INTERVALLI DI CONFIDENZA - TEST DI IPOTESI

Note Title 28/05/2018

$X_1 \_\_ X_n$ campione la cui distribuzione è condizionata da un parametro $\theta \in \mathbb{R}$.

Se $L_i = l_i(X_1 \_\_ X_n)$

$L_s = l_s(X_1 \_\_ X_n)$

sono t.c. $\quad \mathbb{P}(\theta \in (L_i, L_s)) \geq 1-\alpha$,

allora l'intervallo (L_i, L_s) si indica intervallo di confidenza di livello $1-\alpha$

$(L_i, +\infty)$

$(-\infty, L_s)$

$X_1 \_\_ X_n$ campione gaussiano $\mathbb{P}_{X_i} = N(\mu, \sigma^2)$ con σ^2 nota

$$\Rightarrow \left(\overline{X} - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \overline{X} + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$$ è un intervallo di confidenza di livello $1-\alpha$

— o —

$\mathbb{P}_Z = N(0,1) \qquad \mathbb{P}(Z \leq t) = \Phi(t) = 1-\alpha \quad$ sse

$t = z_{1-\alpha}$

$X_1 \_\_ X_n$ campione gaussiano $\mathbb{P}_{X_i} = N(\mu, \sigma^2)$

$\dfrac{(\overline{X}-\mu)}{\sigma/\sqrt{n}}$ ha distribuzione $N(0,1)$

$$1-\alpha = \mathbb{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_{1-\alpha}\right) = \mathbb{P}\left(\overline{X}-\mu \leq \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right) =$$

$$= \mathbb{P}\left(\mu \geq \overline{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right)$$

$$\Rightarrow (L_i, +\infty) = \left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}, +\infty\right)$$

$$\mathbb{P}(Z \geq t) = 1-\alpha$$

$$1-\alpha = 1-\mathbb{P}(Z \leq t) = 1-\Phi(t) \iff \Phi(t) = \alpha \quad t = z_\alpha$$

$$1-\alpha = \mathbb{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq z_\alpha\right) = \mathbb{P}\left(\overline{X}-\mu \geq \frac{\sigma}{\sqrt{n}} z_\alpha\right) =$$

$$= \mathbb{P}\left(\mu \leq \overline{X} - \frac{\sigma}{\sqrt{n}} z_\alpha\right) = \mathbb{P}\left(\mu \leq \overline{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right)$$

$$(-\infty, L_S) = \left(-\infty, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right)$$

$X_1 \ldots X_n$ campione gaussiano $\mathbb{P}_{X_i} = N(\mu, \sigma^2)$ con μ e σ^2 entrambi incogniti

$$T := \frac{(\overline{X}-\mu)\sqrt{n}}{S} \qquad \text{so che} \qquad \mathbb{P}_T = t(n-1)$$

$$\mathbb{P}(|T| \leq t) = 1-\alpha$$

$$1-\alpha = \mathbb{P}(-t \leq T \leq t) = \mathbb{P}(T \leq t) - \mathbb{P}(T \leq -t) =$$

$$= F_T(t) - F_T(-t)$$

$$= F_T(t) - (1 - F_T(t)) = 2F_T(t) - 1$$

$$2F_T(t) = 2-\alpha \qquad F_T(t) = 1-\frac{\alpha}{2} \iff t = t_{n-1, 1-\frac{\alpha}{2}}$$

$$1-\alpha = \mathbb{P}(|T| \leq t_{n-1,1-\frac{\alpha}{2}}) = \mathbb{P}\left(\frac{|(\overline{X}-\mu)\sqrt{n}|}{S} \leq t_{n-1,1-\frac{\alpha}{2}}\right) =$$

$$\mathbb{P}\left(|\overline{X}-\mu| \leq \frac{S}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right) =$$

$$= \mathbb{P}\left(-\frac{S}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}} \leq \mu - \overline{X} \leq \frac{S}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right)$$

$$= \mathbb{P}\left(\overline{X} - \frac{S}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}} \leq \mu \leq \overline{X} + \frac{S}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right)$$

$$(L_i, L_s) = \left(\overline{X} - \frac{S}{\sqrt{n}} t_{n-1, 1-\frac{\alpha}{2}} , \overline{X} + \frac{S}{\sqrt{n}} t_{n-1, 1-\frac{\alpha}{2}} \right)$$

