Variabili aleatorie continue

martedì 19 settembre 2023 15:10

Distribuzione uniforme

Def Una v.a. X segue una legge uniforme in $[a,b] \subset \mathbb{R}$ se ha densità

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

In tal caso scriveremo $X \sim U([a,b])$

Oss
La distribuzione uniforme prevede la stessa probabilità di accadimento su tutti i valori di un intervallo $[a,b]$

Oss
$f(x)$ è effettivamente una densità. Infatti:

$$\int_{\mathbb{R}} f(x)\,dx = \int_a^b \frac{1}{b-a}\,dx = \frac{1}{b-a}[x]_a^b = \frac{\cancel{b-a}}{\cancel{b-a}} = 1$$

Oss
Calcoliamo la sua funzione di ripartizione

$$F(t) = \int_{-\infty}^{t} f(x)\,dx$$

Se $t < a \Rightarrow F(t) = \int_{-\infty}^{t} 0\,dx = 0$

Se $t \in [a,b] \Rightarrow F(t) = \int_{-\infty}^{a} 0 \, dx + \int_{a}^{t} \frac{1}{b-a} \, dx$

$$= \frac{1}{b-a} \left[x \right]_a^t = \frac{t-a}{b-a}$$

Se $t > b \Rightarrow F(t) = \int_{-\infty}^{a} 0 \, dx + \int_{a}^{b} \frac{1}{b-a} \, dx + \int_{b}^{t} 0 \, dx$

$$= 0 + 1 + 0 = 1$$

$$\Rightarrow F(t) = \begin{cases} 0 & \text{se } t < a \\ \frac{t-a}{b-a} & \text{se } t \in [a,b] \\ 1 & \text{se } t > b \end{cases}$$

Oss che F è continua in $\mathbb{R}$.

Media e varianza

Media

$$E[X] = \int_{\mathbb{R}} x \, f(x) \, dx = \int_a^b \frac{x}{b-a} \, dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b =$$

$$= \frac{b^2 - a^2}{2(b-a)} = \frac{\cancel{(b-a)}(b+a)}{2\cancel{(b-a)}} = \frac{a+b}{2}$$

$$VAR(X) = E[X^2] - (E[X])^2$$

$$E[X^2] = \int_{\mathbb{R}} x^2 f(x)\, dx = \int_a^b \frac{x^2}{b-a}\, dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$= \frac{b^3 - a^3}{3(b-a)} = \frac{\cancel{(b-a)}(b^2+a^2+ab)}{3\cancel{(b-a)}} = \frac{a^2+b^2+ab}{3}$$

$$VAR(X) = \frac{a^2+b^2+ab}{3} - \frac{(a+b)^2}{4} = \frac{4a^2+4b^2+4ab-3a^2-3b^2-6ab}{12}$$

$$= \frac{a^2+b^2-2ab}{12} = \frac{(b-a)^2}{12}$$

Esercizio

Il numero di pezzi guasti di una fornitura segue una legge $B(2,p)$ ove p è una v.a. di legge $U([0.5, 1])$.

Qual è la probabilità che un solo pezzo sia guasto

La probabilità p non è assegnata

$\Rightarrow B(2,p)$ è la legge della densità condizionale $p_{X|Y}(x|y)$ ove

X è la v.a. che conta i pezzi guasti

Y è la v.a. che assume il valore di p

$$\Rightarrow f_Y(y) = \begin{cases} \dfrac{1}{1-\frac{1}{2}} & \text{se } x \in [\frac{1}{2}, 1] \\ 0 & \text{altrimenti} \end{cases} = \begin{cases} 2 & \text{se } x \in [\frac{1}{2}, 1] \\ 0 & \text{altrimenti} \end{cases}$$

$$\Rightarrow f(x,y) = p_{X|Y}(x|y)\, f_Y(y) = \begin{cases} 2\binom{2}{x} y^x (1-y)^{2-x} & \text{se } x = 0,1,2 \\ & \quad y \in [0.5, 1] \\ 0 \text{ altrimenti} \end{cases}$$

