**Финансовый университет
при правительстве Российской Федерации**

Шамраева Виктория Викторовна

кандидат физико-математических наук,
доцент кафедры
математики и анализа данных

Теория вероятностей и математическая статистика

НАПРАВЛЕНИЕ ПОДГОТОВКИ: «Прикладная математика - ПМ»

КВАЛИФИКАЦИЯ (СТЕПЕНЬ): бакалавр

Раздел 1. Оценки параметров

Точечные статистические оценки

Раздел 1. Оценки параметров

Статистикой называется любая функция

$$T = T(X_1, X_2, \dots, X_n)$$

от элементов выборки $X_1, X_2, \dots, X_n$

[на конкретной выборке $x_1, x_2, \dots, x_n$ эта статистика принимает конкретное значение $T = T(x_1, x_2, \dots, x_n)$].

Любая статистика является с.в., поскольку она является функцией случайных величин.

Раздел 1. Оценки параметров

Предположим, что генеральное распределение признака X зависит от некоторого параметра $\theta \in \Theta$.

Оценкой $\hat{\theta}$ числовой характеристики или параметра θ называется любая статистика

$$\hat{\theta} = \hat{\theta}(x_1, x_2, \ldots, x_n),$$

используемая в качестве приближенного значения θ.

Иногда, чтобы отметить величину объема выборки, по которой была получена оценка, вместо обозначения $\hat{\theta}$ будем использовать обозначение $\hat{\theta}_n$.

Раздел 1. Оценки параметров

Например, в качестве оценки параметра $a = \boldsymbol{E}(X)$ с.в. X, распределенной по нормальному закону (с неизвестными параметрами), можно использовать:

- результат единичного наблюдения $\hat{a} = X_1;$
- полусумму минимального и максимального элементов выборки
$$\hat{a} = \frac{X_{(\min)} + X_{(\max)}}{2};$$
- выборочную медиану $\hat{a} = \hat{X}_{\mathrm{med}};$
- выборочную моду $\hat{a} = \hat{X}_{\mathrm{mod}};$
- выборочное среднее $\hat{a} = \overline{X}$ и др.

Раздел 1. Оценки параметров

Точечной оценкой неизвестного параметра θ называют число (точку на числовой оси), которое приблизительно равно оцениваемому параметру и может заменить его с **достаточной степенью точности** в статистических расчетах.

Большинство с.в. имеют распределения, зависящие от одного или нескольких параметров.

Так, например, **нормальное распределение** зависит от параметров a и σ^2, **распределение Пуассона** – от параметра λ, **биномиальное** – от параметра p.

Раздел 1. Оценки параметров

Закон распределения $\widehat{\theta}_n$ зависит от **закона распределения случайной величины** $\boldsymbol{X}$ (и, в частности, от самого неизвестного параметра θ).

Кроме того, $\widehat{\theta}_n$ зависит от **числа опытов** $\boldsymbol{n}$, поэтому имеет смысл говорить о последовательности вида

$$\{\widehat{\theta}_n;\ n = 1, 2, \dots\}.$$

Раздел 1. Оценки параметров

Для того чтобы точечные статистические оценки обеспечивали хорошие приближения неизвестных параметров, они должны быть

