Содержание

исходники найдутся тут:
https://github.com/pelmesh619/itmo_conspects

Лекция 1.

Теория вероятности изучает характеристику случайных величин, тогда как математическая статистика решает обратную задачу
Допустим, что у нас есть случайная величина, по ней мы можем найти матожидание, моменты и оценить, какое распределение имеет случайная величина.

Выборки

Def. Выборка - набор данных, полученных в ходе экспериментов. Тогда количество экспериментов n - объем Выборки
Def. Генеральной совокупностью называются все результаты проведенных экспериментов
Def. Выборочной совокупностью называются наблюдаемые данные экспериментов
Не все данные экспериментов мы можем наблюдать, например, выборы, тогда опросы голосовавших - выборочная совокупность, а результаты выборов - генеральная. Очевидно, что выборочная и генеральная совокупности могут иметь различные распределения.
Def. Выборка называется **репрезентативной**, если ее распределение близко к распределению генеральной совокупностью
Пример - ошибка выжившего. Во время Второй Мировой стал вопрос, в каких местах стоит бронировать корпус самолета. Самолеты возвращались с пулевыми отверстиям, и интуитивно казалось, что стоит бронировать те места, которые больше всего пострадали. Однако не были учтены те самолеты, которые не вернулись, а те, которые выжили, выжили благодаря тому, что были прострелены в нелетальных местах, поэтому было принято решение бронировать фюзеляж в менее пострадавших местах
В дальнейшем считаем, что все выборки репрезентативны

Def. 1. Выборкой объема n называется набор из n экспериментаных данных $\vec{X} = (x_1, x_2, \dots, x_n)$ (апостериорное определение)
Def. 2. Выборкой объема n называется набор из n независимых одинаково распределенных случайных величин $\vec{X} = (X_1, X_2, \dots, X_n)$ (априорное определение)

Выборочные характеристики

Можно выборку рассматривать как дискретную случайную величину с одинаковыми вероятностями $p_i = \dfrac{1}{n}$ и вычислить для нее математическое ожидание, дисперсию и функцию распределения

Def. Выборочным средним $\overline{x}$ называется величина $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

Def. Выборочной дисперсией D^* называется величина $D^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^2$ (или $D^* = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{x}^2$)

По закону больших чисел выборочное среднее будет сходиться к матожиданию

Def. Исправленной дисперсией называется величина $S^2 = \frac{n}{n-1}D^* = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2$

Def. Выборочной функцией распределения $F^*(x)$ называется функция $F^*(x) = \frac{\text{число данных } x_i < x}{n}$

Th. Выборочная функция распределения поточечно сходится к теоретической функции распределения:

$$\forall y \in \mathbb{R} F^*(y) \xrightarrow{p} F(y)$$

$F(y) = P(X < y)$

$F_y^* = \frac{1}{n}\sum_{i=1}^{n} I(X_i < y) \xrightarrow[\text{по ЗБЧ}]{p} EI(X_i < y) = P(X_i < y) = P(X_1 < y) = F_{X_1}(y)$

Усилим теорему

Th. Гливенко-Кантелли. $\sup\limits_{x\in\mathbb{R}} |F^*(x) - F(x)| \xrightarrow{p} 0$

Th. Колмогорова. $\sqrt{n}\sup\limits_{x\in\mathbb{R}} |F^*(x) - F(x)| \rightrightarrows K$ - распределение Колмогорова с функцией распределения $F_K(x) = \sum\limits_{j=-\infty}^{\infty}(-1)^j e^{-2j^2x^2},\ x \in [0;\infty)$

Начальная обработка статданных

1. Ранжирование данных - упорядочиваем выборки по возрастанию. В результате получаем вариационный ряд $\vec{X} = (X_{(1)}, X_{(2)}, \dots, X_{(n)})$

   $X_{(1)} = \min X_i; \quad X_{(n)} = \max X_i$

   $X_{(i)} = i$-ая порядковая статистика
2. Объединим повторяющиеся данные - получаем т.н. частотный вариационный ряд

X_i	$X_{(1)}$	...	$X_{(r)}$	$\sum$
n_i	n_1	...	n_r	n

   Иногда часть данных отбрасывается сверху и снизу (по 5, по 10, по 5% и так далее), чтобы сделать выборку репрезентативной

   Тогда $\overline{x} = \frac{1}{n}\sum X_i n_i$, $D^* = \frac{1}{n}\sum (X_i - \overline{x})^2 n_i$

3. Чтобы уменьшить количество вычислений или сделать гистограмму, делают интервальный вариационный ряд: разбиваем данные на интервалы и считаем, сколько данных n_i попало в интервал.

   Тогда n_i - частота интервала A_i

   Есть два основные способа разбиения на интервалы:

   (a) Интервалы одинаковой длины

   (b) Равнонаполненные интервалы (в каждом интервале примерно одинаковое количество данных)

   Число интервалов K такое, что $\frac{K(n)}{n} \longrightarrow 0$ и $K(n) \underset{n\to\infty}{\longrightarrow} \infty$

   Обычно применяют формулу Стерджесса $K \approx 1 + \log_2 n$ или $K \approx \sqrt[3]{n}$

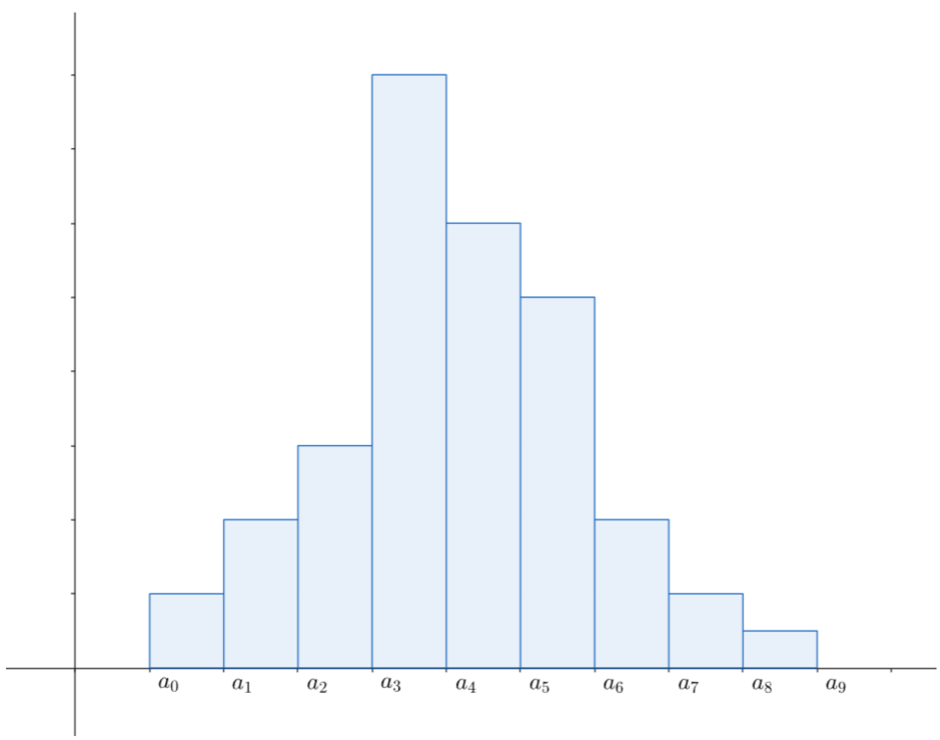
   Пусть получили интервальный вариационный ряд