$\Rightarrow$ La densità marginale di X è

$$f_X(x) = \begin{cases} 2\binom{2}{x} \int_{0.5}^{1} y^x (1-y)^{2-x}\, dy & \text{se } x = 0,1,2 \\ 0 & \text{altrimenti} \end{cases}$$

$\Rightarrow$ La risposta al quesito si ottiene per $x = 1$

$$f_X(1) = 2\binom{2}{1} \int_{0.5}^{1} y(1-y)\, dy = 4 \int_{0.5}^{1} y - y^2\, dy$$

$$= 4 \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_{0.5}^{1} = 4 \left[\frac{1}{2} - \frac{1}{3} - \frac{\left(\frac{1}{2}\right)^2}{2} + \frac{\left(\frac{1}{2}\right)^3}{3} \right]$$

$$= 4 \left[\frac{1}{2} - \frac{1}{3} - \frac{1}{8} + \frac{1}{24} \right] = \cancel{4}^{1}\, \frac{12 - 8 - 3 + 1}{\cancel{24}_{6}}$$

$$= \frac{\cancel{2}^{1}}{\cancel{6}_{3}} = \frac{1}{3}$$

Leggi normali

Def Una v.a. X segue una legge normale standard o gaussiana se ha densità

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

In tal caso scriveremo $X \sim N(0,1)$

Oss

Si può dimostrare che
$$\int_{-\infty}^{+\infty} f(x) = 1$$
$\Rightarrow$ $f(x)$ è una densità

Media e varianza

$$E[X] = \int_{\mathbb{R}} x\, f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\, e^{-\frac{x^2}{2}}\, dx$$

$$= -\frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} =$$

$$= -\frac{1}{\sqrt{2\pi}} \left[\underbrace{\lim_{x \to +\infty} \left(-e^{-\frac{x^2}{2}}\right)}_{=0} - \underbrace{\lim_{x \to -\infty} \left(-e^{-\frac{x^2}{2}}\right)}_{=0} \right] = 0$$

$$\mathrm{VAR}(X) = E[X^2] - (\underbrace{E[X]}_{=0})^2 = E[X^2]$$

$$= \int_{\mathbb{R}} x^2 f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (-x) \cdot \left(-x e^{-\frac{x^2}{2}}\right) dx$$

$$= \frac{1}{\sqrt{2\pi}} \left\{ \left[-x e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -1 \cdot e^{-\frac{x^2}{2}}\, dx \right\}$$

$$= \frac{1}{\sqrt{2\pi}} \left[\underbrace{\lim_{x \to +\infty} \left(-x e^{-\frac{x^2}{2}}\right)}_{=0} - \underbrace{\lim_{x \to -\infty} \left(-x e^{-\frac{x^2}{2}}\right)}_{=0} \right] + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx$$

