность того, что с.в. X приняла значение x_i, если X - дискретная случайная величина.

Если $\varphi(x,\theta)$ удовлетворяет условиям регулярности:

1. область возможных значений исследуемой случайной величины, в которой $\varphi(x,\theta) \neq 0$, не зависит от θ;
2. в выражении $M(\theta) = \int \theta \cdot L(x_1,\dots,x_n,\theta)dx_1dx_2\dots dx_n$ и тождестве $\int L(x_1,\dots,x_n,\theta)dx_1dx_2\dots dx_n = 1$ допустимо дифференцирование по θ под знаком интеграла;
3. величина $I(\theta,x) \neq 0$,

пояснение от ИИ (в конспектах не упоминается): L(x_1, ..., x_n, θ) — функция правдоподобия (совместная вероятность (или плотность) получить именно нашу выборку при данном θ)
Для независимых наблюдений: $\boxed{L(x_1,...,x_n,\theta) = \varphi(x_1,\theta)\cdot \varphi(x_2,\theta)\cdot ...\cdot \varphi(x_n,\theta)}$

тогда для любой оценки $\hat{\theta}$ неизвестного параметра θ имеет место неравенство Рао-Крамера-Фреше:

$$D(\hat{\theta}) \geq \frac{1}{n\cdot I(\theta,x)}, \quad (3.5)$$

где $D(\hat{\theta})$ - дисперсия оценки $\hat{\theta}$ параметра θ; $I(\theta,x)$ - количество информации Фишера о параметре θ, содержащееся в единичном наблюдении:

$$I(\theta,x) = M\left[(\ln\varphi(x,\theta))'_\theta\right]^2. \quad (3.6)$$

По правилу дифференцирования сложной функции

$$(\ln y(x))'_x = \frac{y'(x)}{y(x)}.$$

Поэтому в случае дискретной случайной величины

$$I(\theta,x) = \sum_{i=1}^{n}\left[\frac{\varphi'(x_i,\theta)}{\varphi(x_i,\theta)}\right]^2 \cdot \varphi(x_i,\theta).$$

В случае непрерывной случайной величины

$$I(\theta,x) = \int_{-\infty}^{\infty}\left[\frac{\varphi'(x,\theta)}{\varphi(x,\theta)}\right]^2 \cdot \varphi(x,\theta)dx.$$

Неравенство (5.11) позволяет найти тот минимум дисперсии, который должна иметь дисперсия оценки $D(\hat{\theta})$, чтобы быть эффективной оценкой, т.е.

$$D_{\text{эф}}(\hat{\theta}) = \min D(\hat{\theta}) = \frac{1}{n\cdot I(\theta,x)}.$$

Нижняя граница дисперсии оценки генеральных средней и дисперсии нормально распределенной генеральной совокупности.

- Минимально возможная дисперсия оценки генеральной средней (математического ожидания)

$$\min D(a) = \frac{1}{nI(a)} = \frac{\sigma^2}{n}.$$

- Для генеральной дисперсии

$$\min D(\sigma^2) = \frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n}.$$

Распишем вывод каждой формулы

Пример Найти нижнюю границу дисперсии оценки генеральной средней a и генеральной дисперсии σ^2 повторной выборки нормально распределенной генеральной совокупности.

Решение: В случае нормального закона распределения плотность вероятности

$$\varphi_N(x, a, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Тогда

$$\ln \varphi_N(x, a, \sigma^2) = -\ln \sqrt{2\pi\sigma^2} - \frac{(x-a)^2}{2\sigma^2},$$

и

$$\left[\ln \varphi_N(x, a, \sigma^2)\right]_a' = \left(0 + \frac{x-a}{\sigma^2}\right)^2 = \frac{(x-a)^2}{\sigma^4}.$$

$\partial/\partial a[ln\varphi] = (x-a)/\sigma^2$

*Это и есть **функция вклада (score)** для параметра a. Видно, что ее значение линейно зависит от отклонения наблюдения x от оцениваемого среднего a. Чем дальше точка от a, тем больше вклад.*

Тогда количество информации Фишера

$$I(a) = M\left[\frac{(x-a)^2}{\sigma^4}\right] = \frac{M[(x-a)^2]}{\sigma^4} = \frac{D(X)}{\sigma^4} = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}.$$

Минимально возможная дисперсия оценки генеральной средней (математического ожидания)

$$\boxed{\min D(a) = \frac{1}{nI(a)} = \frac{\sigma^2}{n}.}$$

Для определения минимального разброса оценки дисперсии вычислим производную от $\ln \varphi(x, a, \sigma^2)$ по переменной $\sigma^2 = K$:

$$\left[(\ln \varphi_N(x, a, \sigma^2))_{\sigma^2}'\right]^2 = \left[\left(-\ln \sqrt{2\pi\sigma^2} - \frac{(x-a)^2}{2\sigma^2}\right)_{\sigma^2}'\right]^2 =$$

$$= \left[-\frac{1}{2\sigma^2} + \frac{(x-a)^2}{2\sigma^4}\right]^2 = \frac{1}{4\sigma^4} - \frac{2(x-a)^2}{4\sigma^6} + \frac{(x-a)^4}{4\sigma^8}.$$