— ∘ —

$$1-\alpha = \mathbb{P}(T \le t_{n-1, 1-\alpha}) = \mathbb{P}\left(\frac{(\overline{X}-\mu)\sqrt{n}}{S} \le t_{n-1,1-\alpha} \right) =$$

$$= \mathbb{P}\left(\overline{X} - \mu \le \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \right) = \mathbb{P}\left(\mu \ge \overline{X} - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \right)$$

$$(L_i, +\infty) = \left(\overline{X} - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} , +\infty \right)$$

— ∘ —

$$1-\alpha = \mathbb{P}(T \ge t) = 1 - \mathbb{P}(T \le t) = 1 - F_T(t)$$

$$\Leftrightarrow \quad F_T(t) = \alpha \quad \Leftrightarrow \quad t = t_{n-1,\alpha} = -t_{n-1,1-\alpha}$$

$$1-\alpha = \mathbb{P}(T \ge -t_{n-1,1-\alpha}) = \mathbb{P}\left(\frac{(\overline{X}-\mu)\sqrt{n}}{S} \ge -t_{n-1,1-\alpha} \right)$$

$$= \mathbb{P}\left(\overline{X} - \mu \ge -\frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \right)$$

$$= \mathbb{P}\left(\mu \le \overline{X} + \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \right)$$

$$(-\infty, L_s) = \left(-\infty, \overline{X} + \frac{S}{\sqrt{n}} t_{n-1,1-\alpha} \right)$$

— ∘ —

$X_1 \_\_ X_n$ campione gaussiano $\mathbb{P}_{X_i} = N(\mu, \sigma^2)$ con μ e σ^2 entrambe incognite

$$V := \frac{(n-1)S^2}{\sigma^2} \qquad \mathbb{P}_V = \chi^2_{n-1} = \Gamma\left(\frac{n-1}{2}, \frac{1}{2} \right)$$

$$\mathbb{P}\left(\chi^2_{n-1,\frac{\alpha}{2}} \le V \le \chi^2_{n-1,1-\frac{\alpha}{2}} \right) \overset{?}{=} 1-\alpha$$

$$\mathbb{P}\left(\chi^2_{n-1,\frac{\alpha}{2}} \le V \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right) = \mathbb{P}\left(V \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right) -$$

$$-\mathbb{P}\left(V \le \chi^2_{n-1,\frac{\alpha}{2}}\right) = F_V\left(\chi^2_{n-1,1-\frac{\alpha}{2}}\right) - F_V\left(\chi^2_{n-1,\frac{\alpha}{2}}\right)$$

$$= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1-\alpha$$

$$1-\alpha = \mathbb{P}\left(\chi^2_{n-1,\frac{\alpha}{2}} \le V \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right) =$$

$1-\frac{\alpha}{2} > \frac{\alpha}{2}$

$\alpha < 1$

$$= \mathbb{P}\left(\chi^2_{n-1,\frac{\alpha}{2}} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right)$$

$$= \mathbb{P}\left(\frac{1}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \le \frac{\sigma^2}{(n-1)S^2} \le \frac{1}{\chi^2_{n-1,\frac{\alpha}{2}}}\right) =$$

$$= \mathbb{P}\left(\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right)$$

$$\Rightarrow (L_i, L_s) = \left(\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right)$$