$$= \int_{-\infty}^{+\infty} f(x)\, dx = 1$$

$\Rightarrow E[X] = 0$ e $\mathrm{VAR}(X) = 1$

Per questa ragione si usa il simbolo $N(0,1)$

Oss

La funzione di ripartizione è data da

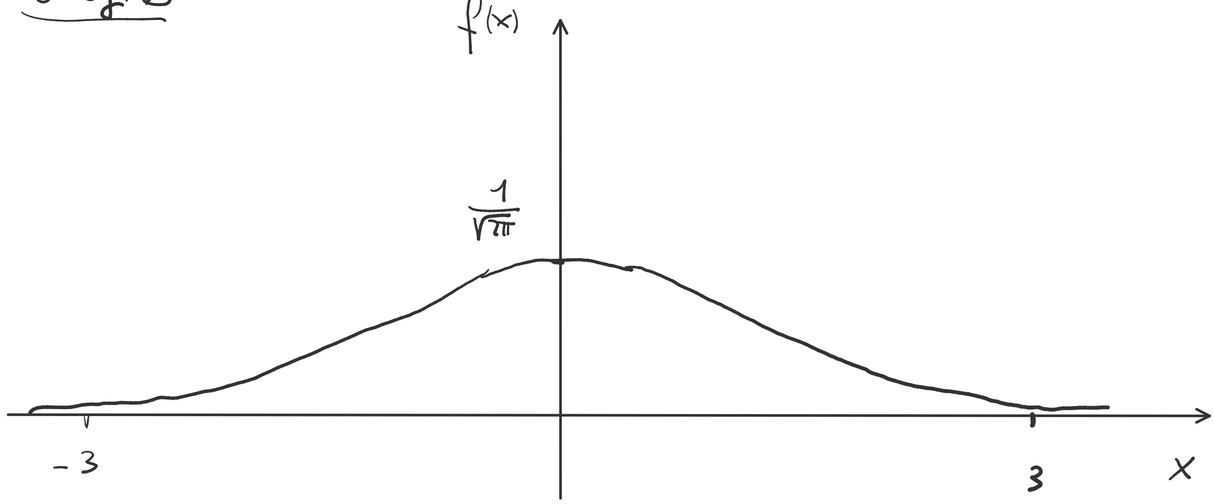
$$\Phi(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

I suoi valori si possono ottenere con procedure numeriche (come vedremo usando Python)

I quantili della normale standard si indicano con ϕ_α

<u>Proprietà</u>: Sia $\alpha \in [0,1]$. Allora $\phi_\alpha = -\phi_{1-\alpha}$

<u>Grafico</u>



Con buona approssimazione $f(x) \approx 0$ per $x < -3$ e $x > 3$

$\boxed{\text{Def}}$ Diremo che una v.a. Y segue una legge normale di parametri $\mu \in \mathbb{R}$ e $\sigma^2 > 0$ e si scrive $Y \sim N(\mu, \sigma^2)$ se $Y = \sigma X + \mu$ con $\sigma > 0$ e $X \sim N(0,1)$.

$\underline{D_3}$

La densità di Y è data da

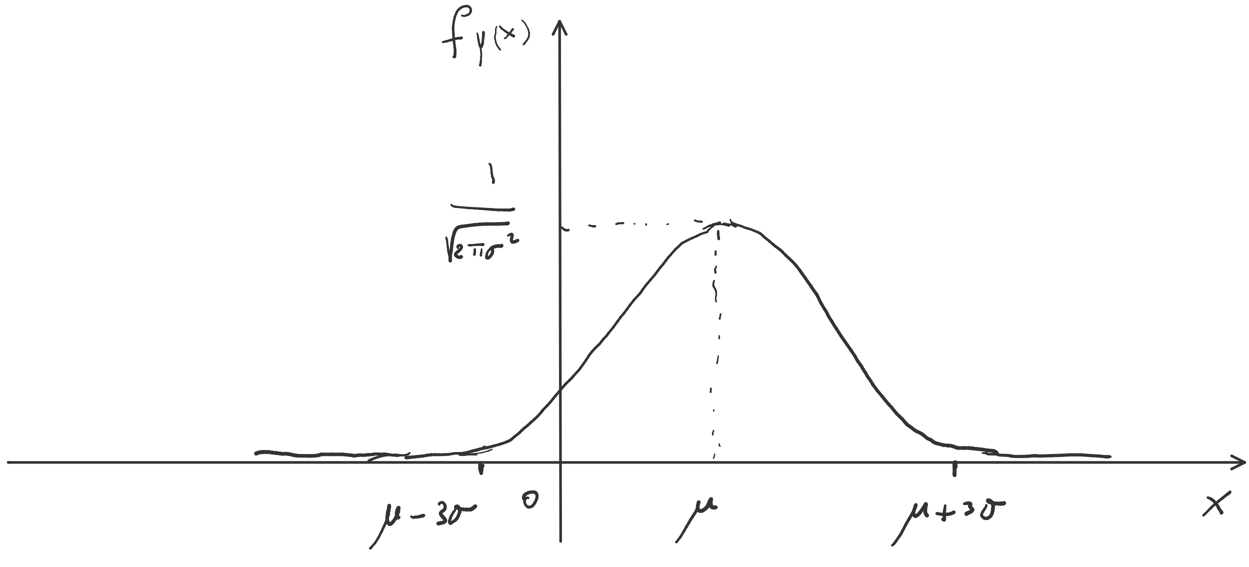
$$f_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$E[Y] = E[\sigma X + \mu] = \sigma \underbrace{E[X]}_{\overset{\shortparallel}{0}} + E[\mu]$$

