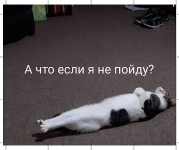
Лекция 4 31.01



но я пошла :(

Робастные оценки

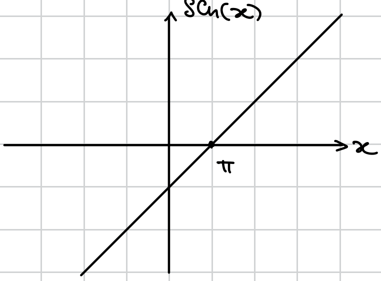
Опр Пусть оценка $\hat{\theta}_n$ построена по выборке $x_1, x_2, \dots, x_n$. Затем добавлено наблюдение x и оценка $\hat{\theta}_{n+1}$ построена по $x_1, \dots, x_n$ и x. Тогда кривой чувствительности, измеряющей влияние наблюдения x на оценку $\hat{\theta}$, называется функция

$$SC_n(x) = \frac{\hat{\theta}_{n+1} - \hat{\theta}_n}{\frac{1}{n+1}} = (n+1)(\hat{\theta}_{n+1} - \hat{\theta}_n)$$

Опр Оценка $\hat{\theta}$ называется B-робастной, если $SC_n(x)$ ограничена.

Пример Пусть $\hat{\theta} = \bar{x}$

$$SC_n(x) = (n+1)\left(\frac{1}{n+1}\left(\sum_{i=1}^n x_i + x\right) - \frac{1}{n}\sum_{i=1}^n x_i\right) = \sum_{i=1}^n x_i + x - \left(\sum_{i=1}^n x_i + \frac{1}{n}\sum_{i=1}^n x_i\right) = x - \bar{x}$$



Выборочная медиана

элемент вариационного ряда

$$\hat{\mu} = \begin{cases} X_{(k+1)}, & \text{если } n = 2k+1 \\ \frac{X_{(k)} + X_{(k+1)}}{2}, & \text{если } n = 2k \end{cases}$$

Опр Пороговой точкой (BP) ε_n^* оценки $\hat{\theta}$, построенной по выборке $x_1, \dots, x_n$ называется

$\varepsilon_n^* = \frac{1}{n}\max\{m : \max_{\xi_1 \dots \xi_n} \sup_{y_1 \dots y_n} |\hat{\theta}(z_1 \dots z_n)| < \infty\}$, где выборка $z_1 \dots z_n$ получена заменой значений $x_{i_1}, \dots x_{i_m}$ на произвольные значения $y_1, \dots y_m$

Доверительный интервал

$x_1, x_2 \dots x_n \sim F(x, \theta)$, $\theta \subset \Theta \subset R^1$

Опр Пусть построены статистики

$T_1(x_1, \dots x_n)$ и $T_2(x_1, \dots, x_n)$, т.ч. $T_1(x) < T_2(x)$

и $P(T_1(x) < \theta < T_2(x)) = 1 - \alpha$, $0 < \alpha < 1$

Тогда интервал $(T_1(x), T_2(x))$ называется доверительным интервалом уровня надежности (доверия) $1-\alpha$ параметра θ

Опр Случайная функция $G(x_1, \dots, x_n, \theta) = G(x, \theta)$ называется центральной (опорной) статистикой, если

1) $G(x, \theta)$ непрерывна и монотонна по θ
2) $F_G(x)$ не зависит от θ

$P(G(x,\theta) < z_{1-\alpha}) = 1 - \alpha$

$P(z_\alpha < G(x,\theta)) = 1 - \alpha$

$P(z_{\alpha/2} < G(x,\theta) < z_{1-\frac{\alpha}{2}}) = 1 - \alpha$

Распределение хи-квадрат (χ^2)

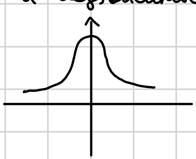
Опр Пусть СВ $\xi_1, \dots, \xi_m \sim N(0;1)$ и независимы

Тогда СВ $\eta = \sum_{i=1}^m \xi_i^2 \sim \chi^2(m)$

Распределение Стьюдента (Student)

