

QUANTILI

Se Z è una v.e. con $\mathbb{P}_Z = N(0,1)$, allora la densità è

$$f_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) > 0 \quad \forall x \in \mathbb{R}$$

$$\Phi(t) = \int_{-\infty}^t f_0(x) dx \quad \Rightarrow \quad \Phi'(t) = f_0'(t) > 0 \quad \forall t \in \mathbb{R}$$

Di conseguenza la funzione $\Phi: t \in \mathbb{R} \mapsto \Phi(t) \in (0,1)$ è suriettiva e strettamente monotona crescente $\Rightarrow$ è invertibile:

$$\forall \alpha \in (0,1) \quad \exists! z = z_\alpha \in \mathbb{R} \text{ t.c. } \Phi(z_\alpha) = \alpha$$

Questo valore si chiama QUANTILE α LIVELLO α DELLA DISTRIBUZIONE $N(0,1)$.

$$\text{So che } \Phi(t) + \Phi(-t) = 1 \quad \text{cioè } \Phi(-t) = 1 - \Phi(t) \quad \forall t \in \mathbb{R}$$

$$t = z_\alpha \quad \Rightarrow \quad \Phi(t) = \alpha \quad \Rightarrow \quad \Phi(-t) = 1 - \alpha \quad \text{cioè}$$

$$\Phi(-z_\alpha) = 1 - \alpha$$

$$\text{cioè } z_{1-\alpha} = -z_\alpha \quad \forall \alpha \in (0,1).$$

$$\text{In particolare } z_{1/2} = 0$$

Poiché Φ è strettamente monotona crescente e

$$\Phi(0) = \frac{1}{2} \quad \Rightarrow \quad z_\alpha > 0 \quad \text{SSE } \alpha > \frac{1}{2}$$

— o —

Sia V una v.e. con $\mathbb{P}_V = \chi_n^2$ — la densità associata

$$f(x) = \begin{cases} \frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad \chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

Considero la legge $F_V(t) = \int_{-\infty}^t f(x) dx \Rightarrow$

$$F_V(t) = \begin{cases} 0 & t \leq 0 \\ \int_0^t \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2} dx & t > 0 \end{cases}$$

$$F'_V(t) = \frac{1}{2^{n/2} \Gamma(n/2)} t^{n/2-1} e^{-t/2} > 0 \quad \forall t > 0$$

$$F'_V(t) = 0 \quad \forall t < 0$$

$$F_V \Big|_{(0, +\infty)} : t \in (0, +\infty) \rightarrow F_V(t) \in (0, 1)$$

è suriettiva e strettamente monotona crescente $\Rightarrow$

è invertibile:

$$\forall \alpha \in (0, 1) \exists! t \text{ T.c. } F_V(t) = \alpha$$

Quanto t si indica $\chi_{n, \alpha}^2$ e si chiama

QUANTILE IN LIVELLO DELLA DISTRIBUZIONE DI PEARSON (o χ^2) A n GRADI DI LIBERTÀ.

— o —

Sia T una v.v. con $P_T = t(n)$ la densità

$$\text{associato è} \quad \tau_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{(n+1)}{2}} \quad \forall x \in \mathbb{R}$$

$$z_n(x) = \frac{1}{\Gamma(\frac{n}{2}) \sqrt{n\pi}} \left(1 + \frac{x^2}{n} \right) \quad \forall x \in \mathbb{R}$$

z_n è una funzione pari

$$z_n(x) > 0 \quad \forall x \in \mathbb{R}$$

Considero la legge $F_T: t \in \mathbb{R} \mapsto \int_{-\infty}^t z_n(x) dx \in (0, 1)$

È suriettiva

$F_T'(t) = z_n(t) > 0 \Rightarrow F_T$ è strettamente monotona crescente.