несмещенными,

состоятельными и

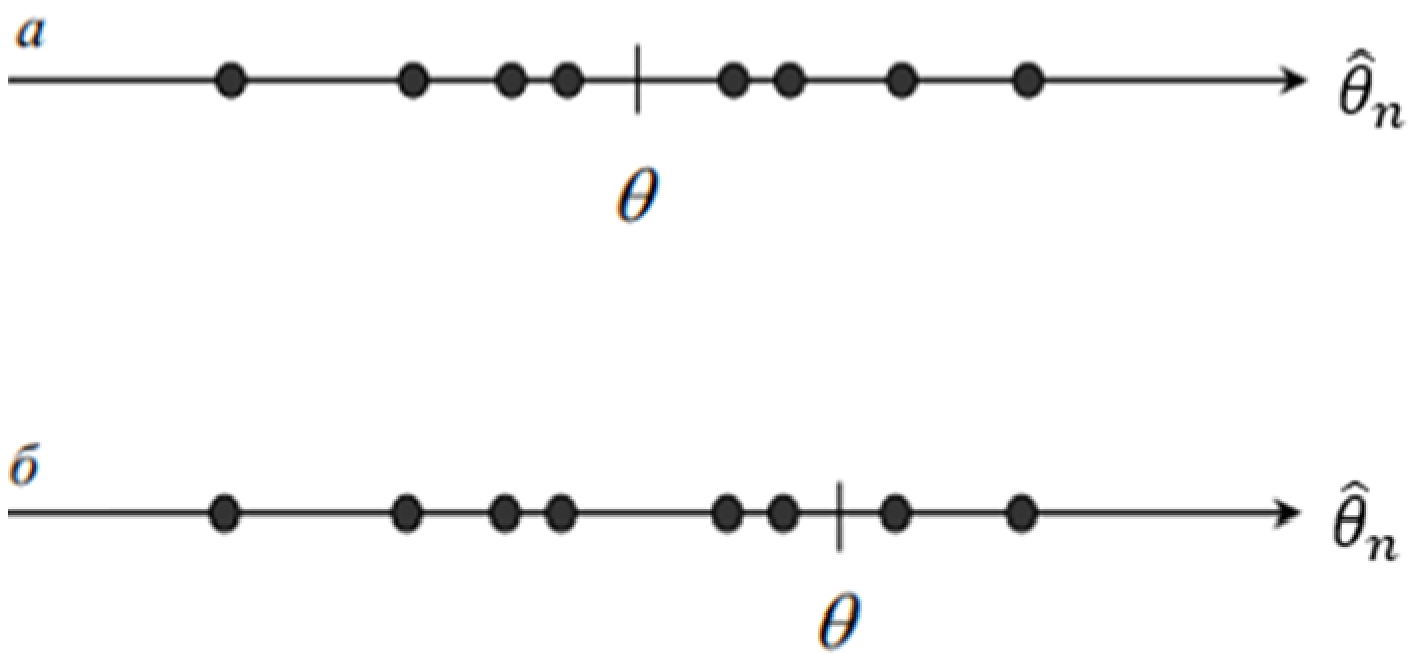
эффективными.

Раздел 1. Оценки параметров

Оценка $\hat{\theta}$ называется **несмещённой оценкой параметра** θ, если она не содержит систематических ошибок, то есть при любом фиксированном объёме выборки n результат осреднения оценки по всем возможным выборкам данного объёма должен приводить к точному значению параметра:

$$E(\hat{\theta}) = \theta, \forall\, \theta \in \Theta.$$

Раздел 1. Оценки параметров



Раздел 1. Оценки параметров

Пример. Проверить, является ли выборочное среднее $\bar{X}$ несмещенной оценкой для неизвестного математического ожидания случайной величины X.

Решение.

Раздел 1. Оценки параметров

Пример. Пусть X_1, X_2, …, X_n— выборка из распределения с плотностью

$$f(x) = \begin{cases} 7e^{7(\theta - x)}, \text{при } x \geq \theta, \\ 0, \text{при } x < \theta. \end{cases}$$

Проверить, является ли оценка , $\hat{\theta} = \bar{X} - \frac{1}{7}$ несмещенной оценкой параметра θ?

Решение.

Раздел 1. Оценки параметров

Пример. Монету подбрасывают n раз. Вероятность выпадения герба при каждом подбрасывании равна p. В ходе опыта монета выпала гербом n_1 раз. Показать несмещенность оценки $\hat{\theta}_n = \frac{n_1}{n}$ вероятности $\theta=p$ выпадения герба в каждом опыте.

Решение.

Раздел 1. Оценки параметров

Пример. Проверить, является ли эмпирический второй начальный момент $\hat{\nu}_2$ несмещенной оценкой для неизвестного теоретического начального момента второго порядка случайной величины ν_2 .

Решение.

Раздел 1. Оценки параметров

Пример. Проверить, является ли эмпирическая дисперсия несмещенной оценкой для неизвестной дисперсии σ_X^2 случайной величины.

Решение.

Раздел 1. Оценки параметров

Разность $E(\hat{\theta}) - \theta$ называется **смещением**.

В примере с дисперсией смещение:

$$\frac{n-1}{n}\sigma_X^2 - \sigma_X^2 = -\frac{\sigma_X^2}{n}\,.$$

То есть при многократном повторении эксперимента эмпирическая дисперсия не будет стремиться к своему истинному значению (как эмпирическая средняя), а будет смещена относительно истинного значения по генеральной совокупности.

Видно, что в случае с дисперсией с увеличением объема выборки **величина смещения стремится к 0**.

