martedì 26 novembre 2019 14:16

STATISTICA INFERENZIALE

$x_1 \_\_\_ x_n$ dati numerici

$(\Omega, \mathcal{E}, \mathbb{P})$ $X_1 \_\_\_ X_n$

$\exists\, \omega \in \Omega$ t.c. $x_1 = X_1(\omega)$

$\vdots$

$x_n = X_n(\omega)$

$X_1 \_\_\_ X_n$ identicamente distribuite $\Rightarrow$ $\mathbb{P}_{X_i}$ è comune

indipendenti

La comune distribuzione sia associata a una densità $f(x)$

e che abbia valore atteso e varianza finite.

$f(x|\theta)$

Sia $f(x)$ sia la comune densità di $X_1 \_\_\_ X_n$

$(X_1, X_2 \_\_\_ X_n)$ ha densità congiunta $f(x_1)f(x_2)\dots f(x_n)$

DEF (CAMPIONE STATISTICO)

Siano $X_1 \_\_\_ X_n$ n variabili aleatorie i.i.d su uno

stesso spazio probabilizzato $(\Omega, \mathcal{E}, \mathbb{P})$. Dico che $X_1 \_\_\_ X_n$

è un campione statistico di cardinalità n.

La comune distribuzione $\mathbb{P}_{X_i}$ $i=1 \_\_ n$ si dice DISTRIBUZIONE

CAMPIONARIA.

DEF (STATISTICA) Sia $X_1 \_\_\_ X_n$ un campione statistico

e sia $\varphi : \mathbb{R}^n \to \mathbb{R}$ una funzione boreliana.

Allora la v.a. $Y := \varphi \circ (X_1 \_\_\_ X_n) : \omega \in \Omega \to \varphi(X_1(\omega) \_\_ X_n(\omega)) \in \mathbb{R}$

si dice una STATISTICA DEL CAMPIONE.

Supponiamo di avere $(X_n)_{n\in\mathbb{N}}$ successione di v.a. i.i.d.

Supponiamo di avere, per ogni n una funzione borelliana $f_n : \mathbb{R}^n \to \mathbb{R}$ e di considerare la statistica $Y_n := f_n(X_1 \dots X_n)$

Se Y_n "serve a stimare un parametro λ che caratterizza la comune distribuzione delle X_i" dico che Y_n è uno

STIMATORE DEL PARAMETRO λ.

STIMATORE CONSISTENTE IN SENSO DEBOLE

La statistica Y_n si dice uno stimatore consistente in senso debole di λ se Y_n converge in probabilità alla v.a. costante λ.

STIMATORE CONSISTENTE IN SENSO FORTE

Se Y_n converge $\mathbb{P}$-q.c. alla v.a. costante λ.

STIMATORE CORRETTO (o NON DISTORTO)

Dico che Y_n è uno stimatore corretto di λ se $\mathbb{E}[Y_n] = \lambda$

N.B. $\mathbb{P}\left(|X - \mathbb{E}[X]| > \delta\right) \leq \dfrac{\mathrm{Var}[X]}{\delta^2} \quad \forall \delta > 0$

Supponiamo che Y_n sia uno stimatore corretto del parametro λ e che $\lim_{n\to\infty} \mathrm{Var}[Y_n] = 0$

Allora
$$\mathbb{P}\left(|Y_n - \lambda| > \delta\right) \leq \frac{\mathrm{Var}[Y_n]}{\delta^2} \longrightarrow 0 \quad \forall \delta > 0$$

cioè Y_n converge in probabilità alla v.a. costante λ

cioè è uno stimatore consistente in senso debole di λ.

MEDIA E VARIANZA CAMPIONARIA

Sia $X_1 \ldots X_n$ un campione statistico

Chiamo MEDIA (CAMPIONARIA) la v.a. $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$

Chiamo VARIANZA CAMPIONARIA la v.a.

$$S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

PROP Sia $X_1 \_\_ X_n$ un campione statistico la cui distribuzione campione ha valore atteso μ e varianza σ^2 (finit.)

Allora

$$\mathbb{E}[\bar{X}] = \mu \qquad Var[\bar{X}] = \frac{\sigma^2}{n}$$

$$\mathbb{E}[S^2] = \sigma^2$$

DIM.