$$= E[\mu] = \mu$$

$$VAR(Y) = VAR(\sigma X + \mu) = VAR(\sigma X) = \sigma^2 VAR(X) = \sigma^2$$

Grafico



Con buona approssimazione $f_Y(x) \approx 0$
per $|x-\mu| > 3\sigma$ $\quad (\Rightarrow x < \mu - 3\sigma$ e $x > \mu + 3\sigma)$

Questa proprietà vien detta "legge del 3σ".

Proprietà

Siano $Y_1 \sim N(\mu_1, \sigma_1^2)$, $Y_2 \sim N(\mu_2, \sigma_2^2)$,

Y_1 e Y_2 indipendenti

$$\Rightarrow \quad Y_1 + Y_2 \sim N(\mu_1 + \mu_2 , \sigma_1^2 + \sigma_2^2)$$

Proprietà

Sia q_α il quantile di ordine α di $Y \sim N(\mu, \sigma^2)$

ϕ_α il quantile di ordine α di $X \sim N(0,1)$

$$\Rightarrow \alpha = F_Y(q_\alpha) = P(Y \leq q_\alpha) = P(\sigma X + \mu \leq q_\alpha)$$
$$= P\left(X \leq \frac{q_\alpha - \mu}{\sigma}\right)$$

D'altra parte

$$\alpha = F_X(\phi_\alpha) = P(X \leq \phi_\alpha)$$

$$\Rightarrow \phi_\alpha = \frac{q_\alpha - \mu}{\sigma} \quad \Rightarrow \quad q_\alpha = \sigma \phi_\alpha + \mu$$

Esempio 20.2.1 *Un lago riceve acqua da due immissari e alimenta un emissario. Misurando la portata in base alla variazione di quota dell'acqua, i due immissari immettono con legge $X_1 \sim N(1,1)$ e $X_2 \sim N(2,2)$ mentre l'emissario viene alimentato con legge $X_3 \sim N(\frac{3}{2}, 3)$. Si determini la legge seguita dall'altezza dell'acqua. Qual è la probabilità che la quota superi il livello di guardia pari a 2? Qual è la probabilità che la quota sia inferiore a 0.5?*

1) La legge seguita dall'altezza dell'acqua è

$$Y = X_1 + X_2 - X_3$$

La v.a. Y segue una legge normale.

Calcoliamo media e varianza

$$E[Y] = E[X_1 + X_2 - X_3] = E[X_1] + E[X_2] - E[X_3]$$
$$= 1 + 2 - \frac{3}{2} = 3 - \frac{3}{2} = \frac{3}{2}$$

Poiché X_1, X_2, X_3 indipendenti si ha

$$VAR(Y) = VAR(X_1 + X_2 - X_3) = VAR(X_1) + VAR(X_2) + VAR(-X_3)$$
$$= VAR(X_1) + VAR(X_2) + (-1)^2 VAR(X_3)$$
$$= VAR(X_1) + VAR(X_2) + VAR(X_3) = 1 + 2 + 3 = 6$$

$$\Rightarrow Y \sim N\left(\frac{3}{2}, 6\right)$$

2) $P(Y > 2) = 1 - P(Y \leq 2) = 1 - P\left(\sqrt{6}\,X + \frac{3}{2} \leq 2\right)$

con $X \sim N(0,1)$

$$= 1 - P\left(X \leq \frac{2 - \frac{3}{2}}{\sqrt{6}}\right) = 1 - P\left(X \leq \frac{1}{2\sqrt{6}}\right)$$
$$= 1 - \Phi\left(\frac{1}{2\sqrt{6}}\right) \approx 0.42$$