Раздел 1. Оценки параметров

Оценка $\hat{\theta}_n$ называется **асимптотически несмещенной оценкой** параметра θ, если $E\left(\hat{\theta}_n\right) \longrightarrow \theta$ при $n \longrightarrow \infty$.

Таким образом,

выборочная дисперсия является асимптотически несмещенной оценкой дисперсии.

Раздел 1. Оценки параметров

При конечном n из оценки $\hat{\sigma}^2$ несложно получить несмещенную оценку дисперсии.

Для этого нужно величину $\hat{\sigma}^2$ умножить на коэффициент $\frac{n}{n-1}$.

Исправленной выборочной дисперсией называется величина

$$s_X^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} = \frac{n}{n-1} \hat{\sigma}_X^2 .$$

Раздел 1. Оценки параметров

Пример. Найти несмещенную оценку дисперсии с.в. на основании данного распределения выборки:

x_i	2	7	9	10
n_i	8	14	10	18

Решение.

Раздел 1. Оценки параметров

Выборочное среднее квадратическое отклонение — это

$$\hat{\sigma}_X = \sqrt{\hat{\sigma}_X^2} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}},$$

а **стандартное отклонение** — это

$$s_X = \sqrt{s_X^2} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{n}{n-1}\hat{\sigma}_X^2}.$$

При переходе к среднеквадратичном отклонению по выборке (корень из выборочной дисперсии) разница становится между ними ещё меньше.

Таким образом, эффект **смещенной дисперсии** проявляется в небольших выборках. В больших выборках можно использовать **генеральную дисперсию**, что как бы не усложняет и не упрощает жизнь.

Раздел 1. Оценки параметров

Оказывается, ≪исправить≫ имеющиеся смещенные оценки возможно не всегда.

Существуют примеры параметров, для которых несмещенные оценки построить невозможно.

Раздел 1. Оценки параметров

Пример (**когда не существует несмещённой оценки**). Производится одно наблюдение X, где X – случайная величина, имеющая пуассоновское распределение с неизвестным параметром θ, то есть $X \sim Pois(\theta)$ и

$$P_\theta(X = k) = \frac{\theta^k}{k!} e^{-\theta}, \qquad k = 0,1, \dots$$

Показать, что для $\tau(\theta) = \frac{1}{\theta}$ **не существует несмещенной оценки**.

Оценка $\hat{\theta}$ называется **несмещённой оценкой параметра θ**, если она не содержит систематических ошибок, то есть при любом фиксированном объёме выборки n результат осреднения оценки по всем возможным выборкам данного объёма должен приводить к точному значению параметра: $E(\hat{\theta}) = \theta, \forall\, \theta \in \Theta$.

Раздел 1. Оценки параметров

Оценка $\hat{\theta}_n$ параметра θ называется **состоятельной**, если она сходится по вероятности к θ, когда объем выборки $n \longrightarrow \infty$.

Раздел 1. Оценки параметров

Достаточное условие состоятельности.

Пусть $\widehat{\theta}_n$ является несмещенной оценкой параметра θ: $E(\widehat{\theta}_n) = \theta$, а ее дисперсия стремится к 0 при $n \longrightarrow \infty$:

$$\lim_{n\to\infty} Var(\widehat{\theta}_n) = 0.$$

Тогда $\widehat{\theta}_n$ обладает свойством **состоятельности**.

Раздел 1. Оценки параметров

Замечание.

Доказанное утверждение остается **справедливым**, если оценка **не является несмещенной**, но обладает **свойством асимптотической несмещенности**.

Раздел 1. Оценки параметров

Пример 1. Покажем, что **выборочное среднее** является **состоятельной** оценкой математического ожидания.

Решение.

Достаточное условие состоятельности. Пусть $\hat{\theta}_n$ является несмещенной оценкой параметра θ: $E(\hat{\theta}_n) = \theta$, а ее дисперсия стремится к 0 при $n \longrightarrow \infty$:

$$\lim_{n\to\infty} Var(\hat{\theta}_n) = 0.$$

Тогда $\hat{\theta}_n$ обладает свойством состоятельности.

Раздел 1. Оценки параметров

Пример 2. Пусть $X_1, X_2, \ldots, X_n$ — выборка из равномерного распределения $Unif([0;\theta])$ с неизвестным параметром $\theta > 0$. Требуется оценить параметр θ. В качестве оценки параметра θ рассматриваются:

$$\hat{\theta}_1 = 2\bar{X}; \quad \hat{\theta}_2 = \frac{n+1}{n} X_{(n)}; \quad \hat{\theta}_3 = (n+1)X_{(1)}.$$

Будут ли оценки несмещёнными? б) Состоятельными?
с) Найти среди них оптимальную.

