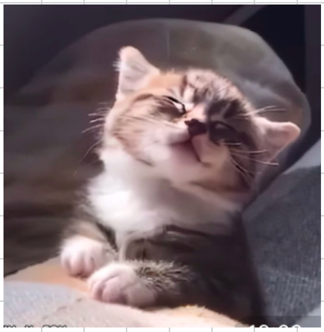
Разбор 5ых задач кр



N1 Робастные оценки

by Полина Свет

$X_1 \dots X_n$ – выборка

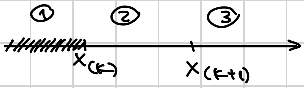
выборочная медиана

$\boxed{n=2k}$

Д-ть $\hat\theta = \begin{cases} \dfrac{X_{(k)}+X_{(k+1)}}{2}, & n=2k \\ X_{(k)}, & n=2k+1 \end{cases}$ – B-робастная оценка

$$SC(x) = \frac{\hat\theta_{n+1}-\hat\theta_n}{\frac{1}{n+1}} = \underbrace{(n+1)(\hat\theta_{n+1}-\hat\theta_n)}_{\text{если это ограничено, то B-робастная}}$$ – кривая чувствительности

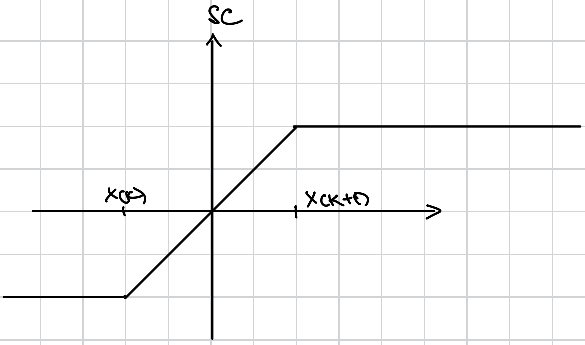
$$\hat\theta_k = \frac{X_{(k)}+X_{(k+1)}}{2}$$



① $x < X_{(k)}$: $\hat\theta_{n+1} = X_{(k)}$

② $x \in [X_{(k)}; X_{(k+1)}]$: $\hat\theta_{n+1} = x$

③ $x > X_{(k+1)}$: $\hat\theta_{n+1} = X_{(k+1)}$



ограничена => у нас B-робастная оценка

N2 Задача про Неймана-Пирсона

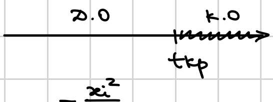
by Максим Клычков

$X_1 \dots X_n \sim N(0,\sigma^2)$

$H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 = \sigma_1^2,\ \sigma_1^2 > \sigma_0^2$

$T(x) = \dfrac{L(x,\sigma_1^2)}{L(x,\sigma_0^2)}$



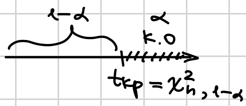
$$\frac{L(x,\sigma_1^2)}{L(x,\sigma_0^2)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{x_i^2}{2\sigma_1^2}}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{x_i^2}{2\sigma_0^2}}} = \underbrace{\left(\frac{\sigma_0}{\sigma_1}\right)^n}_{=\,const} e^{\overbrace{\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)}^{=\,const}\boxed{\sum_{i=1}^n x_i^2}} \geq c_\alpha$$

$$\sum x_i^2 \geq c_\alpha^*$$

$X_i \sim N(0,\sigma_0^2)$

$\sum_{i=1}^n u_i \sim \chi_n^2,\ u_i \sim N(0,1)$

$\sum_{i=1}^n \left(\frac{x_i}{\sigma_0}\right)^2 \sim \chi_n^2$



N3 by Данил Стадник

$$F(x) = \begin{cases} 1-e^{-(x-\theta)^2}, & x>\theta \\ 0, & x \leq \theta \end{cases} \longleftarrow x_1, x_2 \dots x_n$$

$\hat\theta = x_{(1)}$

Д-ть: не является несмещенной, является асимптотич. несмещенной

$$F_{x_{(1)}}(x) = 1-(1-F_x(x))^n = (1-(1-(1-e^{-(x-\theta)^2}))^n = 1-e^{-n(x-\theta)^2}$$

$$E\hat\theta = \int_{-\infty}^{+\infty} x f_{x_{(1)}}(x)dx = \int_\theta^{+\infty} x\, dF_{x_{(1)}}(x) = \int_\theta^{+\infty} x\, d(1-e^{-(x-\theta)^2}) = \theta + \int_\theta^{+\infty} e^{-(\sqrt{n}(x-\theta))^2}dx = \theta + \frac{1}{\sqrt n}\int_0^{+\infty} e^{-(\sqrt n(x-\theta))^2} d(\sqrt n(x-\theta)) = \left| \begin{matrix} t=\sqrt n(x-\theta) \\ x=\theta \Rightarrow t=0 \\ x\to+\infty \Rightarrow t\to+\infty \end{matrix} \right| =$$

$$= \theta + \frac{1}{\sqrt n}\underbrace{\int_0^{+\infty} e^{-t^2}dt}_{\frac{\sqrt\pi}{2}} = \theta + \frac{\sqrt\pi}{2\sqrt n}$$