$\Rightarrow$ è invertibile:

$$\forall \alpha \in (0, 1) \quad \exists! t = t_{n, \alpha} \in \mathbb{R} \quad \text{T.c.} \quad F_T(t_{n, \alpha}) = \alpha$$

$t_{n, \alpha}$ si dice QUANTILE al LIVELLO α DELLA DISTRIBUZIONE t in STUDENT A n GRADI DI LIBERTÀ

$$\begin{aligned} \text{Poiché } z_n(x) \text{ è pari:} \quad & t_{n, 1-\alpha} = -t_{n, \alpha} \quad \forall \alpha \in (0, 1) \\ & t_{n, \frac{1}{2}} = 0 \\ & t_{n, \alpha} > 0 \quad \text{SSE} \quad \alpha \in \left(\frac{1}{2}, 1\right) \end{aligned}$$

STIMATORI DI MASSIMA VEROSIMILIANZA

$\rightarrow x_1, \dots, x_n$ dati numerici

$X_1, \dots, X_n$ i.i.d. con distribuzione campionaria avente densità

Supponiamo di conoscere la densità e meno di un qualche parametro (scalare o vettoriale) -

Se $f(x|\theta)$ è la densità associata alla distribuzione campionaria, la v.e. vettoriale $(X_1, \dots, X_n)$ ha densità

campionante, lo v.e. vettore $(X_1, \dots, X_n)$ ha densità

$$g(x_1, x_2, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta)$$

CASO DISCRETO

$$P(X_1=x_1, X_2=x_2, \dots, X_n=x_n) = \prod_{i=1}^n P(X_i=x_i) = \prod_{i=1}^n f(x_i | \theta)$$

CASO A.C.

$$P(x_i - \delta \leq X_i \leq x_i + \delta, \dots, x_n - \delta \leq X_n \leq x_n + \delta) =$$

$$= \prod_{i=1}^n P(x_i - \delta \leq X_i \leq x_i + \delta) = \prod_{i=1}^n \int_{x_i - \delta}^{x_i + \delta} f(s | \theta) ds \quad \text{--- } \textcircled{A}$$

Se f è continua $\Rightarrow \int_{x_i - \delta}^{x_i + \delta} f(s | \theta) ds \simeq 2\delta f(x_i | \theta)$

$$\textcircled{A} \simeq \prod_{i=1}^n 2\delta f(x_i | \theta) = (2\delta)^n \prod_{i=1}^n f(x_i | \theta)$$

$$\frac{\partial}{\partial \theta} \left(\prod_{i=1}^n f(x_i | \theta) \right)$$

$$h(\theta) = \ln \left(\prod_{i=1}^n f(x_i | \theta) \right)$$

$$\prod_{i=1}^n f(x_i | \theta) \in \max \text{ sse } h(\theta) \in \max$$

$$h(\theta) = \sum_{i=1}^n \ln f(x_i | \theta)$$

$$\bar{\theta} \text{ massimizza } h(\theta) \text{ sse } \bar{\theta} \text{ massimizza } \prod_{i=1}^n f(x_i | \theta)$$

DISTRIBUZIONE DI BERNOLLI

$$x_1, \dots, x_n \in \{0, 1\}$$

$$B(p)$$

$$P(X=1)=p$$

$$P(X=0)=1-p$$

$$B(p) \quad P(X=1)=p \quad P(X=0)=1-p$$

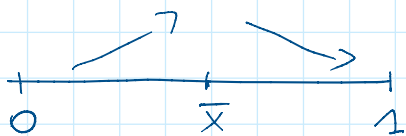
$$f(x_i, p) = \begin{cases} p & \text{se } x_i = 1 \\ 1-p & \text{se } x_i = 0 \end{cases}$$

$$\ln(f(x_i|p)) = \begin{cases} \ln(p) & \text{se } x_i = 1 \\ \ln(1-p) & \text{se } x_i = 0 \end{cases}$$

$$h(p) = \left(\sum_{i=1}^n x_i \right) \ln(p) + \left(n - \sum_{i=1}^n x_i \right) \ln(1-p)$$

$$\frac{dh}{dp}(p) = \frac{1}{p} n\bar{x} + (n - n\bar{x}) \frac{-1}{1-p} = n \frac{\bar{x}(1-p) - p(1-\bar{x})}{p(1-p)}$$

$$= \frac{n}{p(1-p)} (\bar{x} - p) \stackrel{>}{=} 0 \quad p - \bar{x} \stackrel{<}{=} 0 \quad p \stackrel{<}{=} \bar{x}$$



$p = \bar{x}$ massimizza $h(p)$.