Решение.

Достаточное условие состоятельности. Пусть $\hat{\theta}_n$ является несмещенной оценкой параметра θ: $E(\hat{\theta}_n) = \theta$, а ее дисперсия стремится к 0 при $n \longrightarrow \infty$:

$$\lim_{n\to\infty} Var(\hat{\theta}_n) = 0.$$

Тогда $\hat{\theta}_n$ обладает свойством **состоятельности**.

Раздел 1. Оценки параметров

Несмещенная оценка $\hat{\theta}$ параметра θ называется **эффективной**, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра θ, вычисленных по выборкам одного и того же объема n, то есть если $\widehat{\Theta}$ - некоторый класс оценок, то оценка $\hat{\theta}$ называется **эффективной оценкой параметра $\boldsymbol{\theta}$ в классе $\widehat{\boldsymbol{\Theta}}$** если

$$Var\left(\hat{\theta}^*\right) = \min_{\hat{\theta}\in\widehat{\Theta}} Var(\hat{\theta}).$$

Раздел 1. Оценки параметров

Из двух несмещенных оценок $\hat{\theta}^{(1)}$ и $\hat{\theta}^{(2)}$ одного и того же параметра θ оценка $\hat{\theta}^{(1)}$ называется **более эффективной**, чем $\hat{\theta}^{(2)}$, если выполняется неравенство:

$$Var(\hat{\theta}^{(1)}) < Var(\hat{\theta}^{(2)}).$$

Средством исследования эффективности оценок является **неравенство Рао – Крамера**.

Раздел 1. Оценки параметров

Рассмотрим функцию

$$f(x;\theta) = \begin{cases} \boldsymbol{P}\{X = x;\ \theta\}, \text{если } X \text{ — дискретная с.в.} \\ p(x;\ \theta), \text{если } X \text{ — абсолютно непрерывная с.в.} \end{cases}$$

$p(x;\ \theta)$ — плотность распределения генеральной совокупности X, если X — непрерывная случайная величина, и вероятность события $\{X = x\}$, когда X — дискретная случайная величина; θ — оцениваемый параметр.

Функция

$$I(\theta) = E\left(\left[\frac{\partial \ln f(x;\theta)}{\partial\theta}\right]^2\right)$$

называется **количеством информации по Фишеру** в одном наблюдении.

Раздел 1. Оценки параметров

Свойства информационного количества

1. Пусть $I_1(\theta)$ и $I_2(\theta)$ – количества информации о θ содержащиеся в независимых величинах X_1 и X_2 соответственно, а $I(\theta)$ количество информации, содержащиеся в паре (X_1, X_2). Тогда $I(\theta) = I_1(\theta) + I_2(\theta)$.

2. Пусть $X_1, \dots, X_n$ – независимые одинаково распределенные величины и I – количество информации о θ, содержащейся в каждой из них. Тогда количество информации в $(X_1, \dots, X_n)$ равно nI.

Раздел 1. Оценки параметров

Упражнение. Показать, что

$$I(\theta) = -E\left(\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right).$$

Раздел 1. Оценки параметров

Пример. Предположим, что имеет $X \sim N(\theta; \sigma^2)$ (нормальное распределение с неизвестным средним θ и известной дисперсией σ^2).

Определить $I(\theta)$ случайной величины X.

Решение.

Раздел 1. Оценки параметров

Теорема (неравенство Рао–Крамера).

Пусть область $G_n = \{x: f(x,\theta) > 0\}$ всех возможных значений случайной величины, для которых $f(x,\theta) \neq 0$, не зависит от параметра θ и $\hat{\theta}_n$ — несмещенная оценка параметра θ, полученная по выборке объема n. Справедливо неравенство

$$Var(\hat{\theta}_n) \geq \frac{1}{nI(\theta)}.$$

Раздел 1. Оценки параметров

Величину

$$e(\hat{\theta}) = \frac{1}{nI(\theta)Var(\hat{\theta})}$$

называют **показателем эффективности по Рао–Крамеру**.

Для любой **несмещенной оценки** параметра θ выполняются неравенства: $0 < e(\hat{\theta}) \leq 1$.