N4 by Олег slay Сафонов

$X_1, \dots, X_{400}$ — выборка по $\xi \sim R(0;10)$

$\hat F_n(x) = \frac{k}{n}$, где k – количество элементов, которые $\le x$

Д-ть: $\hat F_n(2)$ – несмещенная и несостоятельная $F_\xi(2)$

$\hat F_n(x) = \frac{1}{n}\sum_{i=1}^{n} [x_i \le x]$; $x \sim Bi(n;p)$

$n = 400$ – кол-во элементов из выборки

$F_\xi(x) = \begin{cases} 0, & x<0 \\ \frac{x}{10}, & x \in [0;10] \\ 1, & x>0 \end{cases}$; $F_\xi(2) = \frac{2}{10} = 0,2$

$p = 0,2$

$\hat F_n(x) = \frac{1}{n}\sum_{i=1}^{n} [x_i \le x]$

$x \sim Bi(n;p)$

$x \sim Bi(400;0,2) \Rightarrow Ex = np = 80$, $Dx = npq = 64$

$E[F_n(2)] = F_\xi(2)$? : $E[F_n(2)] \xrightarrow{\text{по ЗБЧ: частота события}} \xrightarrow[n\to\infty]{} p$ $0,2$ $\Rightarrow \hat F_n(x)$ – несмещенная

состоятельная

из лекций $D[\hat F_n(x)] = \frac{F_\xi(x)(1-F_\xi(x))}{\sqrt{n}}$

$P(|\hat F_n(2) - F_\xi(2)| > \varepsilon) = 0 \xRightarrow{\text{по неравенству Чебышева}} P(|\hat F_n(2) - F_\xi(2)| > \varepsilon) \le \frac{D[\hat F_n(2)]}{\varepsilon^2} = \frac{F_\xi(2)\cdot(1-F_\xi(2))}{\sqrt{n}\,\varepsilon^2} = \frac{0,2\cdot 0,8}{\sqrt{n}\cdot\varepsilon^2} \xrightarrow[n\to\infty]{} 0 \Rightarrow \hat F_n(2)$ – состоятельная оценка

$P(|\hat F_n(2) - F_\xi(2)| < 0,04)$

по ЦПТ: $\hat F_n(2) \sim N(np, npq)$

$-0,04 < \frac{1}{n}\sum_{i=1}^{n} [x_i \le x] - F_\xi(2) < 0,04$

$-0,04n < \underbrace{\sum_{i=1}^{n} [x_i \le x] - nF_\xi(2)} < 0,04n$

по ЦПТ: $\sum_{i=1}^{n} [x_i \le x] - nF_\xi(2)) \sim N(0, \overset{64}{npq})$

$$\frac{-0,04\cdot 400}{\sqrt{64}} < \underbrace{\frac{x - 400\cdot 0,2}{\sqrt{64}}}_{\sim\, z \sim N(0,1)} < \frac{0,04\cdot 400}{\sqrt{64}}$$

$$P(|\hat F_n(2) - F_\xi(2)| < 0,04) = 2\Phi_0\left(\overset{=2}{\frac{0,04\cdot 400}{8}}\right) = 2\cdot 0,4772 \approx 0,95$$

Вроде надо серым выделить. Задача 2 (семинар 8)

$H_0: \Delta = 1$ (где $\Delta = \frac{\sigma_y}{\sigma_x}$)

$H_1: \ne, >, <$

X_m – 1-ая выборка

Y_n – 2-ая выборка

ранг эл-та X_i в объед выборке

Критерий Ансари-Брэдли: $A_{m,n} = \sum_{i=1}^{n}\left(\frac{N+1}{2} - \left|R_i - \frac{N+1}{2}\right|\right)$

$Z_N = [X_1, \dots, X_m, Y_1, \dots, Y_m]$ – объединённая выборка, $N = n+m$

$$E\{A_{n,m}\} = \begin{cases} \frac{m(N+2)}{4}, & N = 2k \\ \frac{m(N+1)^2}{4N}, & N = 2k+1 \end{cases}$$

$$D\{A_{n,m}\} = \begin{cases} \frac{m n (N+2)(N-2)}{48(N-1)}, & N = 2k \\ \frac{m\cdot n\cdot(N^2+3)(N+1)}{48N^2}, & N = 2k+1 \end{cases}$$

