Точечное оценивание

неизв θ положили $= \hat{\theta}$ - оценке

Если распр. $\hat{\theta}$ непр., то $P_\theta(\hat{\theta} = \theta) = 0$

Получаем заведомо ложное утверждение

Опр. $X = (X_1 \ldots X_n)$ – выборка из неизв. $P \in \{P_\theta \mid \theta \in \Theta\}$ $\Theta \subset \mathbb{R}$

пара статистик $T_1(x)$, $T_2(x)$ наз-ся дов. интервалом для θ,

если $P(T_1(x) \le \theta \le T_2(x)) \ge \alpha$

Опр. Если $\Theta \subset \mathbb{R}^d$, то $S(x) \subset \mathbb{R}^d$ наз-ся дов. областью, если

$P(\theta \in S(x)) \ge \alpha$

Опр Посл-ть пар статистик $\{T_1^n(x), T_2^n(x)\}$ наз-ся асимптотическим дов. инт. пар-ра θ уровня доверия α, если $\lim\limits_{n\to\infty} P_\theta\left(T_1^n(x) \le \theta \le T_2^n(x)\right) \ge \alpha$

Замечание: Если $X = (X_1 \dots X_n)$ – выборка, то

утв. $P_\theta\left(T_1(X) \le \theta \le T_2(X)\right) \ge \alpha$ и $(T_1(X), T_2(X))$ – дов. инт. имеет смысл

$x = (x_1 \dots x_n)$ – реализация выборки, то

— // —

не корректно

Пример $X_1 \ldots X_n \sim Pois(\theta)$

Построить дов. инт. ур. дов. α

$E_\theta X_1 = \theta$

$D_\theta X_1 = \theta$

Нер-во Чебышёва $P\left(\left|\sum_{i=1}^n X_i - E\sum_{i=1}^n X_i\right| > \varepsilon n\right) \leq \frac{D_\theta \sum_{i=1}^n X_i}{\varepsilon^2 n^2} = \frac{\theta}{\varepsilon^2 n}$

$P(|\overline{X} - \theta| \leq \varepsilon) \geq 1 - \frac{\theta}{\varepsilon^2 n} \geq \alpha$ $\quad \varepsilon = \sqrt{\frac{\theta}{(1-\alpha)n}}$

$$P_\theta\left(|\overline{X} - \theta| \leq \sqrt{\frac{\theta}{n(1-\alpha)}}\right) \geq \alpha$$

т.е. с вер-ю α $\quad \theta - \sqrt{\dfrac{\theta}{n(1-\alpha)}} \leq \bar{X} \leq \theta + \sqrt{\dfrac{\theta}{n(1-\alpha)}}$

Решаем относительно θ:

$$\sqrt{\bar{X} + \frac{1}{4(1-\alpha)n}} - \sqrt{\frac{1}{\alpha(1-\alpha)n}} \leq \theta \leq \sqrt{\bar{X} + \frac{1}{4(1-\alpha)n}} + \sqrt{\frac{1}{\alpha(1-\alpha)n}}$$

доверительный интервал

Обычно в задачах $\alpha = 0.95$ (она же 0.05)

первая магическая константа

Опр Пусть P - распределение на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$F(x)$ — ф.р., $p \in (0, 1)$, тогда

p - квантилью расп. P наз-ся

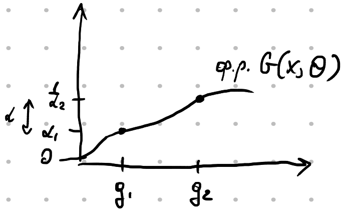
$$u_p = \min\{x : F(x) \geq p\}$$

Центральный интервал

(метод центральной ф-ции)

Пусть $X = (X_1 \ldots X_n)$ — выборка из $P \in \{P_\theta \mid \theta \in \Theta\}$

Рассмотрим $G(X, \theta)$ (центральная ф-ция), распр. которой не зависит от θ



Пусть $\alpha_1, \alpha_2 \in (0,1)$ $\quad \alpha_2 - \alpha_1 = \alpha$

g_1, g_2 — α_1, α_2 - квантили распр $G(X, \theta)$

Тогда $S(x) = \{\theta \in \Theta \mid g_1 \leq G(X, \theta) \leq g_2\}$ — дов. область ур. дов. α

$$P_\theta(\theta \in S(X)) = P_\theta(g_1 \le G(X,\theta) \le g_2) = F(g_2) - F(g_1) = \alpha$$

для разр. ф-ций надо быть аккуратным

Пример $X_1 \dots X_n \sim N(\theta, \sigma^2)$ σ^2 изв.