интервалы	$[a_0; a_1)$	$[a_1; a_2)$	...	$[a_{K-1}; a_K]$	$\sum$
частоты	n_1	n_2	...	n_K	n

Геометрическая интерпретация данных

- Гистограмма

  Строится ступенчатая фигура из прямоугольников, основание i-ого прямоугольника - интервал, высота прямоугольника - $\frac{n_i}{nl_i}$, где l_i - длина интервала

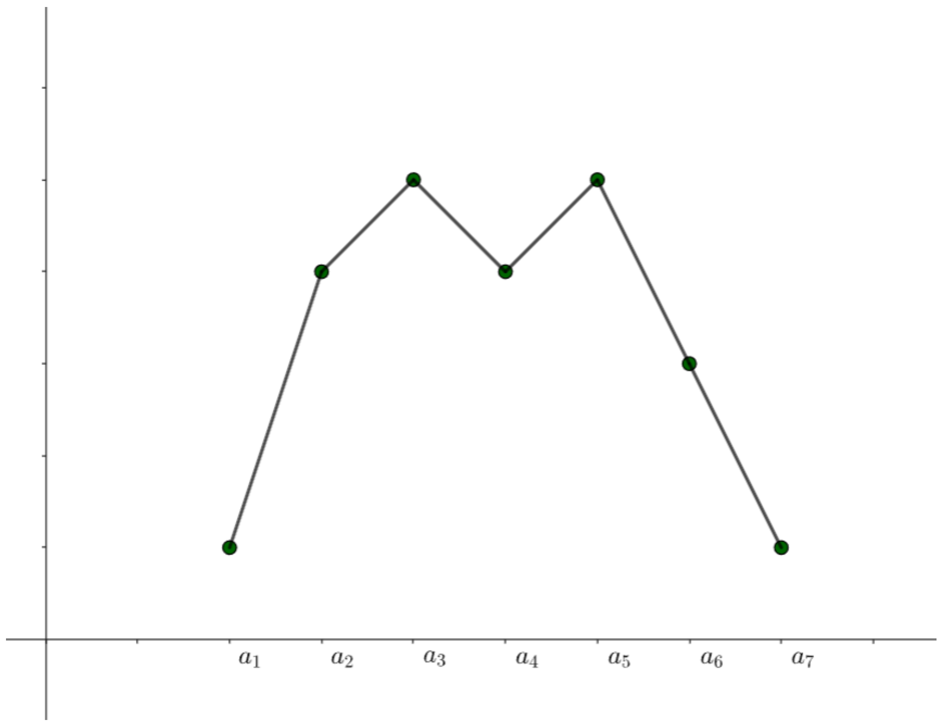


  Визуально можно сделать гипотезу, как ведет себя распределение.

Th. Гистограмма поточечно сходится к теоретической плотности

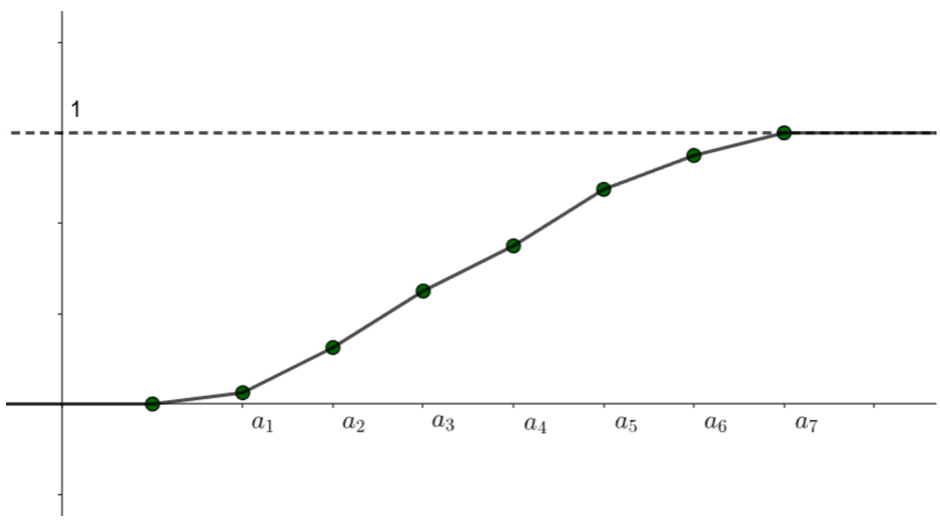
- Полигон

  На оси абсцисс отмечаем значения частотного вариационного ряда, по оси ординат - их частоты. Получившиеся точки соединяем отрезками



- Выборочная функция распределения

  На основе таблицы строится график функции распределения



Она может быть ступенчатой, ломаной или соединена по усмотрению

Лекция 2.

Точечная оценка

Пусть имеется выборка $\vec{X} = (X_1, X_2, \dots, X_n)$ объемом n
Пусть требуется найти приближенную оценку θ^* неизвестного параметра θ
Находим ее при помощи некоторой функции обработки данных $\theta^* = \theta^*(X_1, \dots, X_n)$

Def. Такая функция называется статистикой
Def. А оценка θ^* называется точечной оценкой

Свойство точечных оценок

1. Состоятельность
   Def. Статистика $\theta^* = \theta^*(X_1, \dots, X_n)$ неизвестного параметра называется состоятельной, если $\theta^* \xrightarrow{p} \theta$ при $n \to \infty$

2. Несмещенность
   Def. Оценка θ^* параметра θ называется несмещенной, если математическое ожидание $E\theta^* = \theta$
   Nota. Оценка θ^* называется асимптотически несмещенной, если $E\theta^* \xrightarrow{p} \theta$ при $n \to \infty$

3. Эффективность
   Def. Оценка θ_1^* не хуже θ_2^*, если $E(\theta_1^* - \theta)^2 \leq E(\theta_2^* - \theta)^2$. Или, если θ_1^* и θ_2^* несмещенные, то $D\theta_1^* \leq D\theta_2^*$
   Def. Оценка θ^* называется эффективной, если она не хуже всех остальных оценок
   Nota. Не существует эффективной оценки в классе всех возможных оценок

   Th. В классе несмещенных оценок существует эффективная оценка

4. Асимптотическая нормальность
   Def. Оценка θ^* параметра θ называется асимптотически нормальной, если $\sqrt{n}(\theta^* - \theta) \rightrightarrows N(0, \sigma^2(\theta))$ при $n \to \infty$

Точечные оценки моментов

Def. Выборочным средним $\overline{x}$ называется величина $\overline{x} = \dfrac{1}{n} \sum\limits_{i=1}^{n} X_i$

Def. Выборочной дисперсией D^* называется величина $D^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^2$

Def. Исправленной дисперсией S^2 называется величина $S^2 = \frac{n}{n-1}D^* = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2$

Def. Выборочным средним квадратическим отклонением называется величина $\sigma^* = \sqrt{D^*}$

Def. Исправленным средним квадратическим отклонением называется величина $S = \sqrt{S^2}$

Def. Выборочным k-ым моментом называется величина $\overline{x^k} = \frac{1}{n}\sum_{i=1}^{n}X_i^k$

Def. Модой Mo^* называется варианта x_k с наибольшей частотой $n_k = \max_i(n_1, n_2, \ldots, n_m)$

Def. Выборочной медианой Me^* называется варианта x_i в середине вариационного ряда
$$\begin{cases} \text{Me}^* = X_{(k)}, & \text{если } n = 2k-1 \\ \frac{X_{(k)}+X_{(k+1)}}{2}, & \text{если } n = 2k \end{cases}$$