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} \overbrace{\mathbb{E}[X_i]}^{\mu} = \frac{1}{n} n\mu = \mu$$

$$Var[\bar{X}] = Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} Var\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} \overbrace{Var[X_i]}^{\sigma^2}$$

$$= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}X_i + \bar{X}^2\right)$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - 2\bar{X}\underbrace{\sum_{i=1}^{n} X_i}_{n\bar{X}} + \underbrace{\sum_{i=1}^{n}\bar{X}^2}_{n\bar{X}^2}\right)$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$

$$(n-1)\,\mathbb{E}[S^2] = \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i^2\right) - n\bar{X}^2\right] =$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{n}(X_i - \mu + \mu)^2\right) - n(\bar{X} - \mu + \mu)^2\right] =$$

$$= \sum_{i=1}^{n}\mathbb{E}\left[(\underbrace{X_i - \mu}_{} + \mu)^2\right] - n\,\mathbb{E}\left[(\underbrace{\bar{X} - \mu}_{} + \mu)^2\right] =$$

$$= \sum_{i=1}^{n} \mathbb{E}\left[(X_i-\mu)^2 + 2\mu(X_i-\mu)+\mu^2\right] - n\,\mathbb{E}\left[(\bar{X}-\mu)^2+2\mu(\bar{X}-\mu)+\mu^2\right]$$

$$= \sum_{i=1}^{n} \mathbb{E}\left[(X_i-\mu)^2\right] + 2\mu \sum_{i=1}^{n} \mathbb{E}[X_i-\mu] + \sum_{i=1}^{n} \mu^2$$

$$- n\left(\mathbb{E}\left[(\bar{X}-\mu)^2\right] + 2\mu\,\mathbb{E}[\bar{X}-\mu] + \mu^2\right)$$

$$= \sum_{i=1}^{n} \sigma^2 - 2\mu\cdot 0 + n\mu^2$$

$$- n\left(\frac{\sigma^2}{n} + 2\mu\cdot 0 + \mu^2\right)$$

$$= n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2 = (n-1)\sigma^2 \qquad \Rightarrow \mathbb{E}[S^2] = \sigma^2$$

$$\mathbb{E}[\bar{X}] = \mu \qquad \mathbb{E}[\bar{X}-\mu] = \mathbb{E}[\bar{X}] - \mathbb{E}[\mu] = \mu - \mu = 0$$

TEOREMA CENTRALE DEL LIMITE

Sia $(X_n)_{n\in\mathbb{N}}$ successione di v.a. i.i.d. con valore atteso μ e varianza σ^2 finiti. Per ogni $n\in\mathbb{N}$ sia $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

Allora la v.a

$$Z_n := \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

converge in legge ad una v.a. gaussiana standard Z e la convergenza delle leggi è uniforme

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} \left| P(Z_n \le t) - \Phi(t) \right| = 0$$

N.B. $\quad \mathbb{P}_X = N(\mu,\sigma^2) \qquad Z := \frac{X-\mu}{\sigma} \quad \Rightarrow \mathbb{P}_Z = N(0,1)$

DISTRIBUZIONI $\Gamma(\alpha,\lambda)$ $\quad \alpha,\lambda>0$

Chiamiamo distribuzione Gamma di parametri $\alpha>0$ e $\lambda>0$

Chiamiamo distribuzione Gamma di parametri $\alpha>0$ e $\lambda>0$ la distribuzione assolutamente continua associata alla densità

$$f(x) = \begin{cases} \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x) & x>0 \\ 0 & x \leq 0 \end{cases}$$

dove Γ è la funzione Gamma di Eulero così definita

$$\Gamma(\alpha) := \int_0^{+\infty} x^{\alpha-1} e^{-x}\, dx \qquad \alpha>0$$

1) $f(x)$ è una densità : $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

$$\int_{\mathbb{R}} f(x)\,dx = 1$$

$$\int_{\mathbb{R}} f(x)\,dx = \int_0^{+\infty} \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x)\,dx \qquad \lambda x = y \quad x = \frac{y}{\lambda}$$

$$= \frac{\cancel{\lambda^{\alpha}}}{\Gamma(\alpha)} \int_0^{+\infty} \frac{y^{\alpha-1}}{\cancel{\lambda^{\alpha-1}}} e^{-y} \frac{1}{\cancel{\lambda}} dy = \frac{1}{\Gamma(\alpha)} \underbrace{\int_0^{+\infty} y^{\alpha-1} e^{-y}\,dy}_{\Gamma(\alpha)} =$$

$$= \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1 .$$

N.B. Se X è una v.a. con $P_X = \Gamma(\alpha,\lambda) \Rightarrow X$ è distribuita sulla semiretta $(0,+\infty)$

$$\Gamma(1) = \int_0^{+\infty} x^{1-1} e^{-x}\,dx = \int_0^{+\infty} e^{-x}\,dx = -e^{-x}\Big|_{x=0}^{x \to +\infty} = 1$$