Построить центральной интервал

$$\overline{X} \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

$$\overline{X} - \theta \sim N\left(0, \frac{\sigma^2}{n}\right)$$

$$G(X, \theta) = \sqrt{n}\,\frac{\overline{X} - \theta}{\sigma} \sim N(0,1)$$

$z_\alpha :=$ α-квантиль $N(0,1)$ $\quad z_{\frac{1+\alpha}{2}} = -z_{\frac{1-\alpha}{2}}$

Тогда с вер-ю α $\sqrt{n}\,\frac{|\overline{X} - \theta|}{\sigma} \leq z_{\frac{1+\alpha}{2}}$

Выразим θ

$$P\left(\theta \in \left(\bar{X} - \frac{z_{\frac{1+\alpha}{2}}\sigma}{\sqrt{n}} ; \bar{X} + \frac{z_{\frac{1+\alpha}{2}}\sigma}{\sqrt{n}}\right)\right) \geq \alpha$$

Пусть $\mathcal{P} = \{ P_\theta \mid \theta \in \Theta \}$, θ — пар-р сдвига $[\, p_\theta(x) = p_1(x-\theta) \,]$

Тогда $G(X, \theta) = \bar{X} - \theta$ — центральная ф-ция

Если F_θ — непр, то

$$G(X, \theta) = - \sum_{i=1}^{n} \ln F_\theta(X_i) \sim \Gamma(1, n) \text{ — центр. ф-ция}$$

$$F_\xi(\xi) = U[0, 1]$$

$$F_\theta(X_1) \sim U[0, 1]$$

$$-\ln F_\theta(X_1) \sim Exp(1)$$

$$P_\theta(-\ln F_\theta(X_1) \geq y) = P_\theta(F_\theta(X_1) \leq e^{-y}) = e^{-y}$$

Асимптотический дов. инт. Вальда

Пусть $\hat\theta$ – а.н.о. θ с а.д. $\sigma^2(\theta)$

$$\sqrt{n}\,\frac{\hat\theta-\theta}{\sigma} \xrightarrow{d_\theta} N(0,1)$$

↓ т. Александрова

$$P_\theta\left(\sqrt{n}\,\frac{|\hat\theta-\theta|}{\sigma} \le z_{\frac{1+\alpha}{2}}\right) \longrightarrow \alpha$$

С вероятностью $\to \alpha$

$$\sqrt{n}\,\frac{|\hat\theta-\theta|}{\sigma(\theta)} \le z_{\frac{1+\alpha}{2}}$$

Остаётся выразить θ, но $\sigma(\theta)$ может иметь сложную зависимость от θ

Возьмём $\hat\sigma(\theta)$ – сост. оц. $\sigma(\theta)$

$$\sqrt{n}\,\frac{\hat\theta-\theta}{\hat\sigma(\theta)} = \sqrt{n}\,\frac{\hat\theta-\theta}{\hat\sigma(\theta)}\cdot\frac{\sigma(\theta)}{\sigma(\theta)} = \underbrace{\sqrt{n}\,\frac{\hat\theta-\theta}{\sigma(\theta)}}_{\xrightarrow{d} N(0,1)}\ \underbrace{\frac{\sigma(\theta)}{\hat\sigma(\theta)}}_{\xrightarrow{P} 1} \xrightarrow{d} N(0,1)$$

т. Слуцкого

Интервал $\left(\bar{X} \pm \frac{z_{\frac{1+\alpha}{2}}\,\hat\sigma}{\sqrt{n}}\right)$

Какую взять $\hat\sigma$?

$\hat\sigma(\theta) = \sigma(\hat\theta)$, где $\hat\theta$ — а.н.о.