DISTRIBUZIONE DI POISSON DI PARAMETRO $\lambda > 0$

$$f(x|\lambda) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & x \in \mathbb{N} \\ 0 & \text{altrimenti.} \end{cases}$$

$$x_1, \dots, x_n \in \mathbb{N}$$

$$h(\lambda) = \sum_{i=1}^n \ln(f(x_i|\lambda)) = \sum_{i=1}^n \ln\left(e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}\right) =$$

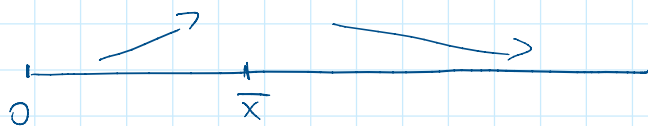
$$= \sum_{i=1}^n \left(\ln(e^{-\lambda}) + \ln(\lambda^{x_i}) - \ln(x_i!) \right) =$$

$$= \sum_{i=1}^n \left(-\lambda + x_i \ln(\lambda) - \ln(x_i!) \right) =$$

$$\frac{n}{\lambda} \dots$$

$$= -n\lambda + \ln(\lambda) n\bar{x} - \sum_{i=1}^n \ln(x_i!)$$

$$\frac{\partial \ell}{\partial \lambda} = -n + \frac{n\bar{x}}{\lambda} = \frac{n}{\lambda} (-\lambda + \bar{x}) \stackrel{SS}{=} 0 \quad \text{SS} \Rightarrow \lambda \leq \bar{x}$$



$\lambda = \bar{x}$ massimizza $\ell(\lambda)$ ovvero massimizza la
derivata congruente.

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

DISTRIBUZIONE GAUSSIANA

$$f(x_i | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

$$\ell(\mu, \sigma) = \sum_{i=1}^n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)\right) =$$

$$= \sum_{i=1}^n \ln((2\pi\sigma^2)^{-1/2}) + \frac{(x_i - \mu)^2}{2\sigma^2} =$$

$$= \sum_{i=1}^n -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\begin{cases} \frac{\partial \ell}{\partial \mu}(\mu, \sigma) = 0 \\ \frac{\partial \ell}{\partial \sigma}(\mu, \sigma) = 0 \end{cases} \quad \begin{cases} + \frac{1}{2\sigma^2} \sum_{i=1}^n 2(x_i - \mu) = 0 \\ -\frac{n}{2} \frac{4\pi\sigma}{2\pi\sigma^2} + \left(\sum_{i=1}^n (x_i - \mu)^2\right) \frac{1}{2} \frac{1}{\sigma^3} = 0 \end{cases}$$

$$\frac{\partial \ell}{\partial \mu} = 0 \quad \text{SS} \Rightarrow \sum_{i=1}^n (x_i - \mu) = 0 \quad \left(\sum_{i=1}^n x_i - \left(\sum_{i=1}^n \mu \right) \right) = 0$$

$$\frac{dh}{d\mu} = 0 \quad \text{SSE} \quad \sum_{i=1}^n (x_i - \mu) = 0 \quad \left(\sum_{i=1}^n x_i - \left(\sum_{i=1}^n 1 \right) \mu \right) = 0$$

$$n\bar{x} - n\mu = 0 \quad \boxed{\text{SSE} \quad \mu = \bar{x}}$$

$\mu = \bar{x}$

Scriviamo in $\frac{dh}{d\mu}$:

$$-\frac{n}{\mu} + \frac{1}{\mu^3} \sum_{i=1}^n (x_i - \bar{x})^2 = 0$$

$(n-1) s^2$

$$-n\mu^2 + (n-1)s^2 = 0$$

$$\text{SSE} \quad \boxed{\mu^2 = \frac{n-1}{n} s^2}$$

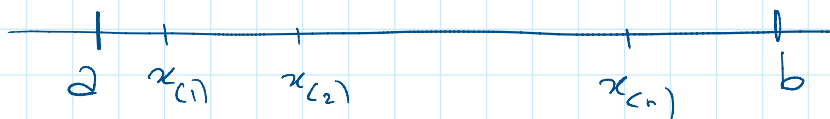
$$\mu^2 = \frac{n-1}{n} s^2$$

**DISTRIBUZIONE UNIFORME
SU UN INTERVALLO**

Se $I = [a, b]$ $f(x|a, b) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{altrimenti} \end{cases}$

$$g(x_1, \dots, x_n | a, b) = \prod_{i=1}^n f(x_i | a, b) =$$

$$= \begin{cases} 0 & \text{se } \exists i \text{ T.c. } x_i < a \vee x_i > b \\ \left(\frac{1}{b-a} \right)^n & \text{se } x_i \in [a, b] \quad \forall i = 1, \dots, n \end{cases}$$



g è massima SSE $a = \min \{x_i : i = 1, \dots, n\}$
 $b = \max \{x_i : i = 1, \dots, n\}$

$$A = \min \{X_1, \dots, X_n\}$$

$$B = \max \{X_1, \dots, X_n\}$$

— ~ —

INTERVALLI DI CONFIDENZA

Sia $X_1, \dots, X_n$ campione statistico e supponiamo che la sua distribuzione sia nota e mono di qualche

la sua distribuzione sia nota e messo di qualche parametro θ

Siano $L_i = l_i(X_1, \dots, X_n)$ due statistiche del
 $L_s = l_s(X_1, \dots, X_n)$ campione.

Dico che (L_i, L_s) è un intervallo di confidenza
di livello $1-\alpha$ per il parametro θ se
 $P(\theta \in (L_i, L_s)) \geq 1-\alpha$.

Dico che la semiretta $(L_i, +\infty)$ è un intervallo di
confidenza unilaterale superiore / ^{di livello $1-\alpha$} per il parametro θ se
 $P(\theta > L_i) \geq 1-\alpha$

Dico che la semiretta $(-\infty, L_s)$ è un intervallo di
confidenza unilaterale inferiore di livello $1-\alpha$ per il para-
metro θ se $P(\theta < L_s) \geq 1-\alpha$

DISTRIBUZIONE ESPONENZIALE

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \lambda e^{-\lambda x} & x > 0 \end{cases}$$

$$\exp(\lambda) = T(1, \lambda)$$

Se $X_1, \dots, X_n$ è un campione statistico con
distribuzione campionaria $\exp(\lambda)$.

$$\text{Considero } S_n = \sum_{i=1}^n X_i \quad P_{S_n} = T(n, \lambda)$$

Se $g(x)$ le densità di S_n :

$$g(x) = \begin{cases} 0 & x \leq 0 \\ \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} & x > 0 \end{cases}$$

Cio $Y = 2\lambda S_n$ $P_Y = \int h(x) dx$ $h(x) = \frac{1}{|2\lambda|} g\left(\frac{x}{2\lambda}\right)$

$$h(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{2\lambda} \frac{\lambda^n}{\Gamma(n)} \left(\frac{x}{2\lambda}\right)^{n-1} e^{-\frac{x}{2}} & x > 0 \end{cases} \Rightarrow$$

$$h(x) = \begin{cases} 0 & x \leq 0 \\ \frac{\left(\frac{1}{2}\right)^n x^{n-1}}{\Gamma(n)} e^{-x/2} & x > 0 \end{cases} \quad \Gamma\left(\frac{2n}{2}, \frac{1}{2}\right) = \chi_{2n}^2$$

$$Y = 2\lambda S_n = 2\lambda \sum_{i=1}^n X_i \quad - \quad \text{Per } \lambda \in (0, 1)$$

$$\mathbb{P}\left(\chi_{2n, \frac{\alpha}{2}}^2 \leq Y \leq \chi_{2n, 1-\frac{\alpha}{2}}^2\right) = \textcircled{A}$$