Теорема (неравенство Рао–Крамера).
Пусть область $G_n = \{x: f(x,\theta) > 0\}$ всех возможных значений случайной величины, для которых $f(x,\theta) \neq 0$, не зависит от параметра θ и $\hat{\theta}_n$ — несмещенная оценка параметра θ, полученная по выборке объема n. Справедливо неравенство

$$Var(\hat{\theta}_n) \geq \frac{1}{nI(\theta)}.$$

Раздел 1. Оценки параметров

Несмещенная оценка параметра θ называется **эффективной**, если

$$e\left(\hat{\theta}\right) = 1$$

и **неэффективной**, если

$$e\left(\hat{\theta}\right) < 1.$$

Величину $e\left(\hat{\theta}\right) = \frac{1}{nI(\theta)Var\left(\widehat{\theta}\right)}$ называют **показателем эффективности по Рао–Крамеру**.

Для любой несмещенной оценки параметра θ выполняются неравенства:

$$0 < e\left(\hat{\theta}\right) \leq 1.$$

Раздел 1. Оценки параметров

Асимптотической эффективностью оценки называется предел

$$e_0\,(\hat{\theta}) = \lim_{n\to\infty} e\big(\hat{\theta}_n\big)\,,$$

если он существует.

Оценку $\hat{\theta}$ называют **асимптотически эффективной**, если

$$e_0(\hat{\theta}) = 1.$$

Величину $e\big(\hat{\theta}\big) = \frac{1}{nI(\theta)Var\left(\hat{\theta}\right)}$ называют **показателем эффективности по Рао–Крамеру**.

Для любой несмещенной оценки параметра θ выполняются неравенства:

$$0\ < e\big(\hat{\theta}\big) \leq\ 1.$$

Раздел 1. Оценки параметров

Пример. Пусть $X \sim N(a;\, \sigma^2)$. Решается задача оценивания математического ожидания a. Показать, что оценка $\hat{\theta} = \bar{X}$ неизвестного параметра $\theta = a\ (= E(X))$ является **эффективной**.

Решение.

Теорема (неравенство Рао–Крамера). Пусть $\hat{\theta}_n$ — несмещенная оценка параметра θ, полученная по выборке объема n. Справедливо неравенство

$$Var(\hat{\theta}_n) \geq \frac{1}{nI(\theta)}, \text{где } I(\theta) = E\left(\left[\frac{\partial \ln f(x;\theta)}{\partial\theta}\right]^2\right)$$

Раздел 1. Оценки параметров

Замечание. Не для всякого закона распределения может быть использовано **неравенство Рао-Крамера**.

Так, **например**, с помощью данного неравенства нельзя проверить является ли эффективной та или иная оценка параметра θ, если наблюдаемая случайная величина распределена равномерно на интервале $(0, \theta)$.

Раздел 1. Оценки параметров

Еще одно **важное свойство оценок — сходимость к нормальному закону**.

Обозначим через $L(X)$ — закон распределения случайной величины X.

И пусть $X_1, X_2, \ldots$ — выборка из генеральной совокупности X, $\hat{\theta}_n$ — некоторая оценка параметра θ, построенная по этой выборке.

Статистика $\hat{\theta}_n$ называется **асимптотически нормальной оценкой параметра $\boldsymbol{\theta}$ с коэффициентом $\boldsymbol{\sigma^2(\theta)}$**, если

$$L\left(\frac{\left(\hat{\theta}_n - \theta\right)\sqrt{n}}{\sigma(\theta)}\right) \xrightarrow[n\to\infty]{} N(0;1).$$

В этом случае закон распределения случайной величины $\hat{\theta}_n$ является приближенно нормальным с параметрами θ и $\sigma^2(\theta)$.

Раздел 1. Оценки параметров

Пример. Рассматривается случайная величина X с неизвестными параметрами:

$$E(X) = a, \ \ Var(X) = \sigma^2.$$

Тогда выборочное среднее X, рассматриваемое в качестве оценки математического ожидания a , является асимптотически нормальным $N\left(a; \frac{\sigma^2}{n}\right)$.

Решение. Действительно, $E(\bar{X}) = a$, $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$ и закон распределения случайной величины

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

сходится к нормальному на основании **центральной предельной теоремы**.