Th. $\overline{x}$ - состоятельная несмещенная оценка теоретического матожидания ÅX $= a$

1) $E\overline{x} = a$

2) $\overline{x} \xrightarrow{p} a$ при $n \to \infty$

1) $E\overline{x} = E\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n}EX_i = \frac{1}{n}nEX_1 = EX_1 = a$

2) $\overline{x} = \frac{\overline{x}_1 + \cdots + \overline{x}_n}{n} \xrightarrow{p} a$ согласно Закону Больших Чисел

Nota. Если второй момент конечен, то $\overline{x}$ - асимптотически нормальная оценка. По ЦПТ $\frac{S_n - nEX_1}{\sqrt{n}\sqrt{DX_1}} = \sqrt{n}\frac{\overline{x} - EX_1}{\sqrt{DX_1}} \rightrightarrows N(0,1)$ или $\sqrt{n}(\overline{x} - EX_1) \rightrightarrows N(0; DX_1)$

Th. Выборочный k-ый момент является состоятельной несмещенной оценкой теоретического k-ого момента

1) $\overline{EX^k} = EX^k$

2) $\overline{X^k} \xrightarrow{p} X^k$

Это следует из предыдущей теоремы, если взять X^k вместо X

Th. Выборочной дисперсией D^* и S^2 являются состоятельными оценками теоретической дисперсией, при этом D^* - смещенная оценка, а S^2 - несмещенная оценка

Заметим, что $D^* = \overline{X^2} - \overline{X}^2$
$ED^* = E(\overline{X^2} - \overline{X}^2) = E\overline{X^2} - E(\overline{X}^2) = EX^2 - E(\overline{X}^2)$
Так как $D\overline{X} = E(\overline{X}^2) - (E\overline{X})^2$, то $EX^2 - E(\overline{X}^2) = EX^2 - ((E\overline{X})^2 + D\overline{X}) = (EX^2 - EX) - D\overline{X} = DX - D\overline{X} = DX - D\left(\dfrac{X_1 + \dots + X_n}{n}\right) = DX - \dfrac{1}{n^2}\sum_{i=1}^{n} DX_i = DX - \dfrac{1}{n^2} nDX_1 = DX - \dfrac{1}{n}DX = \dfrac{n-1}{n}DX$,
то есть D^* - смещенная вниз оценка
$ES^2 = E(\dfrac{n}{n-1}D^*) = \dfrac{n}{n-1}\dfrac{n-1}{n}DX = DX \Longrightarrow S^2$ - несмещенная вниз оценка
2. $D^* = \overline{X^2} - \overline{X}^2 \xrightarrow{p} EX^2 - (EX)^2 = DX$ - состоятельная оценка
$S^2 = \dfrac{n}{n-1}D^* \xrightarrow{p} DX$

Nota. Отсюда видим, что выборочная дисперсия - асимптотически несмещенная оценка. Поэтому при большом (обычно не меньше 100) объеме выборке можно считать обычную выборочную дисперсию

Метод моментов (Пирсона)

Постановка задачи: пусть имеется выборка объема n неизвестного распределения, но известного типа, которое задается k параметрами: $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Требуется дать оценки данным неизвестным параметрам
Идея метода состоит в том, что сначала находим оценки k моментов, а затем с помощью теоретических формул из теории вероятности даем оценки этих параметров
Пусть $\vec{X}$ - выборка из абсолютно непрерывного распределения F_θ с плотностью известного типа, которая задается k параметрами $f_\theta(x, \theta_1, \dots, \theta_k)$
Тогда теоретические моменты находим по формуле $m_i = \int_{-\infty}^{\infty} x^i f_\theta(x, \theta_1, \dots, \theta_k)dx = h_i(\theta_1, \dots, \theta_n)$
Получаем систему из k уравнений с k неизвестными. В эти уравнения подставляем найденные оценки моментов и, решая получившуюся систему уравнений, находим нужные оценки параметров

$$\begin{cases} \overline{x} = h_1(\theta_1^*, \dots, \theta_n^*) \\ \overline{x^2} = h_2(\theta_1^*, \dots, \theta_n^*) \\ \dots \\ \overline{x^k} = h_k(\theta_1^*, \dots, \theta_n^*) \end{cases}$$

Nota. Оценки по методу моментов как правило состоятельные, но часто смещенные

Ex. Пусть $X \in U(a, b)$. Обработав статданные, нашли оценки первого и второго моментов: $\overline{x} = 2.25; \overline{x^2} = 6.75$

Найти оценки параметров a^*, b^*

Плотность равномерного распределения $f_{(a,b)}(x) = \begin{cases} 0, & x < a \\ \frac{1}{b-a} & a \le x \le b, \\ 0, x > b \end{cases}$

$$EX = \int_a^b x\frac{1}{b-a}dx = \frac{a+b}{2}$$
$$EX = \int_a^b x^2\frac{1}{b-a}dx = \frac{a^2+ab+b^2}{3}$$

Получаем:

$$\begin{cases} \overline{x} = \frac{a^*+b^*}{2} \\ \overline{x^2} = \frac{a^{*2}+a^*b^*+b^{*2}}{3} \end{cases} \iff \begin{cases} \frac{a^*+b^*}{2} = 4.5 \\ a^{*2}+a^*b^*+b^{*2} = 20.25 \end{cases} \iff \begin{cases} \frac{a^*+b^*}{2} = 4.5 \\ a^*b^* = 0 \end{cases} \iff \begin{cases} a^* = 0 \\ b^* = 4.5 \end{cases}$$

Лекция 3.

Метод максимального правдоподобия

Пусть имеется выборка $\vec{X} = (X_1, \dots, X_n)$ из распределения известного типа, определяемого неизвестными параметрами $\theta = (\theta_1, \dots, \theta_n)$

Идея метода состоит в следующем: подбираем параметры таким образом, чтобы вероятность получения данной выборки при случайном эксперименте была наибольшей.

Если распределение дискретное, то $P_\theta(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_1 = x_1)\dots P(X_n = x_n)$

Def. Функцией правдоподобия $L(\vec{X}, \theta)$ называется функция $L(\vec{X}, \theta) = P(X_1 = x_1)\dots P(X_n = x_n) = \prod\limits_{i=1}^{n} P(X_i = x_i)$ при дискретном распределении

и $L(\vec{X}, \theta) = f_\theta(x_1)\dots f_\theta(x_n) = \prod\limits_{i=1}^{n} f_\theta(x_i)$ в абсолютно непрерывном распределении

Def. Логарифмической функцией правдоподобия называется функция $\ln L(\vec{X}, \theta)$

Nota. Так как $y = \ln x$ возврастающая функция, точки максимума совпадают, а такую функцию правдоподобия становится легче дифференцировать

Def. Оценкой максимального правдоподобия $\hat{\theta}$ называется значение θ, при котором функция правдоподобия $L(\vec{X}, \theta)$ достигает наибольшего значения (при фиксированных значениях выборки)

Ex. 1. Пусть $\vec{X} = (X_1, \dots, X_n)$ - выборка из распределения Пуассона Π_λ с неизвестным $\lambda > 0$

Мет. Для распределения Пуассона $P(X = x_i) = \frac{\lambda^{x_i}}{x_i!}e^{-\lambda}$

Получаем функцию максимального правдоподобия $L(\vec{X}, \lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!}e^{-\lambda} = \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}e^{-n\lambda} = \frac{\lambda^{n\overline{x}}}{\prod_{i=1}^n x_i!}e^{-n\lambda}$