$$\{a \leq Y \leq b\} = \{Y \leq b\} \setminus \{Y < a\}$$

$$\textcircled{B} \quad \mathbb{P}(Y \leq \chi_{2n, 1-\frac{\alpha}{2}}^2) - \mathbb{P}(Y < \chi_{2n, \frac{\alpha}{2}}^2) =$$

$$= F_Y(\chi_{2n, 1-\frac{\alpha}{2}}^2) - F_Y(\chi_{2n, \frac{\alpha}{2}}^2) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

$$1 - \alpha = \mathbb{P}\left(\chi_{2n, \frac{\alpha}{2}}^2 \leq 2\lambda \sum_{i=1}^n X_i \leq \chi_{2n, 1-\frac{\alpha}{2}}^2\right) =$$

$$\begin{aligned}
 1-\alpha &= \mathbb{P} \left(\chi_{2n, \frac{\alpha}{2}}^2 \leq 2\lambda \sum_{i=1}^n X_i \leq \chi_{2n, 1-\frac{\alpha}{2}}^2 \right) = \\
 &= \mathbb{P} \left(\frac{1}{\chi_{2n, 1-\frac{\alpha}{2}}^2} \leq \frac{1}{2n\lambda \bar{X}} \leq \frac{1}{\chi_{2n, \frac{\alpha}{2}}^2} \right) \\
 &= \mathbb{P} \left(\frac{2n\bar{X}}{\chi_{2n, 1-\frac{\alpha}{2}}^2} \leq \frac{1}{\lambda} \leq \frac{2n\bar{X}}{\chi_{2n, \frac{\alpha}{2}}^2} \right)
 \end{aligned}$$

Abbiamo ottenuto un intervallo di confidenza al livello $1-\alpha$ per il parametro "valore atteso delle distribuzioni" cioè $\frac{1}{\lambda}$

$$(L_i, L_s) = \left(\frac{2n\bar{X}}{\chi_{2n, 1-\frac{\alpha}{2}}^2}, \frac{2n\bar{X}}{\chi_{2n, \frac{\alpha}{2}}^2} \right)$$

CAMPIONE GAUSSIANO A VARIANZA σ^2 NOTA

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

con μ incognito e σ^2 nota

sia Z una v.a.r. $\mathbb{P}_Z = \mathcal{N}(0, 1)$

$$\mathbb{P}(|Z| \leq z_{1-\frac{\alpha}{2}}) = \mathbb{P}(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}})$$

$$= \mathbb{P}(Z \leq z_{1-\frac{\alpha}{2}}) - \mathbb{P}(Z < -z_{1-\frac{\alpha}{2}})$$

$$= \Phi(z_{1-\frac{\alpha}{2}}) - \Phi(-z_{1-\frac{\alpha}{2}}) = \left(1 - \frac{\alpha}{2}\right) - \frac{\alpha}{2} = 1-\alpha$$

$$1-\alpha = \mathbb{P}(|Z| \leq z_{1-\frac{\alpha}{2}})$$

$X \sim \mathcal{N}(\mu, \sigma^2)$ $\mathbb{P}_X = \mathcal{N}(\mu, \sigma^2)$

$$\begin{aligned}
 & \overbrace{X_1 \dots X_n} \\
 & P_{X_i} = N(\mu, \sigma^2) \quad P_{\bar{X}} = N\left(\mu, \frac{\sigma^2}{n}\right) \\
 & \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} =: Z \quad \Rightarrow P_Z = N(0, 1) \\
 & Z = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma}
 \end{aligned}$$

$$1 - \alpha = P\left(\left|\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma}\right| \leq Z_{1 - \frac{\alpha}{2}}\right) =$$

$$= P\left(|\bar{X} - \mu| \leq \frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}}\right) =$$

$$P\left(-\frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}} \leq \mu - \bar{X} \leq \frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}}\right)$$

$$= P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}} \leq \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}}\right)$$

$$(L_i, L_s) = \left[\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{1 - \frac{\alpha}{2}}\right]$$

è un intervallo di confidenza di livello $1 - \alpha$ per il parametro μ .