3) $P(Y \leq 0.5) = P\left(\sqrt{6}\,X + \frac{3}{2} \leq \frac{1}{2}\right) = P\left(X \leq \frac{\frac{1}{2} - \frac{3}{2}}{\sqrt{6}}\right)$

$$= P\left(X \leq -\frac{1}{\sqrt{6}}\right) \approx 0.34$$

Leggi esponenziali

Def Una v.a. X segue una legge esponenziale di parametro $\lambda > 0$ se ha densità

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

In tal caso scriviamo $X \sim Exp(\lambda)$

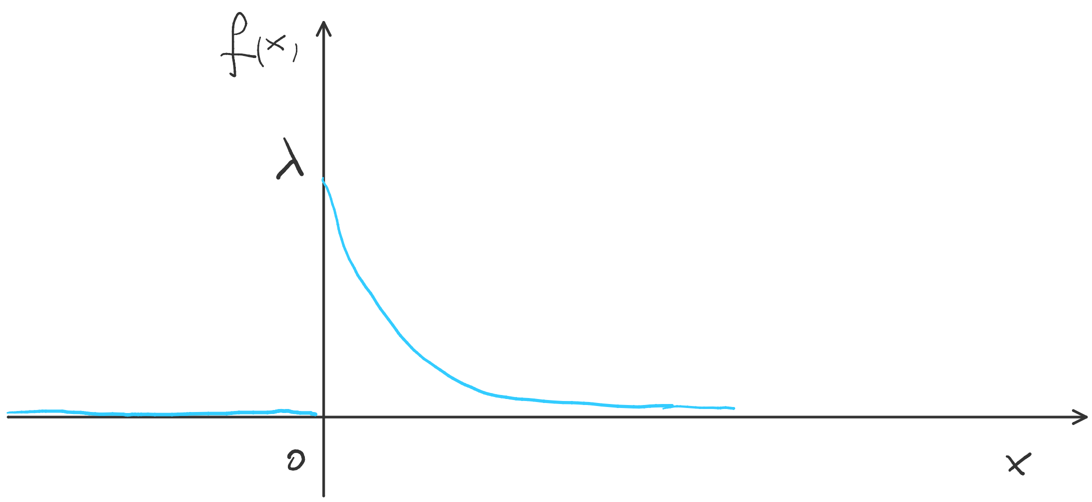
Oss

$f(x)$ è una densità di probabilità.

Infatti

$$\int_{\mathbb{R}} f(x)\,dx = \int_0^{+\infty} \lambda e^{-\lambda x}\,dx = \left[-e^{-\lambda x}\right]_0^{+\infty}$$

$$= \lim_{x \to +\infty} (\underbrace{-e^{-\lambda x}}_{=0}) - (-e^0) = -(-1) = 1$$



Media e Varianza

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

$$= -\int_0^{+\infty} x\left(-\lambda e^{-\lambda x}\right) dx = -\left\{ \Big[x e^{-\lambda x} \Big]_0^{+\infty} - \int_0^{+\infty} e^{-\lambda x}\, dx \right\}$$

$$= -\left\{ \lim_{x \to +\infty} \underset{\overset{\shortparallel}{0}}{\left(x e^{-\lambda x} \right)} - 0 - \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^{+\infty} \right\}$$

$$= -\frac{1}{\lambda} \left\{ \lim_{x \to +\infty} \underset{\overset{\shortparallel}{0}}{e^{-\lambda x}} - e^{0} \right\} = -\frac{1}{\lambda}(-1) = \frac{1}{\lambda}$$

$$VAR(X) = E[X^2] - \left(E[X]\right)^2$$

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

$$= -\int_0^{+\infty} x^2 \left(-\lambda e^{-\lambda x}\right) dx$$

$$= -\left\{ \underset{\overset{\shortparallel}{0}}{\Big[x^2 e^{-\lambda x} \Big]_0^{+\infty}} - \int_0^{+\infty} 2x e^{-\lambda x}\, dx \right\}$$

$$= \frac{2}{\lambda} \int_0^{+\infty} \lambda x e^{-\lambda x}\, dx = \frac{2}{\lambda} E[X] = \frac{2}{\lambda}\frac{1}{\lambda} = \frac{2}{\lambda^2}$$

$$\Rightarrow VAR(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Funzione di ripartizione