$\ln L(\vec{X}, \lambda) = n\overline{x}\ln\lambda - \ln\prod_{i=1}^{n} x_i! - n\lambda$

$\frac{\partial \ln L}{\partial \lambda} = \frac{n\overline{x}}{\lambda} - n = 0 \Longrightarrow \hat{\lambda} = \overline{x}$ - оценка максимального правдоподобия

Убедимся, что этот экстремум - максимум: $\frac{\partial^2 \ln L}{\partial \lambda^2} = -\frac{n\overline{x}}{\lambda} < 0 \Longrightarrow \hat{\lambda} = \overline{x}$ - точка максимума

Ex. 2. Пусть $(X_1, \ldots, X_n)$ из $N(a, \sigma^2)$

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

$L(\vec{X}, a, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-a)^2}{2\sigma^2}} = \frac{1}{\sigma^n(2\pi)^{\frac{n}{2}}}e^{-\frac{\sum_{i=1}^n (x_i-a)^2}{2\sigma^2}}$

$\ln L(\vec{X}, a, \sigma^2) = -n\ln\sigma - \frac{n}{2}\ln 2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - a)^2$

$\frac{\partial \ln L}{\partial a} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} -2(x_i - a) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - a) = \frac{n\overline{x} - na}{\sigma^2}$

$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} - \sum_{i=1}^{n}(x_i - a)^2\frac{1}{2}\cdot(-2)\cdot\sigma^{-3} = \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - a)^2 - \frac{n}{\sigma}$

$\begin{cases} \frac{n\overline{x}-na}{\sigma^2} = 0 \\ \frac{1}{\sigma^3}\sum_{i=1}^n (x_i - a)^2 - \frac{n}{\sigma} = 0 \end{cases} \Longrightarrow \begin{cases} \hat{a} = \overline{x} \\ \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^n (x_i - a)^2 = D^* \end{cases}$

Ex. 3. Пусть $(X_1, \ldots, X_n)$ из $U(0, \theta)$. Найти оценку θ этого распределения.

Воспользуемся методом моментов:

$EX = \frac{a+b}{2} = \frac{\theta}{2} \Longrightarrow \overline{x} = \frac{\theta^*}{2} \Longrightarrow \theta^* = 2\overline{x}$

Воспользуемся методом максимального правдоподобия:

$f_\theta = \begin{cases} 0, & x < 0 \\ \frac{1}{\theta}, & 0 \le x \le \theta \\ 0, & x > \theta \end{cases}$

$X_{(n)} = \max_i(X_1, \ldots, X_n)$

$L(\vec{X}, \theta) = \prod_{i=1}^{n} f_\theta(x_i) = \begin{cases} 0, & \text{если } \theta < X_{(n)} \\ \frac{1}{\theta^n}, & \text{если } \theta \ge X_{(n)} \end{cases}$

$L(\vec{X}, \theta)$ достигает наибольшего значения при наименьшем значении θ^n, то есть при $\hat{\theta} = X_{(n)}$

Сравним оценки:

$\theta^* = 2\overline{x}$ - несмещенная оценка, так как $E\theta^* = 2E\overline{x} = 2EX = \theta$

$E(\theta^* - \theta)^2 = D\theta^* = D2\overline{x} = 4D\overline{x} = 4\frac{D\overline{x}}{n} = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}$

Изучим распределение $X_{(n)}$: $F_{X_{(n)}}(x) = P(X_{(n)} < x) = P(X_1 < x, X_2 < x, \dots, X_n < x) = P(X_1 < x) \dots P(X_n < x) = F_{X_1}(x) \dots F_{X_n}(x) = F^n_{(x_1)}(x)$

$$F_{X_1}(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{\theta}, & 0 \le x \le \theta \\ 1, & x > \theta \end{cases} \Longrightarrow F_{X(n)}(x) = \begin{cases} 0, & x < 0 \\ \frac{x^n}{\theta^n}, & 0 \le x \le \theta \\ 1, & x > \theta \end{cases} \Longrightarrow f_{X(n)}(x) = \begin{cases} 0, & x < 0 \\ n\frac{x^{n-1}}{\theta^n}, & 0 \le x \le \theta \\ 1, & x > \theta \end{cases}$$

$EX_{(n)} = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{\theta^n}\int_0^\theta x^n dx = \frac{nx^{x+1}}{\theta^n(n+1)}\Big|_0^\theta = \frac{n\theta}{n+1}$ - смещенная вниз оценка

$\tilde{\theta} = \frac{n+1}{n}X_{(n)}$ - несмещенная оценка (будем считать, что эффективность не изменилась)

$E\tilde{\theta}^2 = E(\frac{n+1}{n}X_{(n)})^2 = \frac{(x+1)^2}{n^2}EX_{(n)} = \frac{(n+1)^2}{n^2}\int_0^\theta x^2\frac{nx^{n-1}}{\theta^n}dx = \frac{(n+1)^2\theta^2}{n(n+2)}$

$D\tilde{\theta} = E\tilde{\theta}^2 - (E\tilde{\theta})^2 = \frac{\theta^2}{n(n+2)}$

$D\tilde{\theta} = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{3n} = D\theta^*$

Таким образом, оценка по методу правдоподобия сходится быстрее, чем оценка по методу моментов, поэтому она лучше

Отсюда следует, что при равномерном распределении выборочное среднее не является эффективной оценкой для математического ожидания; вместо нее половина максимального элемента выборки будет лучше

Nota. Эффективной здесь будет несмещенная оценка $\frac{n+1}{2n}X_{(n)}$

В общем случае для $U(a, b)$ будет такая эффективная оценка матожидания - $\frac{X_{(1)} + X_{(n)}}{2}$, длины интервала - $\frac{n+1}{n-1}(X_{(n)} - X_{(1)})$

Nota. При методе максимального правдоподобия обычно получаем состоятельные и эффективные оценки, но часто смещенные

Неравенство Рао-Крамера

Пусть $X \in F_\theta$ - семейство распределений с параметром $\theta \in \mathbb{R}$

Def. Носителем семейства распределений F_θ называется множество $C \subset \mathbb{R}$ такое, что $P(X \in C) = 1\ \forall X \in F_\theta$

$$f_\theta(x) = \begin{cases} \text{плотность } f_\theta(x) \text{ при непрерывном распределении} \\ P_\theta(X = x) \text{ при дискретном распределении} \end{cases}$$

Def. Информацией Фишера $I(\theta)$ семейства распределений F_θ называется величина $I(\theta) =$

$E\left(\frac{\partial}{\partial\theta}\ln f_\theta(X)\right)^2$ при условии, что она существует

Def. Семейство распределений F_θ называется регулярным, если:

- существует носитель C семейства F_θ такой, что $\forall x \in C$ функция $\ln f_\theta(x)$ непрерывно дифференцируема по θ
- информация Фишера $I(\theta)$ существует и непрерывна по θ

Th. Пусть $(X_1, \dots, X_n)$ - выборка объема n из регулярного семейства F_θ, $\theta^* = \theta^*(X_1, \dots, X_n)$ - несмещенная оценка параметра θ, дисперсия которой $D\theta^*$ ограничена в любой замкнутой ограниченной области параметра θ

Тогда $\boxed{D\theta^* \geq \frac{1}{nI(\theta)}}$

Следствие: если при данных условиях получили $D\theta^* = \frac{1}{nI(\theta)}$, то оценка θ^* является эффективной (то есть дальше улучшать уже некуда)

Ex. Пусть $(X_1, \dots, X_n)$ из $N(a, \sigma^2)$ (то есть $F_a = N(a, \sigma^2)$, σ^2 зафиксируем)