Свойства основных выборочных характеристик случайных величин

Раздел 1. Оценки параметров

Будем использовать обозначения:

$$a = E(X), \sigma^2 = Var(X).$$

Раздел 1. Оценки параметров

Выборочное среднее $\bar{X}$ ($\hat{\theta} = \bar{X}$; $\theta = a$)

1. $\bar{X}$ — несмещенная оценка математического ожидания: $E(\bar{X}) = a$.

2. $\bar{X}$ — состоятельная оценка: $\bar{X} \xrightarrow{P} a$.

3. $\bar{X}$ — эффективная оценка в классе всех линейных оценок вида $\sum_{i=1}^{n} c_i X_i$, где $\sum_{i=1}^{n} c_i = 1$.

4. В случае, когда оценивается математическое ожидание a нормально распределенной случайной величины X, оценка $\hat{\theta} = \bar{X}$ является эффективной.

5. Выборочное среднее асимптотически нормально $N\left(a; \frac{\sigma^2}{n}\right)$.

$$a = E(X), \sigma^2 = Var(X)$$

Раздел 1. Оценки параметров

Выборочная дисперсия $\hat{\sigma}^2$ **(**$\hat{\theta} = \hat{\sigma}^2$**;** $\theta = \sigma^2$**)**

1. Выборочная дисперсия $\hat{\sigma}^2$, используемая в качестве оценки дисперсии генеральной совокупности σ^2, является смещенной оценкой: $E\left(\hat{\theta}\right) = E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma_X^2 \neq \theta = \sigma_X^2$).

2. Если генеральная совокупность имеет моменты до четвертого порядка включительно, то $\hat{\theta} = \hat{\sigma}^2$ является состоятельной оценкой дисперсии σ^2: $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$.

3. Выборочная дисперсия является оценкой асимптотически нормальной $N\left(\sigma^2; \frac{\mu_4 - \sigma^2}{n}\right)$.

4. Исправленная выборочная дисперсия $\hat{S}^2 = \frac{n}{n-1}\hat{\sigma}^2$ несмещенная оценка неизвестного параметра σ^2:

$$E\left(\hat{S}^2\right) = \sigma^2$$

$$a = E(X), \sigma^2 = Var(X)$$

Раздел 1. Оценки параметров

Из состоятельности выборочной дисперсии следует, что оценка $\hat{\theta} = \hat{S}^2$ — состоятельная.

Свойством эффективности оценка $\hat{\theta} = \hat{S}^2$ не обладает.

В случае нормального закона распределения:

$$X \sim N(a; \sigma^2)$$

справедливо равенство:

$$e(\hat{S}^2) = \frac{n-1}{n} < 1.$$

$$a = E(X), \sigma^2 = Var(X)$$

Выборочная оценка дисперсии при известном и неизвестном среднем

Раздел 1. Оценки параметров

Если **генеральное среднее** $\boldsymbol{a} = \boldsymbol{E}(\boldsymbol{X})$ **известно**, для приближенного вычисления неизвестной генеральной дисперсии $Var(X)$ используется ее несмещенная оценка:

$$s_0^2 = \frac{1}{n}\left((X_1 - a)^2 + \cdots + (X_n - a)^2\right).$$

Если $\boldsymbol{a} = \boldsymbol{E}(\boldsymbol{X})$ **не известно**, используется исправленная выборочная дисперсия (несмещенная оценка):

$$s_X^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} = \frac{n}{n-1}\hat{\sigma}_X^2.$$

Раздел 1. Оценки параметров

В случае, если **генеральное среднее** $\boldsymbol{a} = \boldsymbol{E}(\boldsymbol{X})$ **известно**, то оценка $\hat{\theta} = S_0^2$ $(\theta = \sigma^2)$ обладает следующими **свойствами**:

1. Оценка $\hat{\theta} = S_0^2$ является **несмещенной** оценкой дисперсии.

Доказательство.

2. Если для случайной величины X существует четвертый центральный момент $E[(X - a)^4]$, то оценка $\hat{\theta} = S_0^2$ обладает свойством **состоятельности**.

Доказательство.

Раздел 1. Оценки параметров

3. В случае, когда оценивается дисперсия σ^2 нормально распределенной случайной величины X с известным математическим ожиданием a, оценка $\hat{\theta} = S_0^2$ является **эффективной**.

Раздел 1. Оценки параметров

Выборочный начальный момент $\hat{\nu}_k$ **(**$\hat{\theta} = \hat{\nu}_k; \theta = \nu_k$**)**

В качестве оценки начального момента k -го порядка используется соответствующая выборочная характеристика

$$\hat{\nu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k.$$

1. $\hat{\nu}_k$ — **несмещенная** оценка начального момента.