Вычислим количество информации Фишера

$$I(\sigma^2) = M\left[(\ln \varphi_N(x, a, \sigma^2))_{\sigma^2}'\right]^2 = M\left[\frac{1}{4\sigma^4}\right] - M\left[\frac{(x-a)^2}{2\sigma^6}\right] + M\left[\frac{(x-a)^4}{4\sigma^8}\right] =$$

$$= \frac{1}{4\sigma^4} - \frac{M(x-a)^2}{2\sigma^6} + \frac{M(x-a)^4}{4\sigma^8} = \frac{1}{4\sigma^4} - \frac{\sigma^2}{2\sigma^6} + \frac{3\sigma^4}{4\sigma^8} = \frac{1}{2\sigma^4}.$$

Здесь в последней строке учтено, что для нормального распределения $M(x-a)^4 = \mu_4 = 3\sigma^4$ (см. Теорию вероятностей).

В результате получим, что эффективная оценка генеральной дисперсии σ^2 повторной выборки для нормально распределенной генеральной совокупности должна иметь минимальную дисперсию

$$\boxed{\min D(\sigma^2) = \frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n}.}$$

Алгоритм нахождения нижней границы дисперсии оценки (на примере нормального распределения)

1. **Записать плотность распределения** $\varphi(x, \theta)$.
2. **Взять натуральный логарифм** от плотности: $\ln \varphi(x, \theta)$.
3. **Найти частную производную** логарифма по интересующему параметру θ.
4. **Возвести эту производную в квадрат** (готовимся вычислить математическое ожидание квадрата).
5. **Вычислить количество информации Фишера** $I(\theta, x)$ как **математическое ожидание** от полученного квадрата производной.
6. **Подставить** в формулу нижней границы: $\min D(\hat{\theta}) = \frac{1}{nI(\theta,x)}$.

Пояснения

1. Смысл неравенства Рао-Крамера-Фреше (РКФ)

Неравенство РКФ отвечает на фундаментальный вопрос: **Какова наилучшая возможная точность оценки неизвестного параметра?** Оно устанавливает **теоретический предел**, ниже которого дисперсия любой несмещенной оценки опуститься не может.

Аналогия: Представьте, что вы пытаетесь измерить длину стола шариковой ручкой. Ваши измерения будут колебаться вокруг истинного значения. Неравенство РКФ говорит: "Какой бы метод измерения (оценки) вы ни придумали, разброс ваших результатов (дисперсия) не может быть меньше, чем $1/(n * I)$". Это как "предел Шеннона" в теории информации, но для статистических оценок.

2. Пояснение компонент формулы РКФ

Формула:
$D(\hat{\theta}) \geq 1/[n * I(\theta, x)]$

- $I(\theta, x)$ — **Информация Фишера (в одном наблюдении).** Это **ключевое понятие**.
  - **Смысл:** Количество информации, которое одно наблюдение x несет об неизвестном параметре θ. Чем больше $I(\theta, x)$, тем "чувствительнее" распределение к изменению параметра θ, тем легче его оценить по данным.
  - **Как вычисляется:** $I(\theta, x) = M[\frac{\delta}{\delta_\theta} ln\varphi(x, \theta))^2]$.
    - $ln\varphi(x, \theta)$ — Логарифмическая функция правдоподобия для одного наблюдения. С ней удобнее работать математически.
    - $(\partial/\partial\theta ln\varphi(x, \theta))$ — **Вклад одного наблюдения (score function).** Показывает, насколько резко меняется "правдоподобие" данных при малом изменении параметра θ. Если правдоподобие сильно меняется (производная большая), значит, данные дают четкий сигнал о значении параметра.
    - **Квадрат и математическое ожидание** $M[...]$**:** Мы усредняем квадрат этого "вклада" по всем возможным значениям x. Это дает нам **среднюю "силу сигнала"** об параметре, содержащуюся в одном наблюдении. Большая дисперсия score function означает большую информацию.
- $1/[n * I(\theta, x)]$ — **Нижняя граница Крамера-Рао (НГКР).** Это и есть тот самый **минимально достижимый уровень "шума"** (дисперсии) для любой несмещенной оценки. Если вы нашли оценку, дисперсия которой в точности равна НГКР ($D(\hat{\theta}) = 1/(nI)$), то вы нашли **эффективную оценку** — лучшую из всех возможных несмещенных оценок.

Итог: Неравенство Рао-Крамера — это не просто формула, а мощный теоретический инструмент. Оно задает **планку** для точности, позволяет **сравнивать** оценки, **планировать** объемы данных и глубже **понимать**, от чего зависит сложность оценки того или иного параметра распределения.