Проверим эффективность $a^* = \overline{x}$

Плотность $f_a(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$, носитель - вся прямая $\mathbb{R}$

$\ln f_a(x) = -\ln\sigma - \frac{1}{2}\ln 2\pi - \frac{(x-a)^2}{2\sigma^2}, \qquad a \in (-\infty, \infty)$

$\frac{\partial}{\partial a}\ln f_a(x) = \frac{1}{2\sigma^2}\cdot 2(x-a) = \frac{x-a}{\sigma^2}$ - непрерывна для всех $a \in \mathbb{R}$

$I(a) = E\left(\frac{\partial}{\partial a}\ln f_a(X)\right)^2 = E\left(\frac{X-a}{\sigma^2}\right)^2 = \frac{1}{\sigma^4}E(X-a)^2 = \frac{E(X-EX)^2}{\sigma^4} = \frac{DX}{\sigma^4} = \frac{1}{\sigma^2}$ - непрерывна по a

Из этого следует, что $N(a, \sigma^2)$ - регулярное семейство относительно параметра a

$Da^* = D\overline{x} = \frac{DX}{n} = \frac{\sigma^2}{n}$ - ограничена по параметру a

По неравенству Рао-Крамера $Da^* = \frac{\sigma^2}{n} = \frac{1}{nI(a)} = \frac{1}{n}\sigma^2$; из этого следует, что a^* - эффективная оценка параметра a

Nota. Аналогично можно показать, что S^2 - несмещенная эффективная оценка для параметра σ^2

Лекция 4.

Основные распределения математической статистики

Def. Случайная величина имеет нормальное распределение $\xi \in N(a, \sigma^2)$ с параметрами a и σ^2, если ее плотность имеет вид $f_\xi(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$
На практике нормальное распределение встречается чаще всего в силу ЦПТ

Def. Распределение $N(0,1)$ с параметрами $a = 0, \sigma^2 = 1$ называется стандартным нормальным распределением. Его плотность равна $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. В дальнейшем такую случайную величину будем называть стандартной нормалью

Свойства

1. $a = E\xi \qquad \sigma^2 = D\xi$
2. Линейность: $\xi \in N(a, \sigma^2)$, то $\eta = b\xi + \gamma \in N(ab + \gamma, b^2\sigma^2)$
3. Стандартизация: Если $\xi \in N(a, \sigma^2)$, то $\eta = \dfrac{\xi - a}{\sigma} \in N(0,1)$
4. Устойчивость относительно суммирования: если $\xi_1 \in N(a_1, \sigma_1^2)$, $\xi_2 \in N(a_2, \sigma_2^2)$, независимы то $\xi_1 + \xi_2 \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

Распределение «хи-квадрат»

Def. Распределение «хи-квадрат» H_n со степенями свободы n называется распределение суммы квадратов независимых стандартных нормальных величин: $\chi_n^2 = X_1^2 + X_2^2 + \dots + X_n^2$, где $X \in N(0,1)$ и независимы

Свойства

1. $E\chi_n^2 = n$

   Так как $\forall i\ X_i \in N(0,1)$, то $EX_i^2 = DX_i^2 + (EX_i)^2 = 1 \Longrightarrow E(X_i^2 + \dots X_n^2) = \sum\limits_{i=1}^{n} X_i^2 = n$

2. Устойчивость относительно суммирования: если $X \in H_n$, $Y \in H_m$, независимы, то $X + Y \in H_{n+m}$ (по определению)
3. $\dfrac{\chi_k^2}{k} \xrightarrow[k\to\infty]{p} 1$ (по Закону Больших Чисел)

Распределение Стьюдента

Def. Пусть $X_0, X_1, \dots, X_k$ - независимые стандартные нормальные величины. Распределением Стьюдента T_k с k степенями свободы называется распределение случайной величины $t_k = \dfrac{X_0}{\sqrt{\frac{1}{k}(X_1^2 + \dots + X_k^2)}} = \dfrac{X_0}{\sqrt{\frac{\chi_k^2}{k}}}$

Свойства

1. $Et_k = 0$ - в силу симметрии
2. $t_k \rightrightarrows N(0,1)$ (на практике при $k \geq 100$ распределение Стьюдента можно считать стандартным нормальным)

Распределение Фишера-Снедекера

Def. Распределением Фишера-Снедекера $F_{n,m}$ (другое название - F-распределение) со степенями свободы n и m называется распределение случайной величины $f_{n,m} = \dfrac{\frac{\chi_n^2}{n}}{\frac{\chi_m^2}{m}}$, где χ_n^2 и χ_m^2 - независимые случайные величины с распределением «хи-квадрат»

Свойства

1. $Ef_{n,m} = \dfrac{n}{n-2}$
2. $f_{n,m} \xrightarrow[n,m\to\infty]{p} 1$

Математическое ожидание и дисперсия случайного вектора

Пусть $\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ - случайный вектор, где случайная величина X_i - компонента (координата) случайного вектора

Def. Математическим ожидание случайного вектора называется вектор с координатами из математических ожиданий компонент: $E\vec{X} = \begin{pmatrix} EX_1 \\ \vdots \\ EX_n \end{pmatrix}$

Def. Дисперсией случайного вектора (или матрицей ковариаций) случайного вектора $\vec{X}$ называется матрица $D\vec{X} = E(\vec{X} - E\vec{X})(\vec{X} - E\vec{X})^T$, состоящая из элементов $d_{ij} = \operatorname{cov}(X_i, X_j)$

Nota. На главной диагонали стоят дисперсии компонент: $d_{ii} = DX_i$

Nota. $D\vec{X}$ - симметричная положительно определенная матрица

Свойства

1. $E(A\vec{X}) = AE\vec{X}$
2. $E(\vec{X}+\vec{B}) = E\vec{X}+\vec{B}$, где $\vec{B}$ - вектор чисел
3. $D(A\vec{X}) = A\cdot D\vec{X}\cdot A^T$
4. $D(\vec{X}+\vec{B}) = D\vec{X}$

Многомерное нормальное распределение

Def. Пусть случайный вектор $\vec{\xi} = \begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix}$ имеет вектор средних $\vec{a} = E\vec{\xi}$, K - симметричная положительно определенная матрица. Вектор $\vec{\xi}$ имеет нормальное распределение в $\mathbb{R}^n$ с параметрами $\vec{a}$ и K, если его плотность $f_{\vec{\xi}}(\vec{X}) = \dfrac{1}{\left(\sqrt{2\pi}\right)^n\sqrt{\det K}}e^{-\frac{1}{2}(\vec{X}-\vec{a})^TK^{-1}(\vec{X}-\vec{a})}$

Свойства

1. Матрица $K = D\vec{\xi} = \big(\operatorname{cov}(\xi_i,\xi_j)\big)$ - матрица ковариаций
2. При $\vec{a}=\vec{0}$ и $K=E$ имеем вектор из независимых стандартных нормальных величин

При $\vec{a} = \vec{0}$ и $K = E$: $f_{\vec{\xi}}(X_1,\ldots,X_n) = \dfrac{1}{\left(\sqrt{2\pi}\right)^n}e^{-\frac{1}{2}\begin{pmatrix}X_1 & \ldots & X_n\end{pmatrix}E\begin{pmatrix}X_1 & \ldots & X_n\end{pmatrix}^T} =$ $\dfrac{1}{\left(\sqrt{2\pi}\right)^n}e^{-\frac{1}{2}(X_1^2+\dots+X_n^2)} = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}X_1^2}+\dots+\dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}X_n^2}$
Так как плотность распалась на произведение плотностей стандартного нормального распределение, то все компоненты имеют стандартное нормальное распределение