Опр Пусть $\xi_0, \xi_1 \dots, \xi_m \sim N(0,1)$ и независимы

Тогда СВ $\zeta = \frac{\xi_0}{\sqrt{\frac{1}{m}\sum_{i=1}^m \xi_i^2}} \sim t(m)$



$m = 1$ — распределение Коши (мат. ожид ∞)

$m = 2$ — (нет ожид конеч, дисперс ∞)

Распределение Фишера

Опр Пусть СВ $\xi_1 \sim \chi^2(m)$, $\xi_2 \sim \chi^2(n)$ и ξ_1 и ξ_2 – независим.
Тогда СВ $F = \dfrac{\frac{1}{m}\xi_1}{\frac{1}{n}\xi_2} \sim F(m,n)$

Теорема Фишера

Пусть $X_1, \dots, X_n$ порождена СВ $X \sim N(m, \sigma^2)$

Тогда

1) $\dfrac{nS^2}{\sigma^2} = \sum\limits_{i=1}^{n}\left(\dfrac{x_i - \bar{x}}{\sigma}\right)^2 \sim \chi^2(n-1)$

2) $\bar{X}$ и S^2 – независимые СВ

---

① $x_1, \dots x_n \sim N(\theta, \sigma^2)$, σ^2 – известный

ДИ → доверительный интервал

ДИ для θ

$\hat{\theta} = \bar{x} \sim N(\theta, \frac{\sigma^2}{n})$

$\dfrac{\sqrt{n}(\bar{x} - \theta)}{\sigma} \sim N(0,1)$

$G(x, \theta)$

$$P\left(z_{\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{x} - \theta)}{\sigma} < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\frac{-z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}} - \bar{x} < -\theta < \frac{z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}} - \bar{x}\right) = 1 - \alpha$$

$$P\left(\bar{x} - \frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{\sqrt{n}} < \theta < \frac{z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$\left(\bar{x} - \frac{z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}} ; \bar{x} + \frac{z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right)$$

---

② $x_1, \dots x_n \sim N(m, \theta_2^2)$, m – известный

ДИ для θ_2^2

$\sum\limits_{i=1}^{n}\left(\dfrac{x_i - m}{\theta_2}\right)^2 \sim \chi^2(n)$

$$P\left(\chi^2_{n, \frac{\alpha}{2}} < \frac{\sum_{i=1}^{n}(x_i - m)^2}{\theta_2^2} < \chi^2_{n, 1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\frac{\sum_{i=1}^{n}(x_i - m)^2}{\chi^2_{n, \frac{\alpha}{2}}} < \theta_2^2 < \frac{\sum_{i=1}^{n}(x_i - m)^2}{\chi^2_{n \frac{\alpha}{2}}}\right) = 1 - \alpha$$

---

③ $x_1, \dots, x_n \sim N(\theta_1, \theta_2^2)$, θ_1 и θ_2 неизвестн

ДИ для θ_2^2

$\sum\limits_{i=1}^{n}\left(\dfrac{x_i - \bar{x}}{\theta_2}\right) = G(x, \theta_2) \sim \chi^2(n-1)$

$$P\left(x^2_{n-1 \frac{\alpha}{2}} < \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\theta_2^2} < x^2_{n-1 \ 1-\frac{\alpha}{2}}\right) < 1 - \alpha$$

$$P\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{x^2_{n-1 \ 1-\frac{\alpha}{2}}} < \theta_2^2 < \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{x^2_{n-1 \ \frac{\alpha}{2}}}\right) = 1 - \alpha$$

---

④ $x_1 \dots x_n \sim N(\theta_1, \theta_2^2)$ θ_1 и θ_2 неизвестн

ДИ для θ_1

$$\frac{\sqrt{n}\left(\frac{\bar{x} - \theta_1}{\sigma}\right)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sigma}\right)^2}} = \frac{\sqrt{n}(\bar{x} - \theta_1)}{s} \sim t(n-1)$$

$$P\left(t_{n-1 \frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{x} - \theta_1)}{s} < t_{n-1 \ 1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\overline{x} - \frac{t_{n-1,\,1-\frac{\alpha}{2}} \cdot \tilde{s}}{\sqrt{n}} < \theta_1 < \overline{x} + \frac{t_{n-1,\,1-\frac{\alpha}{2}} \cdot \tilde{s}}{\sqrt{n}}\right) = 1-\alpha$$