Проранжировали эл-ты и подчеркиваем эл-ты из выборки X_m

Ранг	1	2	3	4	5	6	7	8	…	22
Вес	42,2	46	46,4	46,5	47	47,7	48,3	48,6		59,8

$N = 11 + 11 = 22$

$$A^* = \frac{A - \mathbb{E}\{A\}}{\sqrt{\mathbb{D}\{A\}}} \sim N(0,1)$$

$t_M = (\qquad\qquad) = 53 = 52 + 1$

$\mathbb{E}\{A\}$ = <по формуле>

$\mathbb{D}[A]$ = <по формуле>

$t_{H,скор} =$

$$\mathcal{D}\{A_{m,n}\} = \begin{cases} \frac{mn(N+2)(N-2)}{48(N-1)}, & \text{если } N\text{-четное} \\ \frac{mn(N^2+3)(N+1)}{48N^2}, & \text{если } N\text{-нечетное} \end{cases}$$

$$A^*_{m,n} = \frac{A_{m,n} - E\{A_{m,n}\}}{\sqrt{\mathcal{D}\{A_{m,n}\}}} \overset{n\to\infty}{\sim} N(0,1)$$

③ $X_m = [X_1, \dots, X_m]^T$ $(m = 11)$ – размеры деталей до наладки

$Y_n = [Y_1, \dots, Y_n]^T$ $(n = 11)$ – размеры деталей после наладки

Размер деталей до и после наладки соответствует задан. нормативу (по условию) $\Rightarrow$ медианы СВ X и Y одинаковы $\mu_x = \mu_y = \mu$ и параметр μ совпадает с нормативным размером

Т.к. наладка станка производится с целью уменьшения отклонения размеров изготовл. деталей от размера, требуемого стандартом, то можно считать, что распределен. СВ X и Y различаются лишь параметром масштаба Δ, т.е.

$F_X(t) = F(t - \mu)$; $F_Y(t) = F\left(\frac{t-\mu}{\Delta}\right)$

а $F(\mu) = 0.5$

Тогда $H_0: \Delta = 1$ – выборки однородны и наладка не привела к ожидаемому рез-ту

$H_a: \Delta < 1$ – $\mathcal{D}[X] > \mathcal{D}[Y]$, т.е. точность увеличилась

$\Delta = \frac{\sigma_y}{\sigma_x}$

$N = m + n = 22$

$$T(Z_n) = A_{m,n} = \sum_{i=1}^{m} \left(\frac{N+1}{2} - \left| R_i - \frac{N+1}{2} \right| \right)$$

Объединим выборки и проранжируем объединенную выборку:

Ранг	1	2	3	4	5	6	7	8	9	10
Вес	42.2	46	46.4	46.5	47	47.7	48.3	48.6	48.8	49.1

11	12	13	14	15	16	17	18	19	20	21	22
49.3	49.7	49.8	51.2	51.5	51.6	52	52.4	52.9	56.1	56.6	59.8

Вектор искомых реализаций рангов:

$[r_1, \dots, r_{11}]^T = [18, 20, 8, 4, 2, 1, 9, 21, 22, 12, 16]^T$

$$A_{11,11} = \frac{23}{2} - \left(18 - \frac{23}{2}\right) + \frac{23}{2} - \left(20 - \frac{23}{2}\right) + \frac{23}{2} - \left(\frac{23}{2} - 8\right) + \frac{23}{2} - \left(\frac{23}{2} - 4\right) + \frac{23}{2} - \left(\frac{23}{2} - 2\right) + \frac{23}{2} - \left(\frac{23}{2} - 1\right) + \frac{23}{2} - \left(\frac{23}{2} - 9\right) + \frac{23}{2} - \left(21 - \frac{23}{2}\right) + \frac{23}{2} - \left(22 - \frac{23}{2}\right) + \frac{23}{2} - \left(12 - \frac{23}{2}\right) + \frac{23}{2} - \left(16 - \frac{23}{2}\right) = 53.$$

Таблицы точного распределения $A_{m,n}$ составлены только для выборок объема $m+n \le 20$, поэтому для построения КО воспользуемся аппроксимацией

$$A^*_{m,n} = \frac{A_{m,n} - E\{A_{m,n}\}}{\sqrt{\mathcal{D}\{A_{m,n}\}}} \sim N(0,1)$$

при $N^* = \min(m,n) \to \infty$

…ытать 225 доброволь-
…фиксировано существенное улучш…
…й интервал уровн…

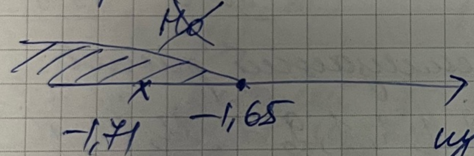
$N = m+n = 22$ — четное $\Rightarrow$

$$E\{A_{11,11}\} = \frac{m(N+2)}{4} = \frac{11\cdot 24}{4} = 66$$

$$D\{A_{11,11}\} = \frac{mn(N+2)(N-2)}{48(N-1)} = \frac{11\cdot 11\cdot 24\cdot 20}{48\cdot 21} = 57{,}62$$

$$\Rightarrow A^*_{11,11} = \frac{53-66}{\sqrt{57{,}62}} = -1{,}71$$

$$t_{кр} = z_{\alpha} = z_{0.05} = -1.65$$



$\Rightarrow \cancel{H_0} \Rightarrow$ точность изготовления деталей увеличилась.