2. Если случайная величина X имеет начальный момент ν_{2m}, то оценка $\hat{\nu}_k$ неизвестного параметра ν_k является **состоятельной** оценкой момента ν_k при любом $k \leq m$.

3. Выборочный начальный момент является асимптотически нормальным $N\left(\nu_k; \frac{\nu_{2k}-\nu_k^2}{n}\right)$.

Раздел 1. Оценки параметров

Выборочный центральный момент $\hat{\mu}_k$ ($\hat{\theta} = \hat{\mu}_k$; $\theta = \mu_k$)

В качестве оценки начального момента k-го порядка используется соответствующая выборочная характеристика

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k.$$

1. Свойством несмещённости оценка $\hat{\mu}_k$ центрального момента не обладает.

2. Если случайная величина X имеет начальный момент ν_{2m}, то выбранный в качестве оценки $\hat{\mu}_k$ центрального момента μ_k выборочный центральный момент $\hat{\mu}_k$ обладает свойством состоятельности при любом $k \le m$.

3. Выборочный центральный момент является асимптотически нормальным $N\left(\mu_k; \frac{\mu_{2k} - 2k\mu_{k-1}\mu_{k+1} - \mu_k^2 + k^2\mu_{k-1}^2}{n}\right)$.

Раздел 1. Оценки параметров

Выборочный коэффициент асимметрии $\hat{\mu}_k$ **(**$\hat{\theta} = \widehat{As}$**;** $\theta = As$**)**

В качестве оценки коэффициента асимметрии используется выборочный коэффициент асимметрии

$$\widehat{As} = \frac{\hat{\mu}_3}{\hat{\sigma}^3}.$$

1. Свойством несмещённости оценка $\widehat{As}$ коэффициента асимметрии не обладает.

2. Выборочный коэффициент асимметрии является оценкой состоятельной.

3. В случае, когда оценивается коэффициент асимметрии нормально распределенной случайной величины X, выборочный коэффициент асимметрии является асимптотически нормальным $N\left(0; \frac{6(n-2)}{(n+1)(n+3)}\right)$.

Раздел 1. Оценки параметров

Исправленная выборочная асимметрия

$$\tilde{A}_X = \frac{\sqrt{n(n-1)}}{n-2}\hat{A}_X.$$

Раздел 1. Оценки параметров

Выборочный коэффициент эксцесса $\widehat{Ex}$ **(**$\hat{\theta} = \widehat{Ex}$**;** $\theta = Ex$**)**

В качестве оценки коэффициента эксцесса используется выборочный коэффициент эксцесса

$$\widehat{Ex} = \frac{\hat{\mu}_4}{\hat{\sigma}^4} - 3.$$

1. Свойством несмещённости оценка $\widehat{Ex}$ коэффициента эксцесса не обладает.

2. Выборочный коэффициент эксцесса является оценкой состоятельной.

3. В случае, когда оценивается коэффициент эксцесса нормально распределенной случайной величины X, выборочный коэффициент эксцесса является асимптотически нормальным

$$N\left(-\frac{6}{n+1}; \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}\right).$$

Исправленный выборочный эксцесс

$$\tilde{E}_X = \frac{(n-1)[(n+1)\hat{E}_X + 6]}{(n-2)(n-3)}.$$

Раздел 1. Оценки параметров

Оценка вероятности события $\widehat{\boldsymbol{p}}$ ($\hat{\theta} = \hat{p}$; $\theta = p$)

Рассмотрим задачу оценивания неизвестной вероятности p в биномиальной модели, где p имеет смысл вероятности ≪успеха≫ в любом из n независимых повторных испытаний (испытаний по схеме Бернулли), в которых было зафиксировано k ≪успехов≫.

Оценка вероятности в виде относительной частоты ≪успеха≫ в n испытаниях: $\hat{p} = \frac{k}{n}$ обладает следующими **свойствами**.

1. Оценка $\hat{\theta} = \hat{p}$ является несмещенной (с док.).
2. Оценка $\hat{\theta} = \hat{p}$ является состоятельной.

3. Оценка $\hat{\theta} = \hat{p}$ является эффективной (с док.).

4. Оценка $\hat{\theta} = \hat{p} = \frac{k}{n}$ является асимптотически нормальной $N\left(p; \frac{p(1-p)}{n}\right)$.