— o —

$$1-\alpha = \mathbb{P}(V \le t) = F_V(t) \iff t = \chi^2_{n-1,1-\alpha}$$

$$1-\alpha = \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} \le \chi^2_{n-1,1-\alpha}\right) = \mathbb{P}\left(\frac{\sigma^2}{(n-1)S^2} \ge \frac{1}{\chi^2_{n-1,1-\alpha}}\right) =$$

$$= \mathbb{P}\left(\sigma^2 \ge \frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha}}\right)$$

$$(L_i, +\infty) = \left(\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha}}, +\infty\right)$$

— o —

$$1-\alpha = \mathbb{P}(V \geq t) = 1 - \mathbb{P}(V \leq t) = 1 - F_V(t) \iff$$

$$F_V(t) = \alpha \iff t = \chi^2_{n-1,\alpha}$$

$$1-\alpha = \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} \geq \chi^2_{n-1,\alpha}\right) = \mathbb{P}\left(\frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{\chi^2_{n-1,\alpha}}\right)$$

$$= \mathbb{P}\left(\sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{n-1,\alpha}}\right)$$

$$(\cancel{-\infty}\ 0, L_S) = \left(\cancel{-\infty}\ 0, \frac{(n-1)S^2}{\chi^2_{n-1,\alpha}}\right)$$

TEST DI IPOTESI

IPOTESI NULLA H_0

1. L'ipotesi è corretta e in base ai dati io l'accetto
2. L'ipotesi è corretta ma in base ai dati io la rifiuto.

ERRORE DI 1ª SPECIE

3. L'ipotesi è errata e in base ai dati io la rifiuto
4. L'ipotesi è errata ma in base ai dati io la accetto

ERRORE DI 2ª SPECIE

— o —

Lancio una monetina e voglio testare l'ipotesi nulla
H_0 : la monetina non è truccata.

x_1 ———— x_n $\qquad x_i = \begin{cases} 1 & \text{se ottengo testa al lancio i-esimo} \\ 0 & \text{se ottengo croce al lancio i-esimo} \end{cases}$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

X_1 ———— X_n $\qquad \mathbb{P}_{X_i} = B(p)$ $\qquad H_0 : p = \frac{1}{2}$

$$\mathbb{E}[\bar{X}] = p, \qquad Var[\bar{X}] = \frac{p(1-p)}{n}$$

Accetto H_0 se $\left| \bar{x} - \frac{1}{2} \right| < \varepsilon$

$$\alpha = \mathbb{P}\left(\left| \bar{X} - \frac{1}{2} \right| \geq \varepsilon \;\middle|\; p = \frac{1}{2} \right) \qquad \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

X e Y v.a. indipendenti
$\mathbb{P}_X = B(n, p) \qquad \mathbb{P}_Y = B(k, p)$
$\Rightarrow \mathbb{P}_{X+Y} = B(n+k, p)$

$$\mathbb{P}(X=j) = \binom{n}{j} p^j (1-p)^{n-j} \qquad \forall j = 0, \ldots n$$
$$\mathbb{P}(Y=s) = \binom{k}{s} p^s (1-p)^{k-s} \qquad \forall s = 0, \ldots k$$

$$X+Y(\Omega) = \{0, \_\_\_ n+k\}$$

$$\ell \in \{0, \dots, n+k\} \quad \mathbb{P}(X+Y=\ell) = \sum_{r=0}^{k} \mathbb{P}(X+Y=\ell, \; Y=r)$$

$$= \sum_{r=0}^{k} \mathbb{P}(X=\ell-r, \; Y=r) = \sum_{r=0}^{k} \mathbb{P}(X=\ell-r)\mathbb{P}(Y=r)$$

$$= \sum_{r=0}^{k} \binom{n}{\ell-r} p^{\ell-\cancel{r}} (1-p)^{n-(\ell-\cancel{r})} \binom{k}{r} \cancel{p^{r}} (1-p)^{k-\cancel{r}}$$

$$= p^{\ell}(1-p)^{n+k-\ell} \underbrace{\sum_{r=0}^{k} \binom{n}{\ell-r}\binom{k}{r}}_{\binom{n+k}{\ell}}$$