тогда $\sigma(\hat\theta)$ — а.н.о. $\Rightarrow$ состоятельная

Пример $X_1 \ldots X_n \sim N(\theta, \sigma^2)$, σ – неизв.

ЦПТ: $\overline{X}$ – а.н.о. θ с а.д. σ^2

S – сост. оценка σ (ДЗ)

Получим $\left(\overline{X} \pm \frac{z_{\frac{1+\alpha}{2}} S}{\sqrt{n}} \right)$

Пример $X_1 \ldots X_n \sim Pois(\theta)$

ЦПТ $\overline{X}$ — а.н.о. θ с ас.д. θ

Получаем $\left(\overline{X} \pm \dfrac{z_{\frac{1+\alpha}{2}}\sqrt{\overline{X}}}{\sqrt{n}}\right)$

Пример $X_1 \dots X_n \sim U[0, \theta]$

ЦПТ $2\overline{X}$ – а.н.о. θ с ас. д. $\sigma^2(\theta) = \frac{\theta^2}{3}$

Получаем интервал $\left(2\overline{X} \pm \frac{z_{\frac{1+\alpha}{2}} \frac{2\overline{X}}{\sqrt{3}}}{\sqrt{n}} \right)$

Более хороший интервал в ДЗ / На семинаре

Точные доверительные интервалы для нормальной модели

Пусть $X = (X_1 \dots X_n)$ — выборка из распр $N(a, \sigma^2)$

1. Дов. интервал для a, если σ известно

$$\left(\overline{X} \pm \frac{z_{\frac{1+\alpha}{2}} \sigma}{\sqrt{n}} \right)$$

хи-квадрат

2. Дов. интервал для σ, если a известно

$$\frac{X_i - a}{\sigma} \sim N(0,1) \qquad \underbrace{\sum_{i=1}^{n}\left(\frac{X_i - a}{\sigma}\right)^2}_{= G(X,\sigma)} \sim \chi^2_n \text{ — ц.ф.}$$

хи-квадрат

$$P\left(\chi^2_{n,\frac{1-\alpha}{2}} \le \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - a)^2 \le \chi^2_{n,\frac{1+\alpha}{2}}\right) = \alpha$$

Получаем $\left(\sqrt{\dfrac{\sum (X_i - a)^2}{\chi^2_{n,\frac{1+\alpha}{2}}}}\,;\ \sqrt{\dfrac{\sum (X_i - a)^2}{\chi^2_{n,\frac{1-\alpha}{2}}}}\right)$

Стьюдент

3. Дов. интервал для a, если σ не изв.

Теорема. Если $X_1 \dots X_n \sim N(\theta, \sigma^2)$, то

1) $\overline{X}$, S^2 нез-мы

2) $\frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$

3) $\xi = \sqrt{n-1}\,\frac{\overline{X}-\theta}{S} \sim T_{n-1}$ (стьюдента)

▲ $k = n-1$; $y = \frac{nS^2}{\sigma^2}$;

$$\frac{\xi}{\sqrt{y/k}} = \sqrt{n}\,\frac{\overline{X}-\theta}{\sigma}\sqrt{n-1}\,\frac{\sigma}{\sqrt{n}\,S}$$

// доказали какую-то часть

По 3) $G(X, \theta) = \sqrt{n-1}\,\dfrac{\bar{X} - \theta}{S} \sim T_{n-1}$

с вер-ю α $\sqrt{n-1}\,\dfrac{|\bar{X} - \theta|}{S} \le T_{n-1,\,\frac{1+\alpha}{2}}$

4. Дов. интервал для σ, если a неизв.

по 2) $G(X, \sigma) = \dfrac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$

Получаем интервал $\left(\sqrt{\dfrac{nS^2}{\chi^2_{n-1, \frac{1+\alpha}{2}}}}, \sqrt{\dfrac{nS^2}{\chi^2_{n-1, \frac{1-\alpha}{2}}}} \right)$

5 Дов. область для (a, σ) — ДЗ

Построить дов. инт.

Дек. произв. по 1)

Теорема (об ортогональном разложении Гаусовского вектора)

Пусть $\xi \sim N(a, \sigma^2 I_n)$

$\mathbb{R}^n = L_1 \oplus \ldots \oplus L_k$ — разложение в прям. сумму ортог. подпр-в.

