Общий случай

4.3 продолж.

$X = (X_1 \ldots X_n) \sim P \in \mathcal{P}$, $\quad \mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$ – дом. сем-во с $p_\theta(x)$

$L_X(\theta) = \prod_{i=1}^{n} p_\theta(x_i)$ – ф-ция правдоподобия

Рассмотрим гипотезы $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_1$

статистика отнош. правд.: $\quad \lambda(x) = 2 \log \dfrac{L_X(\hat{\theta}_1)}{L_X(\hat{\theta}_0)}$

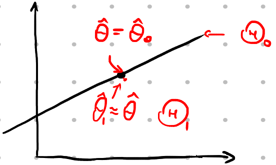
где $\hat{\theta}_0, \hat{\theta}_1$ – ОМП пар-ра θ на Θ_0, Θ_1

Замечание: Обычно рассматривается случай $\Theta \subset \mathbb{R}^D$

$$\Theta_0 \sqcup \Theta_1 = \Theta$$

$$\dim \Theta_0 = d < D$$

Тогда ОМП $\hat{\theta}_1$ по мн-ву Θ_1 совпадает с глоб. ОМП $\hat{\theta}$



$\Rightarrow$ статистика $$\lambda(x) = 2 \log \frac{L_x(\hat{\theta})}{L_x(\hat{\theta}_0)} = 2 \log \frac{\sup\limits_{\theta \in \Theta} L_x(\theta)}{\sup\limits_{\theta \in \Theta_0} L_x(\theta)}$$

Теорема Пусть $\Theta \subset \mathbb{R}^d$ и $\Theta_0 = \{\theta \in \Theta \mid \theta_{d+1} = \underline{\theta^0_{d+1}} \dots \theta_D = \underline{\theta^0_D}\}$

известные константы

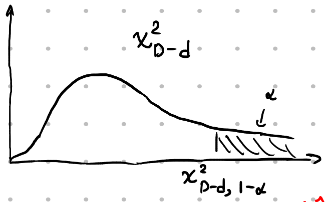
Тогда $\lambda(x) \xrightarrow{d_0} \chi^2_{D-d}$

Пример $\Theta = \mathbb{R}^5$ $\quad H_0: \theta_4 = \theta_5 = 0 \quad D = 5 \quad d = 3$

$$\lambda(x) \xrightarrow{d_0} \chi^2_2$$

$\Rightarrow$ Критерий:

$$S = \{\lambda(x) > \chi^2_{D-d,\,1-\alpha}\}$$



$\lambda(x)$ – большая $\Rightarrow L(\hat\theta) > L(\hat\theta_0) \Rightarrow$

$\Rightarrow$ альт лучше $\Rightarrow$ отверг.

Замечание: при $\lambda(x) \approx 0 \Rightarrow L_x(\hat\theta_0) \approx L_x(\hat\theta) \Rightarrow$

$\Rightarrow H_0$ правдоподобно

Непараметрический подход 5

Эмпирическое распределение 5.1

$X = (X_1 \ldots X_n) \sim P \in \mathcal{P}$, где $\mathcal{P} = \{$все расп.$\}$

Опр. Расп. $\hat{P}_n$ наз-ся эмпирическим расп по выборке X,

если $\forall B \in \mathcal{B}_\mathcal{X}$: $\hat{P}_n(B) = \frac{1}{n} \sum\limits_{i=1}^{n} I\{X_i \in B\}$

Св-ва:

1) $\hat{P}_n(B)$ — случ. величина, доля эл-тов X, попавших в B

2) $\hat{P}_n$ — случ. вер. мера

3) $n\hat{P}_n(B) \sim Bin(n, P(B)) \Rightarrow E\hat{P}_n(B) = P(B)$

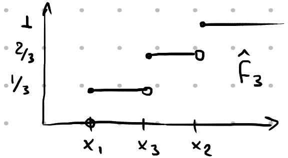
$$D\hat{P}_n(B) = \frac{P(B)(1-P(B))}{n}$$

4) $\hat{P}_n(B) \xrightarrow{\text{п.н.}} P(B)$

Рассмотрим $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, тогда $\hat{P}_n$ соотв. эмпирическая функция

Распр. (ЭФР) $\hat{F}_n(x) = \hat{P}_n((-\infty; x]) = \frac{1}{n}\sum\limits_{i=1}^{n} I\{X_i \le x\}$

Утв. $\hat{F}_n(x) \xrightarrow{\text{п.н.}} F(x)$ ← ф.р. для P



Теорема Гливенко – Кантелли

Если F – ф.р. для P и $X = (X_1 \ldots X_n)$ – выборка из распр. P

то $D_n = \sup\limits_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| \xrightarrow{\text{п.н.}} 0$

Замечание $D_n = \sup\limits_{B\in\mathcal{A}} |\hat{P}_n(B) - P(B)|$, $\mathcal{A} = \{(-\infty; x] \mid x \in \mathbb{R}\}$

Обобщается на другие $\mathcal{A}$, напр $\mathcal{A} = \{$ все объединения инт $\le k$ шт.$\}$

нельзя $\mathcal{A} = \{$ все конечные об. инт.$\}$

Теорема Вапника – Червоненкиса

$$\sup_{B \in \mathcal{A}} |\hat{P}_n(B) - P(B)| \xrightarrow{п.н.} 0$$

$\Leftrightarrow$ конечна VC-размерность $\mathcal{A}$ при разбиении $\mathbb{R}^d$ мн-ми из $\mathcal{A}$

Теорема Колмогорова – Смирнова

$$\sqrt{n}\, D_n \xrightarrow{d} \xi \qquad \xi \sim \text{распр. Колмогорова}$$

+ F – непр. ф-ция распр.

$$F_\xi(x) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2x^2} I\{x \geq 0\}$$

Вывод: скорость сходимости $n^{1/2}$

Метод подстановки

$X = (X_1 \ldots X_n) \sim P$ с. ф.р. F

$\theta = G(P)$ — параметр, значение кот. хотим оценить

$\uparrow$

функционал от распр.

$\hat{\theta} = G(\hat{P}_n)$ — оценка методом подстановки

Пример 1) $\theta = G(P) = E_P f(X_1) = \int_{\mathbb{R}} f(x)\, dF(x)$

$$\hat{\theta} = G(\hat{P}) = E_{\hat{P}_n} f(X_1) = \int_{\mathbb{R}} f(x)\, d\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} f(x_i) = \overline{f(X)}$$

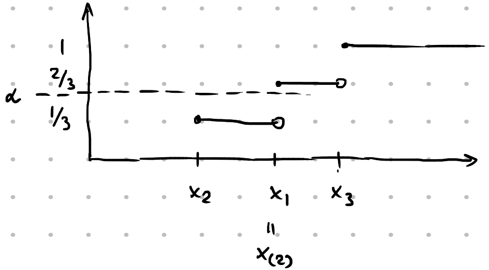
$$f(x) = x \Rightarrow \hat{\theta} = \overline{X}$$

Пример 2) $\theta = G(P) = D_P X_1 = E_P X_1^2 - (E_P X_1)^2$

$$\hat{\theta} = E_{\hat{P}_n} X_1^2 - E_{\hat{P}_n} X^2 = \overline{X^2} - \overline{X}^2 = S$$

3) $\theta = G(P) = F^{-1}(\alpha) = \min\{x \mid F(x) \geq \alpha\}$

$$\hat{\theta} = G(\hat{P}_n) = \min\{x \mid \hat{F}_n(x) \geq \alpha\} = X_{(\lceil n\alpha \rceil)}$$



$\alpha = 1/2$

$\lceil 3 \cdot 1/2 \rceil = 2$