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\, dt$$

Se $x \le 0 \implies F_X(x) = 0$

Se $x > 0$

$$\implies F_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = -\int_0^x -\lambda e^{-\lambda t}\, dt$$

$$= -\left[e^{-\lambda t} \right]_0^x = -\left[e^{-\lambda x} - e^0 \right]$$

$$= -e^{-\lambda x} + 1 = 1 - e^{-\lambda x}$$

$$\implies F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{se } x \le 0 \end{cases}$$

<u>Oss</u>

È utile introdurre la <u>funzione di sopravvivenza</u>

$$S(x) = 1 - F_X(x)$$

$$\implies S(x) = \begin{cases} e^{-\lambda x} & \text{se } x \ge 0 \\ 1 & \text{se } x < 0 \end{cases}$$

Abbiamo che

$$S(x) = 1 - F_X(x) = 1 - P(X \le x) = P(X > x)$$

Proprietà

Vale la proprietà di mancanza di memoria, cioè

$$P(X \geq x+t \mid X \geq x) = P(X \geq t) \qquad \forall x, t \geq 0$$

Dim

$\forall\, x, t \geq 0$

$$P(X \geq x+t \mid X \geq x) \overset{\text{T. d. Bayes}}{=} \frac{P(X \geq x+t,\ X \geq t)}{P(X \geq x)} =$$

$$= \frac{P(X \geq x+t)}{P(X \geq x)} = \frac{S(x+t)}{S(x)} = \frac{e^{-\lambda(x+t)}}{e^{-\lambda x}} = e^{-\cancel{\lambda x} - \lambda t + \cancel{\lambda x}}$$

$$= e^{-\lambda t} = S(t) = P(X \geq t)$$

Oss

Le leggi esponenziali vengono utilizzate per descrivere tempi di guasto di tipo accidentale, cioè non dovuti ad effetto di rodaggio o di invecchiamento.
Si tratta del periodo di vita intermedio di un dispositivo

Esempio 20.3.1 *Siano A e B due elementi collegati in serie. Dette T_A e T_B le v.a. descriventi la durata di singoli elementi, si determini la legge dell'intero dispositivo supponendo T_A e T_B di legge esponenziale ed indipendenti tra loro.*

$$T_A \sim \text{EXP}(\lambda_A), \quad \lambda_A > 0$$

$$T_B \sim \text{EXP}(\lambda_B), \quad \lambda_B > 0$$

L'intero dispositivo si guasta non appena si guasta uno solo dei due elementi

Se $t \leq 0 \Rightarrow F_T = 0$

Se $t > 0$

$F_T(t) = P(T \leq t) = P(\min(T_A, T_B) \leq t)$

$= 1 - P(\min(T_A, T_B) \geq t) = 1 - P(T_A \geq t, T_B \geq t)$

$= 1 - P(T_A \geq t)\, P(T_B \geq t) = 1 - S_A(t)\, S_B(t)$

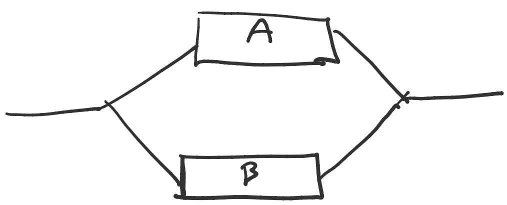
T_A, T_B indipendenti

$= 1 - e^{-(\lambda_A + \lambda_B)t}$

Esempio 20.3.2 *Siano A e B due elementi collegati in parallelo. Dette T_A e T_B le v.a. descriventi la durata di singoli elementi, si determini la legge dell'intero dispositivo supponendo T_A e T_B di legge esponenziale ed indipendenti tra loro.*