Раздел 1. Оценки параметров

Эмпирическая функция распределения $\widehat{F}(x)$

($\hat{\theta} = \hat{F}(x)$; $\theta = F(x)$)

В качестве оценки значения функции распределения $F(x)$ в точке $x \in \mathbb{R}$ используется эмпирическая функция распределения $\hat{F}(x)$.

Перечислим основные свойства.

Раздел 1. Оценки параметров

1. $\hat{\theta} = \hat{F}(x)$ является несмещенной и состоятельной оценкой функции распределения $F(x)$ генеральной совокупности X в точке $x \in \mathbb{R}$.

2. Значение эмпирической функции распределения в каждой точке x является эффективной оценкой для значения в этой точке теоретической функции распределения.

Раздел 1. Оценки параметров

Выборочный коэффициент корреляции $\hat{\rho}_{xy}$
($\hat{\theta} = \hat{\rho}_{xy}$**;** $\theta = \rho_{xy}$**)**

Пусть для случайного вектора (X, Y) имеется выборка $(X_1, Y_1), \ldots, (X_n, Y_n)$ объема n.
Выборочный коэффициент корреляции определяется по формуле

$$\hat{\rho}_{xy} = \frac{\overline{XY} - \bar{X}\bar{Y}}{\hat{\sigma}_x \hat{\sigma}_y}.$$

Замечание. В формуле для выборочного коэффициента корреляции используются обычные выборочные стандартные отклонения $\hat{\sigma}_x$ и $\hat{\sigma}_y$, а не исправленные характеристики $\hat{s}_x$, $\hat{s}_y$. Только при этом условии выполняется свойство: $-1 \leq \hat{\rho}_{xy} \leq 1$.

Перечислим основные свойства.

Раздел 1. Оценки параметров

1. Оценка $\hat{\theta} = \hat{\rho}_{xy}$ коэффициента корреляции ρ_{xy} — смещенная, состоятельная, асимптотически нормальная $N(a; \sigma^2)$ с параметрами $a = \rho_{xy}$,

$$\sigma^2 = \frac{\rho_{xy}^2}{4n}\left(\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} + \frac{4\mu_{22}}{\mu_{11}^2} - \frac{4\mu_{31}}{\mu_{11}\mu_{20}} - \frac{\mu_{13}}{\mu_{11}\mu_{02}}\right),$$

где μ_{ij} — теоретические центральные моменты порядка i и j:

$$\mu_{ij} = E\left([X - E(X)]^i \left[Y - E(Y)\right]^j\right).$$

2. Для нормального двумерного вектора (X, Y) выборочный коэффициент корреляции асимптотически нормальный с параметрами $a = \rho_{xy}$ и $\sigma^2 = \frac{(1-\rho_{xy}^2)^2}{n}$.

Раздел 1. Оценки параметров

3. Сходимость выборочного коэффициента корреляции к нормальному закону медленная. Существенно быстрее сходится к нормальному закону случайная величина

$$Z = \frac{1}{2}\lg\left(\frac{1+\hat{\rho}_{xy}}{1-\hat{\rho}_{xy}}\right).$$

В случае нормального закона распределения вектора (X, Y) случайная величина Z асимптотически нормальна

$$N\ (\lg\left(\frac{1+\hat{\rho}_{xy}}{1-\hat{\rho}_{xy}}\right); \frac{1}{n-3}).$$

Раздел 1. Оценки параметров

Выборочные квантили

В качестве оценки квантиля x_α порядка α используется выборочный квантиль $\hat{X}_\alpha$:

$$\hat{X}_\alpha = \begin{cases} X_{([n\alpha]+1)}, & \text{если } n\alpha \text{ дробное,} \\ X_{(n\alpha)}, & \text{если } n\alpha \text{ целое} \end{cases}.$$

Здесь $[x]$ $-$целая часть числа x.

Перечислим основные свойства.

Раздел 1. Оценки параметров

1. Выборочный квантиль является оценкой состоятельной.

2. Для непрерывной случайной величины X с плотностью распределения $f(x)$ выборочный квантиль является асимптотически нормальным:

$$N\left(x_\alpha; \frac{\alpha(1-\alpha)}{nf^2(x_\alpha)}\right).$$

В частности, выборочная медиана $\widehat{Me}$ асимптотически нормальна

$$N\left(Me; \frac{1}{4nf^2(Me)}\right).$$

Теория вероятностей и математическая статистика

Конец лекции