Далее вектором из независимых стандартных нормальных величин для краткости будем называть стандартным нормальным вектором

3. $\sqsupset\vec{X}$ - стандартный нормальный вектор, B - невырожденная матрица, тогда вектор $\vec{Y} = B\vec{X}+\vec{a}$ имеет многомерное нормальное распределение с параметрами $\vec{a}$ и $K = BB^T$
4. $\sqsupset\vec{Y}\in N(\vec{a},K)$. Тогда вектор $\vec{X} = B^{-1}(\vec{Y}-\vec{a})$ - стандартный нормальный вектор, где $B = \sqrt{K}$
   Следствие. Эквивалентное определение: Многомерное нормальное распределение - это то, которое получается из стандартного нормального вектора при помощи невырожденного преобразования и сдвиг
5. $\sqsupset\vec{X}$ - стандартный нормальный вектор, C - ортогональная матрица. Тогда $\vec{Y} = C\vec{X}$ - стандартный нормальный вектор

Так как C - ортогональная, то $C^T = C^{-1}$. Тогда по третьему свойству $K = CC^T = E$, а по второму свойству $\vec{Y}$ - стандартный нормальный вектор

6. $\sqsupset$ случайный вектор $\xi \in N(\vec{a}, K)$. Тогда его координаты независимы тогда и только тогда, когда они не коррелированы (то есть матрица ковариаций K диагональная)
   Следствие. Если плотность совместного распределения случайных величин ξ и η ненулевая, то они независимы тогда и только тогда, когда их коэффициент корреляции равен нулю

Многомерная центральная предельная теорема

Th. Среднее арифметическое независимых одинаково распределенных случайных векторов слабо сходится к многомерному нормальному распределению

Лемма Фишера

Пусть вектор $\vec{X}$ - стандартный нормальный вектор, C - ортогональная матрица, $\vec{Y} = C\vec{X}$. Тогда $\forall 1 \le k \le n-1$ случайная величина $T(\vec{X}) = \sum\limits_{i=1}^{n} X_i^2 - Y_1^2 - Y_2^2 - \dots Y_k^2$ не зависит от $Y_1, Y_2, \dots, Y_k$ и имеет распределение «хи-квадрат» со степенями свободы $n-k$

Так как C - ортогональное преобразование, то $\|\vec{X}\| = \|\vec{Y}\|$, то есть $\sum\limits_{i=1}^{n} X_i^2 = \sum\limits_{i=1}^{n} Y_i^2 \Longrightarrow$

$$T(\vec{X}) = \sum_{i=1}^{n} X_i^2 - Y_1^2 - Y_2^2 - \dots Y_k^2 = Y_{k+1}^2 + \dots + Y_n^2$$

Согласно свойству 5 $Y_i \in N(0,1)$ и независимы, то по определению «хи-квадрат» $T(\vec{X}) \in H_{n-k}$ и не зависит от $Y_1, \dots, Y_k$

Основная теорема

Эта теорема также известна как основное следствие леммы Фишера

Th. Пусть $(X_1, \dots, X_n)$ - выборка из нормального распределения $N(a, \sigma^2)$, $\overline{x}$ - выборочное среднее, S^2 - исправленная дисперсия.

Тогда справедливы следующие высказывания:

1. $\sqrt{n}\dfrac{\overline{x}-a}{\sigma} \in N(0,1)$
2. $\displaystyle\sum_{i=1}^n \frac{(X_i-a)^2}{\sigma^2} \in H_n$
3. $\displaystyle\sum_{i=1}^n \frac{(X_i-\overline{x})^2}{\sigma^2} = \frac{nD^*}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \in H_{n-1}$
4. $\sqrt{n}\dfrac{\overline{x}-a}{S} \in T_{n-1}$
5. $\overline{x}$ и S^2 независимы

1. Так как $X_i \in N(a,\sigma^2)$, то $\displaystyle\sum_{i=1}^n X_i \in N(na, n\sigma^2) \Longrightarrow \overline{x} \in N\left(a, \frac{\sigma^2}{n}\right) \Longrightarrow \overline{x}-a \in N\left(0, \frac{\sigma^2}{n}\right) \Longrightarrow$ $\dfrac{\sqrt{n}}{\sigma}(\overline{x}-a) \in N(0,1)$
2. Так как $X_i \in N(a,\sigma^2)$, то $\dfrac{X_i-a}{\sigma} \in N(0,1)$ и $\displaystyle\sum_{i=1}^n \frac{(X_i-a)^2}{\sigma^2} \in H_n$ по определению
3. $\displaystyle\sum_{i=1}^n \frac{(X_i-\overline{x})^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i-a}{\sigma} - \frac{\overline{x}-a}{\sigma}\right)^2 = \sum_{i=1}^n (z_i-\overline{z})^2$, где $z_i = \dfrac{X_i-a}{\sigma} \in N(0,1)$, $\overline{z} = \displaystyle\frac{1}{n}\sum_{i=1}^n z_i = \frac{\sum_{i=1}^n X_i - na}{\sigma} = \frac{\overline{x}-a}{\sigma}$

   Поэтому можно считать, что изначально $X_i \in N(0,1)$

   $T(\vec{X}) = \displaystyle\sum_{i=1}^n (X_i-\overline{x})^2 = nD^* = n(\overline{x^2}-\overline{x}^2) = \sum_{i=1}^n X_i^2 - n\overline{x}^2 = \sum_{i=1}^n X_i^2 - Y_1^2$, где $Y_1 = \sqrt{n}\overline{x} = \dfrac{X_1}{\sqrt{n}} + \dots + \dfrac{X_n}{\sqrt{n}}$

   Строчка $\left(\dfrac{1}{\sqrt{n}}, \dots, \dfrac{1}{\sqrt{n}}\right)$ имеет длину 1, поэтому ее можно дополнить до ортогональной матрицы C, тогда Y_1 - первая компонента $\vec{Y} = C\vec{X}$, и согласно лемме Фишера $T(\vec{X}) = \displaystyle\sum_{i=1}^n (X_i-\overline{X})^2 \in H_{n-1}$
5. Согласно лемме Фишера $T(\vec{X}) = (n-1)S^2$ не зависит от $Y_1 = \sqrt{n}\overline{x} \Longrightarrow S^2$ и $\overline{x}$ - независимы
4. $\sqrt{n}\dfrac{\overline{x}-a}{S} = \sqrt{n}\dfrac{\overline{x}-a}{\sigma} \cdot \dfrac{1}{\sqrt{\frac{S^2(n-1)}{\sigma^2}\cdot\frac{1}{n-1}}} = \dfrac{\sqrt{n}\frac{\overline{x}-a}{\sigma}}{\frac{\chi^2_{n-1}}{n-1}}$