La densità associata alla distribuzione $\Gamma(1,\lambda)$

$$f(x) = \begin{cases} \frac{\lambda}{1} x^{1-1} e^{-\lambda x} & x \geq 0 \\ 0 & x<0 \end{cases} = \begin{cases} \lambda e^{-\lambda x} & x>0 \\ 0 & x \leq 0 \end{cases}$$

$\Rightarrow \Gamma(1,\lambda) = Exp(\lambda) \quad \forall \lambda>0 .$

$\Rightarrow \Gamma(1,\lambda) = Exp(\lambda) \quad \forall \lambda > 0$.

$$\Gamma(\alpha+1) = \int_0^{+\infty} x^{(\alpha+1)-1} e^{-x}\, dx = \int_0^{+\infty} \underset{f}{x^{\alpha}}\, \underset{g'}{e^{-x}}\, dx =$$

$$= \cancel{x^{\alpha}(-e^{-x})\Big|_0^{+\infty}}^{=0} - \int_0^{+\infty} \alpha x^{\alpha-1}(-e^{-x})\, dx =$$

$$= \alpha \int_0^{+\infty} x^{\alpha-1} e^{-x}\, dx = \alpha\Gamma(\alpha) \qquad \Gamma(\alpha+1) = \alpha\Gamma(\alpha) \quad \forall \alpha > 0$$

Ma $\Gamma(1) = 1$

$\Gamma(2) = 1 \cdot \Gamma(1) = 1$

$\Gamma(3) = 2 \cdot \Gamma(2) = 2 \cdot 1$

$\Gamma(4) = 3 \cdot \Gamma(3) = 3 \cdot 2 \cdot 1$

$\vdots$

$\Gamma(n) = (n-1)!$

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} x^{-1/2} e^{-x}\, dx \qquad x = y^2 \qquad y = x^{1/2} \qquad dx = 2y\, dy$$

$$= \int_0^{+\infty} \frac{1}{y} e^{-y^2}\, 2y\, dy = 2\int_0^{+\infty} e^{-y^2}\, dy = \int_{\mathbb{R}} e^{-y^2}\, dy = \sqrt{\pi}$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt{\pi} \qquad n = 1$$

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi} \qquad n = 2$$

$$\Gamma\left(\frac{7}{2}\right) = \frac{5}{2}\Gamma\left(\frac{5}{2}\right) = \frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi} \qquad n = 3$$

$$\Gamma\left(\frac{2n+1}{2}\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi} \qquad \Gamma\Big(\underset{\frac{2n}{2}}{\underset{\shortparallel}{n}}\Big) = (n-1)!$$

PROP Sia X una v.a. con $\mathbb{P}_X = \Gamma(\alpha,\lambda)$. Allora

$$\mathbb{E}[X] = \frac{\alpha}{\lambda} \qquad \mathrm{Var}[X] = \frac{\alpha}{\lambda^2}$$

DIM

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f(x)\,dx = \int_0^{+\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx =$$

$$= \frac{\cancel{\lambda^\alpha}}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\cancel{\alpha}+1}} \underbrace{\int_0^{+\infty} \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{(\alpha+1)-1} e^{-\lambda x}\,dx}_{\int_{\mathbb{R}} \text{della densità associata a } \Gamma(\alpha+1,\lambda) \Rightarrow = 1}$$

$$= \frac{\alpha\Gamma(\alpha)}{\lambda\Gamma(\alpha)} \cdot 1 = \frac{\alpha}{\lambda}$$

$$\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 f(x)\,dx = \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^2 \cdot x^{\alpha-1} e^{-\lambda x}\,dx =$$

$$= \frac{\cancel{\lambda^\alpha}}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\lambda^{\cancel{\alpha}+2}} \underbrace{\int_0^{+\infty} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{(\alpha+2)-1} e^{-\lambda x}\,dx}_{\int_{\mathbb{R}} \text{della densità di } \Gamma(\alpha+2,\lambda) \Rightarrow \text{vale } 1}$$

$$= \frac{(\alpha+1)\alpha\cancel{\Gamma(\alpha)}}{\cancel{\Gamma(\alpha)}\lambda^2} = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 =$$

$$= \frac{\alpha}{\lambda^2}\left(\alpha + 1 - \alpha\right) = \frac{\alpha}{\lambda^2}$$