$\sum_{i=1}^{n} X_i$ ha distribuzione $B(n,p)$

$$\alpha = \mathbb{P}\left(\left| \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{2} \right| \geq \varepsilon \;\middle|\; p=\frac{1}{2} \right) =$$

$$= \mathbb{P}\left(\left| \sum_{i=1}^{n} X_i - \frac{n}{2} \right| \geq n\varepsilon \;\middle|\; p=\frac{1}{2} \right)$$

$$Y := \sum_{i=1}^{n} X_i \qquad \mathbb{P}_Y = B(n,p) = B\left(n, \frac{1}{2}\right)$$

$$\alpha = \mathbb{P}\left(\left| Y - \frac{n}{2} \right| \geq n\varepsilon \right)$$

$$= \mathbb{P}\left(Y \geq \frac{n}{2} + n\varepsilon \right) + \mathbb{P}\left(Y \leq \frac{n}{2} - n\varepsilon \right)$$

Errore di 2^a specie: l'ipotesi H_0 è falsa ma in base ai dati io l'accetto

$$|\bar{x} - \tfrac{1}{2}| < \varepsilon$$

$$\beta(p_0) = \mathbb{P}\left(|\bar{X} - \tfrac{1}{2}| < \varepsilon \mid p = p_0 \right) \qquad p_0 \neq \frac{1}{2}$$

$Y := \sum_{i=1}^{n} X_i$ ha distribuzione $\mathcal{B}(n, p_0)$

$$\beta(p_0) = \mathbb{P}\left(\left| n\overline{X} - \frac{n}{2} \right| < n\varepsilon \;\middle|\; p = p_0 \right) =$$

$$= \mathbb{P}\left(\left| Y - \frac{n}{2} \right| < n\varepsilon \right) = \qquad p_0 : \left(0, \tfrac{1}{2}\right) \cup \left(\tfrac{1}{2}, 1\right) \to [0,1]$$

$$= \mathbb{P}\left(\frac{n}{2} - n\varepsilon < Y < \frac{n}{2} + n\varepsilon \right)$$

$X_1 \_\_\_ X_n$ campione statistico

H_0 ipotesi nulla

TEST PARAMETRICO $\quad \Theta \subseteq \mathbb{R}^k$

$H_0 : \quad \theta \in \Theta_0 \quad$ con $\Theta_0 \subseteq \Theta \qquad H_A : \theta \in \Theta \setminus \Theta_0$

TEST NON PARAMETRICO $\quad \mathcal{F}$

$H_0 : \quad F(x) \in \mathcal{F}_0 \qquad H_A : F(x) \in \mathcal{F} \setminus \mathcal{F}_0$

CRITERIO DI ACCETTAZIONE

$f(X_1 \_\_\_ X_n) \in \mathbb{R}$

STABILISCO una regione di accettazione $\mathcal{A} \subseteq \mathbb{R}$

considero $\mathcal{C} := \mathbb{R} \setminus \mathcal{A}$ regione di rifiuto

Accetto H_0 se $f(x_1 \_\_\_ x_n) \in \mathcal{A} \qquad f(X_1 \_ X_n) = \overline{X}$

Rifiuto H_0 se $f(x_1 \_\_\_ x_n) \notin \mathcal{A} \qquad \mathcal{A} = \left(\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon\right)$

PROBABILITÀ DI COMMETTERE ERRORE DI 1ª SPECIE

$\alpha = \mathbb{P}(\varphi(X_1 \_\_\_ X_n) \notin \mathcal{A} \mid H_0 \text{ vera})$

PROBABILITÀ DI COMMETTERE ERRORE DI 2ª SPECIE

$\beta = \mathbb{P}(\varphi(X_1 \_\_\_ X_n) \in \mathcal{A} \mid H_A \text{ vera})$

$(\mu, \sigma) \in \mathbb{R} \times (0, +\infty)$

$H_0 : \mu = 5$ $\quad\quad (\mu, \sigma) \in \{5\} \times (0, +\infty)$