$y_j = \mathrm{proj}_{L_j} \xi$ — проекция ξ на L_j

Тогда

- $y_1 \ldots y_k$ — нез. в совок.
- $E y_j = \mathrm{proj}_{L_j} a$
- $\frac{1}{\sigma^2} \| y_j - E y_j \|^2 \sim \chi^2_{\dim L_j}$

▲ Рассмотрим ОНБ в $\mathbb{R}^n$

$d_j = \dim L_j$

$\underbrace{e_1 \ldots e_{d_1}}_{\text{базис в } L_1} \ \ldots \ \underbrace{\ldots e_{d_k}}_{\text{базис в } L_k}$

I_j – индексы базисных векторов в L_j

$B = (e_1 \ldots e_n)$ – ортог. матрица $n \times n$

$\zeta_i = \langle e_i, \xi \rangle = e_i^T \xi$ – длина проекция ξ на e_i

$$\zeta = \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix} = \begin{pmatrix} e_1^T \xi \\ \vdots \\ e_n^T \xi \end{pmatrix} = B^T \xi$$

$$\xi = \sum_{i=1}^{n} \langle e_i, \xi \rangle e_i = \sum_{i=1}^{n} \zeta_i e_i = (e_1 \ldots e_n) \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix} = B \zeta$$

$$E \zeta = E B^T \xi = B^T E \xi = B^T a$$

$$D \zeta = B^T D \xi B = B^T \sigma I_n B = \sigma I_n$$

Вывод ζ — гауссовский вектор с нез. компонентами

$$y_j = \operatorname{proj}_{L_j} \xi = \sum_{i \in I_j} \langle \xi, e_i \rangle e_i = \sum_{i \in I_j} \zeta_i e_i$$

Компоненты вектора ζ не пересекаются в разных y_j

$\Rightarrow$ $y_1 \dots y_k$ — нез. в совок.

$$E\, y_j = \sum_{i \in I_j} \langle E\xi, e_i \rangle e_i = \mathrm{proj}_{L_j} a$$

$$\frac{1}{\sigma^2} \| y_j - E y_j \|^2 = \frac{1}{\sigma^2} \Big\| \sum_{i \in I_j} \langle \xi - a, e_i \rangle e_i \Big\|^2 = \Big| \text{базис ортонорм} \Big| =$$

$$= \frac{1}{\sigma^2} \sum_{i \in I_j} (\xi_i - E\xi_i)^2 = \sum_{i \in I_j} \left(\frac{\xi_i - E\xi_i}{\sigma} \right)^2 \sim \chi^2_{\dim L_j}$$

Док-во п.1 и п.2:

1) $\overline{X}$, S^2 — нез $\quad L = \left\langle \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\rangle$

$$\mathbb{R}^n = L \oplus L^{\perp}$$

$$\mathrm{proj}_L X = \underset{c \in \mathbb{R}}{\mathrm{argmin}} \left\| X - \begin{pmatrix} c \\ \vdots \\ c \end{pmatrix} \right\|^2 = \underset{c}{\mathrm{argmin}} \sum_{i=1}^{n} (X_i - c)^2 =$$

$$= \begin{pmatrix} \bar{X}_1 \\ \vdots \\ \bar{X}_n \end{pmatrix}$$

$$\mathrm{proj}_{L^{\perp}} X = X - \mathrm{proj}_L X = \begin{pmatrix} X_1 - \bar{X}_1 \\ \vdots \\ X_n - \bar{X}_n \end{pmatrix}$$

$\Rightarrow$ $\bar{X}$ и S^2 нез-мы как ф-ции от $\mathrm{proj}_L X$ и $\mathrm{proj}_{L^{\perp}} X$

n. 2)

$$\frac{1}{\sigma^2} \| \operatorname{proj}_{L^\perp} X - \underset{\substack{\parallel \\ 0}}{E \operatorname{proj}_{L^\perp} X} \|^2 = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{n S^2}{\sigma^2} \sim \chi^2_{\underbrace{n-1}_{\dim L^\perp}}$$