PROP. Siano X e Y v.a. indipendenti con

$$\mathbb{P}_X = \Gamma(\alpha,\lambda) \quad \text{e} \quad \mathbb{P}_Y = \Gamma(\beta,\lambda)$$

$\mathrm{Exp}(\lambda) = \Gamma(1,\lambda)$

Allora $\mathbb{P}_{X+Y} = \Gamma(\alpha+\beta,\lambda)$

DIM Se X e Y sono due v.a. indipendenti con distribuzione A.C.

$$\mathbb{P}_X = f(x)\,dx \qquad \mathbb{P}_Y = g(x)\,dx$$

$\mathbb{P}_X = f(x)\,dx \qquad \mathbb{P}_Y = g(x)\,dx$

$\Rightarrow \mathbb{P}_{X+Y} = h(x)\,dx \qquad h(x) = \int_{\mathbb{R}} \underbrace{f(y)\,g(x-y)}\,dy$

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & x>0 \\ 0 & x \le 0 \end{cases}$$

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

$f(y)\,g(x-y) \neq 0 \quad$ SSE $\begin{cases} y>0 \\ x-y>0 \end{cases}$

SSE $\begin{cases} y>0 \\ y<x \end{cases}$

1. $x \le 0 \Rightarrow \not{\emptyset} \Rightarrow h(x) = 0$
2. $x>0 \qquad y \in (0,x)$

$x>0 \Rightarrow h(x) = \int_0^x \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \cancel{e^{-\lambda y}} \cdot \frac{\lambda^\beta}{\Gamma(\beta)} (x-y)^{\beta-1} e^{-\lambda(x-\cancel{y})}\,dy$

$$= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda x} \int_0^x y^{\alpha-1}(x-y)^{\beta-1}\,dy \qquad \begin{array}{l} y = xt \\ dy = x\,dt \\ x-y = x(1-t) \end{array}$$

$$= \frac{\lambda^{\alpha+\beta} e^{-\lambda x}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha-1} t^{\alpha-1} x^{\beta-\cancel{1}} (1-t)^{\beta-1} \cancel{x}\,dt$$

$$= \underbrace{\frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} \left(\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,dt \right)}_{C} x^{\alpha+\beta-1} e^{-\lambda x}$$

$$h(x) = \begin{cases} C\, x^{\alpha+\beta-1} e^{-\lambda x} & x>0 \\ 0 & x \le 0 \end{cases}$$

$$\Rightarrow \int_{\mathbb{R}} h(x)\,dx = 1 \quad \text{SSE} \quad C = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha+\beta)}$$

$$\Rightarrow \mathbb{P}_{X+Y} = \Gamma(\alpha+\beta, \lambda)$$

DISTRIBUZIONE DI PEARSON A n GRADI DI LIBERTÀ O DISTRIBUZIONE χ^2_n

Se $\alpha = \frac{n}{2}$ e $\lambda = \frac{1}{2}$ la distribuzione $\Gamma(\alpha, \lambda) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ si dice DISTRIBUZIONE DI PEARSON A n gradi di libertà o anche distribuzione χ^2_n

OSS. Se X è una v.a. con $\mathbb{P}_X = \chi^2_n$

$$\Rightarrow E[X] = \frac{n/2}{1/2} = n$$

$$Var[X] = \frac{n/2}{1/4} = 2n$$

OSS. Se X e Y sono v.a. indipendenti con

$$\mathbb{P}_X = \chi^2_n \quad \text{e} \quad \mathbb{P}_Y = \chi^2_k$$

$$\Rightarrow \mathbb{P}_{X+Y} = \Gamma\left(\frac{n}{2} + \frac{k}{2}, \frac{1}{2}\right) \quad \text{cioè} \quad \mathbb{P}_{X+Y} = \chi^2_{n+k}$$

TEOREMA Sia Z una v.a. gaussiana standard: $\mathbb{P}_Z = N(0,1)$

Allora $\mathbb{P}_{Z^2} = \chi^2_1$.