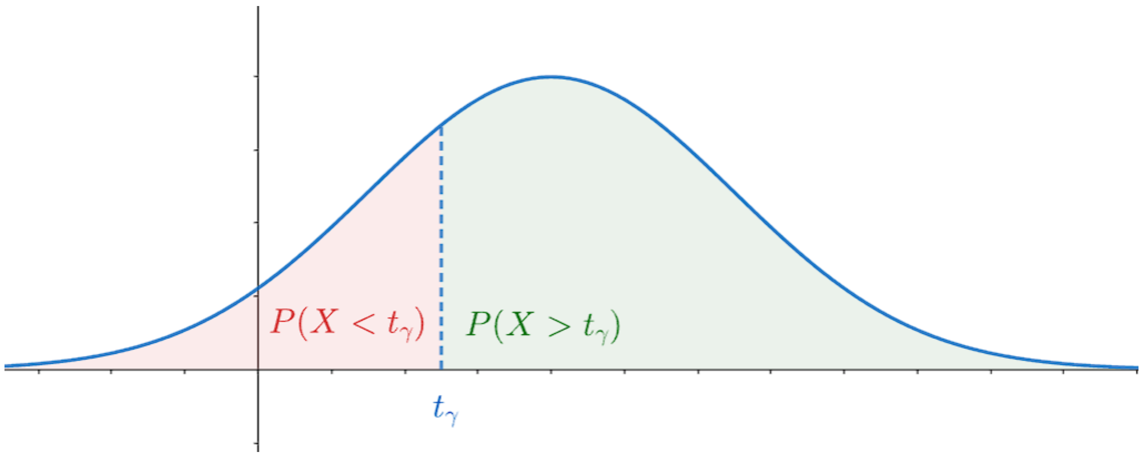
   Так как по пятому пункту числитель и знаменатель независимы, по определению получаем распределение Стьюдента

Лекция 5.

Квантильное распределение

Предполагаем, что распределение абсолютно непрерывное и $F(x)$ - фунцкия распределения

Def. 1. Число t_γ называется квантилем распределения уровня γ, если значения функции распределения $F(t_\gamma) = \gamma$ или $P(X < t_\gamma) = \gamma$ $(t_\gamma = F^{-1}(\gamma))$



Ex. Медиана - квантиль уровня $\dfrac{1}{2}$

Def. 2. Число t_α называется квантилем уровня значимости α, если $P(X > t_\alpha) = \alpha$ или $F(t_\alpha) = 1 - \alpha$
Ясно, что $\gamma = 1 - \alpha$

Интервальные оценки

Недостатки точечных оценок - неизвестно насколько они далеки от реального значения параметра и насколько им можно доверять. Особенно это заметно при малых выборках. Поэтому мы указываем интервал, в котором лежит этот параметр с заданной вероятностью (надежностью) γ. Такие оценки называются интервальными (доверительными)

Def. Интервал $(\theta_\gamma^-; \theta_\gamma^+)$ называется доверительным интервалом параметра θ надежности γ, если вероятность $P(\theta_\gamma^- < \theta < \theta_\gamma^+) = \gamma$

Nota. Если имеем дискретную случайную величину, то $P(\theta_\gamma^- < \theta < \theta_\gamma^+) \geq \gamma$
Nota. Так как параметр θ - константа, то бесмысленно говорить о его попадании в интервал. Правильно: доверительный интервал накрывает параметр θ с вероятностью γ

Nota. 1. $\alpha = 1 - \gamma$ называется уровнем значимости доверительного интервала

Nota. 2. Обычно пытаются строить симметричный доверительный интервал относительно несмещенной оценки θ^*

Nota. 3. Возникает вопрос, какой уровень γ выбрать для исследования. Стандартные уровени надежности γ: $0.9,\ 0.95,\ 0.99,\ 0.999$. Самый мейнстримный - 0.95. В малых выборках используют 0.9

Вспомним основную теорему:

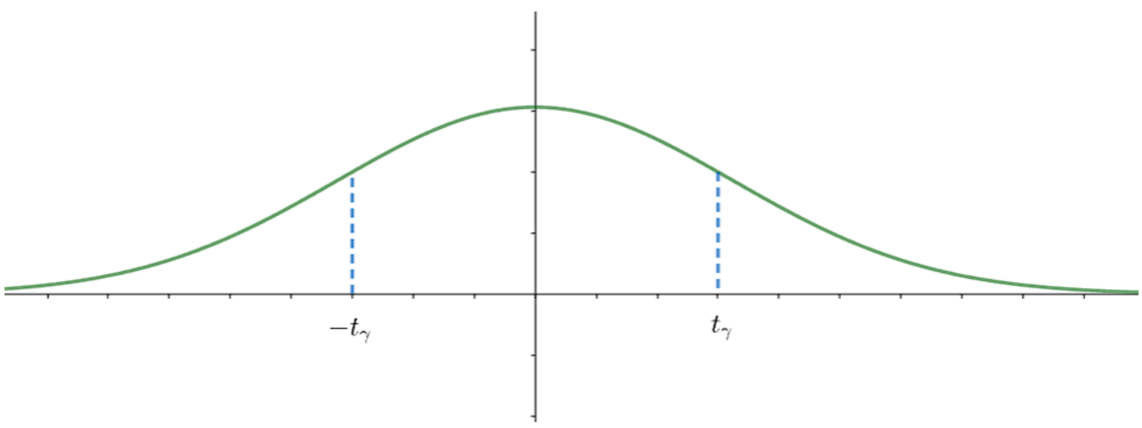
$\sqsupset (X_1, \dots, X_n)$ - выборка объема n из $N(\alpha, \sigma^2)$

$\overline{x}$ - выборочное среднее, S^2 - исправленная дисперсия, D^* - выборочная дисперсия

Тогда:

1. $\sqrt{n}\dfrac{\overline{x} - a}{\sigma} \in N(0, 1)$
2. $\displaystyle\sum_{i=1}^{n}\left(\frac{X_i - a}{\sigma}\right)^2 = \frac{n\tilde{\sigma^2}}{\sigma^2} \in H_n$, где $n\tilde{\sigma^2} = \displaystyle\sum_{i=1}^{n}(X_i - a)^2$
3. $\displaystyle\sum_{i=1}^{n}\left(\frac{X_i - \overline{x}}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{nD^*}{\sigma^2} \in H_{n-1}$
4. $\sqrt{n}\dfrac{\overline{x} - a}{S} \in T_{n-1}$
5. $\overline{x}$ и S^2 - независимы

Nota. Если $F(x)$ - функция симметричного относительно $x = 0$ распределения, то $P(|X| < t) = 2F(t) - 1$



Доверительные интервалы для параметров нормального распределения

Пусть $\vec{X} = (X_1, \dots, X_n)$ - выборка объема n из $N(a, \sigma^2)$. Хотим найти интервалы для параметров a и σ^2

I. Доверительный интервал для параметра a при известном значении σ^2

По пункту 1 из теоремы $\sqrt{n}\dfrac{\overline{x}-a}{\sigma} \in N(0,1)$

$P\left(-t_\gamma < \sqrt{n}\dfrac{\overline{x}-a}{\sigma} < t_\gamma\right) = P\left(\left|\sqrt{n}\dfrac{\overline{x}-a}{\sigma}\right| < t_\gamma\right) = 2F_0(t_\gamma) - 1 = \gamma$

$F_0(t_\gamma) = \dfrac{1+\gamma}{2} \Longrightarrow t_\gamma$ - квантиль уровня $\dfrac{1+\gamma}{2}$ для $N(0,1)$, где $F_0(x) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz$

Решая неравенство, получаем $-t_\gamma < \sqrt{n}\dfrac{\overline{x}-a}{\sigma} < t_\gamma$

$-t_\gamma\dfrac{\sigma}{\sqrt{n}} < \overline{x} - a < t_\gamma\dfrac{\sigma}{\sqrt{n}}$

$\overline{x} - t_\gamma\dfrac{\sigma}{\sqrt{n}} < a < \overline{x} + t_\gamma\dfrac{\sigma}{\sqrt{n}}$ - симметричный интервал относительно $\overline{x}$