$T_A \sim EXP(\lambda_A)$, $\lambda_A > 0$

$T_B \sim EXP(\lambda_B)$, $\lambda_B > 0$



L'intero dispositivo si guasta quando entrambi gli elementi si guastano

$F_T(t) = 0$ per $t \leq 0$

se $t > 0$

$F_T(t) = P(T \leq t) = P(\max(T_A, T_B) \leq t)$

$$= P(T_A \le t, T_B \le t) = P(T_A \le t)\, P(T_B \le t)$$

$\uparrow$ perché T_A, T_B indipendenti

$$= F_A(t)\, F_B(t) = (1 - e^{-\lambda_A t})(1 - e^{-\lambda_B t})$$

Leggi chi-quadro

Def Una v.a. X segue una legge chi-quadro a n gradi di libertà se ha densità

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\, \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} & \text{se } x > 0 \\ 0 & \text{altrimenti} \end{cases}$$

ove $n \in \mathbb{N}$ e Γ rappresenta la cosiddetta funzione gamma, definita da

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

In tal caso scriveremo $X \sim \chi^2(n)$

Media e varianza

$$E[X] = n \quad , \quad VAR(X) = 2n$$

Proprietà

Sia $X \sim N(0,1) \implies X^2 \sim \chi^2(1)$

Dim

Sia $f_X(x)$ la densità di X

$F_X(t)$ la funzione di ripartizione di X

Si ha che

$P(X^2 \leq t) = 0$ se $t < 0$

Per $t \geq 0$ si ha

$$F_{X^2}(t) = P(X^2 \leq t) = P(-\sqrt{t} \leq X \leq \sqrt{t})$$
$$= F_X(\sqrt{t}) - F_X(-\sqrt{t})$$

$$\implies f_{X^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}}\left(f_X(\sqrt{x}) + f_X(-\sqrt{x})\right) & \text{se } x > 0 \\ 0 & \text{altrimenti} \end{cases}$$

Nel nostro caso si ha

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}} \cdot \frac{1}{\sqrt{2\pi}}\left[e^{-\frac{x}{2}} + e^{-\frac{x}{2}}\right] = \frac{1}{\sqrt{x}}\frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}}$$
$$= \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}}$$

Ora che $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ si ha che f_{X^2} è la densità di una legge $\chi^2(1)$

Legge di Student

Def Una v.a. T segue una legge di Student a n gradi di libertà se ha densità

$$f(t) = \frac{A_0}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}$$

con $A_0 = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$

In tal caso scriveremo $T \sim t(n)$

Media e varianza

$E[T] = 0$ se $n > 1$, altrimenti non è definita

$VAR(X) = \frac{n}{n-2}$ se $n > 2$, altrimenti è infinita.

Oss

Il quantile di ordine α di $T \sim t(n)$ si indica con $t_\alpha(n)$

Vale la proprietà $t_\alpha(n) = -t_{1-\alpha}(n)$

Oss

La legge di Student approssima la gaussiana standard per $n \to \infty$.

Solitamente per $n \geq 30$ non si fa differenza.

Oss

La legge $\chi^2(n)$ e la legge $t(n)$ hanno applicazioni in statistica che vedremo più avanti.

Leggi normali multivariate

È una v.a. continua multidimensionale t.c. ciascuna componente è una legge normale.
Consideriamo il caso di due componenti che prende il nome di legge normale bivariata

Sia $X = (X_1, X_2)$, $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N_2(\mu_2, \sigma_2^2)$
con X_1, X_2 indipendenti.

La densità congiunta è data da

$$f(x_1, x_2) = \frac{1}{2\pi \sigma_1 \sigma_2} \exp\left(-\frac{1}{2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right)\right)$$

Medie

$\mu = (\mu_1, \mu_2)$

Matrice di covariante

$$C = \left(cov(X_1, X_2) \right) = \begin{pmatrix} VAR(X_1) & COV(X_1, X_2) \\ COV(X_2, X_1) & VAR(X_2) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$