DIM. Se $\mathbb{P}_X = f(x)\,dx \Rightarrow \mathbb{P}_{X^2} = g(x)\,dx$

$$\text{e} \quad g(x) = \begin{cases} \frac{1}{2\sqrt{x}}\left(f(\sqrt{x}) + f(-\sqrt{x})\right) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Applico con $X = Z$ e quindi con $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$

$$\Rightarrow f(-\sqrt{x}) = f(\sqrt{x}) \qquad \forall x \in \mathbb{R}$$

$$\Rightarrow f(-\sqrt{x}) = f(\sqrt{x}) \qquad \forall x \in \mathbb{R}$$

$$\Rightarrow \text{per } x > 0 \quad g(x) = \frac{1}{2\sqrt{x}}\left(f(\sqrt{x}) + f(\sqrt{x})\right) = \frac{f(\sqrt{x})}{\sqrt{x}}$$

$$= x^{-1/2} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\sqrt{x})^2\right) = x^{-1/2}\frac{1}{\sqrt{2\pi}} e^{-x/2} =$$

$$= \frac{1}{\sqrt{2\pi}} x^{\frac{1}{2}-1} e^{-x/2} \qquad \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \qquad \chi^2_1$$

PROP $X_1 \ldots X_n$ v.a. indipendenti. $\mathbb{P}_{X_i} = N(\mu_i, \sigma_i^2)$

$$\sum_{i=1}^{n}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \text{ ha distribuzione } \chi^2_n .$$

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

Allora $\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2$ ha distribuzione χ^2_n

TEOREMA Sia $X_1 \ldots X_n$ un campione gaussiano di valore atteso μ e varianza σ^2.

Allora $\overline{X}$ e S^2 sono v.a. indipendenti.

Per ogni $i = 1 \ldots n$ sia $Z_i = \frac{X_i - \mu}{\sigma}$ $\Rightarrow \mathbb{P}_{Z_i} = N(0,1)$

Allora $\overline{Z} = \frac{\overline{X} - \mu}{\sigma}$

Inoltre $\overline{Z}$ e $\sum_{i=1}^{n}(Z_i - \overline{Z})^2$ sono v.a. indipendenti

e la distribuzione di $\sum_{i=1}^{n}(Z_i - \overline{Z})^2$ è la distribuzione di Pearson a $(n-1)$ gradi di libertà.

N.B. $\mathbb{P}_{X_1} = \mathbb{P}_{X_2} = \ldots = \mathbb{P}_{X_n} = N(\mu, \sigma^2)$ $\quad X_1 \ldots X_n$ indip.

$$\mathbb{P}_{X_1 + X_2 \ldots + X_n} = N(n\mu, n\sigma^2)$$

$$\mathbb{P}_{\frac{X_1 + \ldots + X_n}{n}} = N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \mathbb{P}_{\overline{X}} = N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\mathbb{P}_{\underbrace{\frac{X_1+\dots+X_n}{n}}_{\bar{X}}} = N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \mathbb{P}_{\bar{X}} = N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\mathbb{P}_{\bar{Z}} = N\left(0, \frac{1}{n}\right)$$

$n=2$. $\quad Z_1 \quad Z_2 \qquad \mathbb{P}_{Z_1} = \mathbb{P}_{Z_2} = N(0,1) \qquad Z_1$ e Z_2 indip.

$$\bar{Z} = \frac{Z_1+Z_2}{2}$$

$$\sum_{i=1}^{2}(Z_i-\bar{Z})^2 = \left(Z_1 - \frac{Z_1+Z_2}{2}\right)^2 + \left(Z_2 - \frac{Z_1+Z_2}{2}\right)^2 =$$

$$= \left(\frac{Z_1-Z_2}{2}\right)^2 2 = \left(\frac{Z_1-Z_2}{\sqrt{2}}\right)^2$$

$$\frac{(Z_1-Z_2)^2}{4} \cdot 2 = \frac{(Z_1-Z_2)^2}{2} = \left(\frac{Z_1-Z_2}{\sqrt{2}}\right)^2$$

$$\mathbb{P}_{Z_i} = N(0,1) \qquad \mathbb{P}_{-Z_2}$$

$$\mathbb{P}_X = f(x)\,dx \qquad \mathbb{P}_{-X} = g(x)\,dx$$

$$\mathbb{P}_X = f(x)\,dx \qquad \mathbb{P}_{aX+b} = g(x)\,dx \quad g(x) = \frac{1}{|a|} f\left(\frac{x-b}{a}\right)$$