Доверительный интервал надежности γ: $\left(\overline{x} - t_\gamma\dfrac{\sigma}{\sqrt{n}}, \overline{x} + t_\gamma\dfrac{\sigma}{\sqrt{n}}\right)$, где t_γ - квантиль $N(0,1)$ уровня $\dfrac{1+\gamma}{2}$

II. Доверительный интервал для параметра a при неизвестном σ^2

Из пункта 4 из теоремы $\sqrt{n}\dfrac{\overline{x}-a}{S} \in T_{n-1}$

$P\left(-t_\gamma < \sqrt{n}\dfrac{\overline{x}-a}{S} < t_\gamma\right) = P\left(\left|\sqrt{n}\dfrac{\overline{x}-a}{S}\right| < t_\gamma\right) = 2F_{T_{n-1}}(t_\gamma) = \gamma$

$F_{T_{n-1}}(t_\gamma) = \dfrac{1+\gamma}{2} \Longrightarrow t_\gamma$ - квантиль T_{n-1} уровня $\dfrac{1+\gamma}{2}$

Аналогично с примером выше получаем интервал $\left(\overline{x} - t_\gamma\dfrac{S}{\sqrt{n}}, \overline{x} + t_\gamma\dfrac{S}{\sqrt{n}}\right)$, где t_γ - квантиль T_{n-1} уровня $\dfrac{1+\gamma}{2}$

III. Доверительный интервал для параметра σ^2 при неизвестном a

По пункту 3 из теоремы $\displaystyle\sum_{i=1}^{n}\left(\frac{X_i - \overline{x}}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{nD^*}{\sigma^2} \in H_{n-1}$

Пусть χ_1^2 и χ_2^2 - квантили H_{n-1} уровней $\dfrac{1-\gamma}{2}$ и $\dfrac{1+\gamma}{2}$

Тогда $P\left(\chi_1^2 < \dfrac{(n-1)S^2}{\sigma^2} < \chi_2^2\right) = F_{H_{n-1}}(\chi_1^2) - F_{H_{n-1}}(\chi_2^2) = \dfrac{1-\gamma}{2} - \dfrac{1+\gamma}{2} = \gamma$

$\chi_1^2 < \dfrac{(n-1)S^2}{\sigma^2} < \chi_2^2$

$\dfrac{1}{\chi_2^2} < \dfrac{\sigma^2}{(n-1)S^2} < \dfrac{1}{\chi_1^2}$

$\dfrac{(n-1)S^2}{\chi_2^2} < \sigma^2 < \dfrac{(n-1)S^2}{\chi_1^2}$ или $\dfrac{nD^*}{\chi_2^2} < \sigma^2 < \dfrac{nD^*}{\chi_1^2}$

Получаем интервал $\left(\dfrac{(n-1)S^2}{\chi_2^2}, \dfrac{(n-1)S^2}{\chi_1^2}\right)$, где χ_1^2 и χ_2^2 - квантили H_{n-1} уровней $\dfrac{1-\gamma}{2}$ и $\dfrac{1+\gamma}{2}$

Nota. Данный интервал не симметричен относительно неизвестного параметра σ^2

IV. Доверительный интервал для параметра σ^2 при известном a

По пункту 2 из теоремы $\sum\limits_{i=1}^{n}\left(\dfrac{X_i - a}{\sigma}\right)^2 = \dfrac{n\tilde{\sigma^2}}{\sigma^2} \in H_{n-1}$

Пусть χ_1^2 и χ_2^2 - квантили H_n уровней $\dfrac{1-\gamma}{2}$ и $\dfrac{1+\gamma}{2}$

Тогда $P\left(\chi_1^2 < \dfrac{n\tilde{\sigma^2}}{\sigma^2} < \chi_2^2\right) = F_{H_n}(\chi_1^2) - F_{H_n}(\chi_2^2) = \dfrac{1-\gamma}{2} - \dfrac{1+\gamma}{2} = \gamma$

Аналогично получаем интервал $\left(\dfrac{n\tilde{\sigma^2}}{\chi_2^2}, \dfrac{n\tilde{\sigma^2}}{\chi_1^2}\right)$, где χ_1^2 и χ_2^2 - квантили H_n уровней $\dfrac{1-\gamma}{2}$ и $\dfrac{1+\gamma}{2}$, $n\tilde{\sigma^2} = \sum\limits_{i=1}^{n}(X_i - a)^2$

Nota. $\tilde{\sigma^2} - D^* = \dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i - a)^2 - \dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{x})^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i^2 - 2aX_i + a^2 - X_i^2 + 2\overline{x}X_i - \overline{x}^2) = \dfrac{1}{n}(na^2 - 2an\overline{x} + 2\overline{x}\cdot n\overline{x} - n\overline{x}^2) = a^2 - 2a\overline{x} + \overline{x}^2 = (a-\overline{x})^2 \Longrightarrow \tilde{\sigma^2} = D^* + (a-\overline{x})^2$

Получаем $\left(\dfrac{n(D^* + (a-\overline{x})^2)}{\chi_2^2}, \dfrac{n(D^* + (a-\overline{x})^2)}{\chi_1^2}\right)$

Асимптотические доверительные интервалы

Def. Интервал $(\theta_\gamma^-, \theta_\gamma^+)$ называется асимптотическим доверительным интервалом параметра θ уровня γ, если $P(\theta_\gamma^- < \theta < \theta_\gamma^+) \underset{n\to\infty}{\longrightarrow} \gamma$

Ex. Доверительный интервал вероятности события A

Пусть $p = P(A), q = 1-p, n$ - число испытаний или объем выборки $(X_1, \ldots, X_n)$, где $X_i \in \{0, 1\}$

$p^* = \dfrac{n_A}{n} = \overline{x}$ - оценка p

Согласно Центральной предельной теореме $\sqrt{n}\dfrac{p^* - p}{DX_1} = \sqrt{n}\dfrac{p^* - p}{\sqrt{pq}} \rightrightarrows N(0,1)$

Так как $p^* \xrightarrow{p} p$, то $\sqrt{n}\dfrac{p^* - p}{\sqrt{p^*(1-p^*)}} = \sqrt{n}\underbrace{\dfrac{p^* - p}{\sqrt{p(1-p)}}}_{\rightrightarrows N(0,1)}\underbrace{\dfrac{\sqrt{p(1-p)}}{\sqrt{p^*(1-p^*)}}}_{\xrightarrow{p} 1} \rightrightarrows N(0,1)$

$P\left(\left|\sqrt{n}\dfrac{p^* - p}{\sqrt{p^*(1-p^*)}}\right| < t_\gamma\right) \underset{n\to\infty}{\longrightarrow} 2F_0(t_\gamma) - 1 = \gamma$

$F_0(t_\gamma) = \dfrac{1+\gamma}{2}$, t_γ - квантиль $N(0,1)$ уровня $\dfrac{1+\gamma}{2}$

Получаем $\left|\sqrt{n}\dfrac{p^* - p}{\sqrt{p^*(1-p^*)}}\right| < t_\gamma$

$|p^* - p| < t_\gamma\dfrac{\sqrt{p^*(1-p^*)}}{\sqrt{n}}$

Итак, $\left(-t_\gamma \dfrac{\sqrt{\overline{x}(1-\overline{x})}}{\sqrt{n}}, t_\gamma \dfrac{\sqrt{\overline{x}(1-\overline{x})}}{\sqrt{n}}\right)$, где t_γ - квантиль $N(0,1)$ уровня $\dfrac{1+\gamma}{2}$