$$\mathbb{P}_{-X} = g(x)\,dx \quad g(x) = f(-x) = f(x)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) = f(-x) \qquad \Rightarrow \mathbb{P}_{-Z_2} = N(0,1)$$

$$\mathbb{P}_{Z_1-Z_2} = \mathbb{P}_{Z_1+(-Z_2)} = N(0+0, 1+1) = N(0,2)$$

$$\mathbb{P}_{\frac{Z_1-Z_2}{\sqrt{2}}} = N(0,1) \quad \Rightarrow \quad \mathbb{P}_{\left(\frac{Z_1-Z_2}{\sqrt{2}}\right)^2} = \chi_1^2$$

$$\sum_{i=1}^{n}(Z_i-\bar{Z})^2$$

$$Z_i = \frac{X_i-\mu}{} \qquad \bar{Z} = \frac{\bar{X}-\mu}{}$$

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad \bar{Z} = \frac{\bar{X} - \mu}{\sigma}$$

$$Z_i - \bar{Z} = \frac{X_i - \cancel{\mu}}{\sigma} - \frac{\bar{X} - \cancel{\mu}}{\sigma} = \frac{X_i - \bar{X}}{\sigma}$$

$$\sum_{i=1}^{n} (Z_i - \bar{Z})^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 =$$

$$\frac{(n-1) S^2}{\sigma^2}$$

Pongo $V := \frac{(n-1) S^2}{\sigma^2} = \frac{(n-1)}{\sigma^2} \frac{1}{n-1} \sum (X_i - \bar{X})^2$

So che $\mathbb{P}_V = \chi^2_{n-1}$

DISTRIBUZIONE t DI STUDENT A n GRADI DI LIBERTÀ

Si indica $t(n)$ ed è la distribuzione A.C. associata alla densità

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

$$\mathbb{E}[X] = 0 \qquad \mathrm{Var}[X] = \begin{cases} +\infty & n = 1,2 \\ \frac{n}{n-2} & n \geq 3 \end{cases}$$

Poiché τ_n è pari, se F è la legge associata allora $F(-t) = 1 - F(t) \quad \forall t \in \mathbb{R}$.

TEOREMA Siano Y e Z due v.a. indipendenti t.c.

$$\mathbb{P}_Z = N(0,1) \quad e \quad \mathbb{P}_Y = \chi^2_n$$

Sia $T := \frac{Z\sqrt{n}}{\sqrt{Y}}$. Allora $\mathbb{P}_T = t(n)$.

COROLLARIO Sia $X_1 \ldots X_n$ un campione gaussiano con valore atteso μ e varianza σ^2.

Allora la v.a.

$$T := \frac{(\overline{X}-\mu)\sqrt{n}}{S} \qquad S := \sqrt{S^2}$$

ha distribuzione t di Student con n-1 gradi di libertà

DIM $\mathbb{P}_{X_i} = N(\mu, \sigma^2)$

$$\mathbb{P}_{\overline{X}} = N\left(\mu, \frac{\sigma^2}{n}\right) \qquad Z := \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \Rightarrow \text{è gaussiana}$$

$$\mathbb{E}[Z] = \frac{1}{\sigma/\sqrt{n}} \mathbb{E}[\overline{X}-\mu] = \frac{\sqrt{n}}{\sigma}\left(\mathbb{E}[\overline{X}]-\mu\right) = 0$$

$$\mathrm{Var}[Z] = \frac{1}{\frac{\sigma^2}{n}} \mathrm{Var}[\overline{X}-\cancel{\mu}] = \frac{n}{\sigma^2} \cdot \frac{\sigma^2}{n} = 1$$

$$\Rightarrow \mathbb{P}_Z = N(0,1)$$

So che $V := \frac{(n-1)S^2}{\sigma^2}$ ha distribuzione χ^2_{n-1}

$$\Rightarrow \frac{Z\sqrt{n-1}}{\sqrt{V}} \text{ per il Teo ha distribuzione } t(n-1)$$

$$\frac{Z\sqrt{n-1}}{\sqrt{V}} = \frac{\frac{\overline{X}-\mu}{\cancel{\sigma}/\sqrt{n}} \cdot \cancel{\sqrt{n-1}}}{\cancel{\sqrt{n-1}}\ \frac{S}{\cancel{\sigma}}} = \frac{(\overline{X}-\mu)\sqrt{n}}{S}$$