

NOZIONI DI CALCOLO DELLE PROBABILITA'

- Spazi di probabilità e sigma-algebre
- Variabili aleatorie
- Distribuzione di probabilità: funzione di distribuzione e di densità di v.a. discrete e continue
- Esempi: uniforme discreta e continua, binomiale, esponenziale e gaussiana
- Valore atteso
- Varianza e covarianza
- Sigma-algebra generata da una v.a.
- Indipendenza: eventi, sigma-algebre e v.a.
- Probabilità condizionata
- Distribuzioni di probabilità congiunte, marginali e condizionate
- Media condizionata

Spazio di probabilità
($\Omega, \mathcal{F}, P$)

Def. $\mathcal{F}$ σ -algebra

- 1) $\emptyset \in \mathcal{F}$
- 2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- 3) $\{A_n\}_{n \geq 1} \subset \mathcal{F} \Rightarrow \bigcup_n A_n \in \mathcal{F}$
- 4) $\Omega \in \mathcal{F}$
- 5) $\{A_n\}_{n \geq 1} \subset \mathcal{F} \Rightarrow \bigcap_n A_n \in \mathcal{F}$

Formule di De Morgan :

$$\bigcap_{n \geq 1} A_n = \left(\bigcup_{n \geq 1} A_n^c \right)^c$$

$$\left(\bigcap_n A_n \right)^c = \bigcup_n A_n^c$$

Def. misura di probabilità
P: $\mathcal{A} \subset \mathcal{F} \rightarrow P(A) \in [0, 1]$

$$1) P(\Omega) = 1$$

$$2) P(A \cup B) = P(A) + P(B) \text{ se } A \cap B = \emptyset$$

$$P(\emptyset) = 0$$

$$\cancel{P(\emptyset) = 0} \quad \emptyset = \Omega^c$$

$$\Rightarrow 3) P(A) = 1 - P(A^c)$$

$$\Omega = A \cup A^c$$

$$4) P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$5) A \subset B \Rightarrow P(A) \leq P(B)$$

ESEMPLI di σ -algebra:

• $\mathcal{F}^B = \{\emptyset, \Omega\}$ σ -algebra banale

• $\mathcal{A} \in \mathcal{F}$ $\sigma(\mathcal{A}) = \{A, A^c, \emptyset, \Omega\}$ σ -algebra generata da un evento

• $\mathcal{F}_1$ e $\mathcal{F}_2$ sono 2 σ -algebra $\Rightarrow \mathcal{B} = \mathcal{F}_1 \cap \mathcal{F}_2$ è una σ -algebra

- $\emptyset \in \mathcal{B}$ perché $\emptyset \in \mathcal{F}_1$ e $\emptyset \in \mathcal{F}_2$

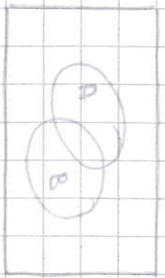
$$\{A_n \in \mathcal{F}_1 \quad \forall n \Rightarrow \bigcup_n A_n \in \mathcal{F}_1 \quad \bigcap_n A_n \in \mathcal{F}_1\} \Rightarrow \bigcup_n A_n \in \mathcal{F}_1 \cap \mathcal{F}_2 = \mathcal{B}$$

• Unione di σ -algebra non è una σ -algebra

esempio: $\sigma(A)$ e $\sigma(B)$ $\sigma(A) \cup \sigma(B) = \{A, A^c, \emptyset, \Omega, B, B^c\}$

non è una σ -algebra perché $A \cap B, A \cup B, A \cap B^c, A \cup B^c, A^c \cap B, A^c \cap B^c \notin \sigma(A) \cup \sigma(B)$

Si indica con $\mathcal{F}_1, \mathcal{F}_2$ o più piccole σ -algebra contenente $\mathcal{F}_1 \cup \mathcal{F}_2$ nell'esempio precedente $\sigma(A) \vee \sigma(B) = \sigma(A, B)$



generate dai due eventi contiene anche $A \cup B, A \cap B, A \cap B^c, A^c \cap B, A \cup B^c, A^c \cap B^c$

Varianti costruite

$$(\mathbb{R}, \mathcal{F}, P)$$

$$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$$

$$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$$

μ da misura di Lebesgue

$$\mu(a, b) = b - a \quad \text{lunghezza}$$

$$\mu((a, b) \times (c, d)) = (b - a)(d - c)$$

area dell'angolo

σ -algebra dei boreliani, generate dagli intervalli

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a, b); a, b \in \mathbb{R} \mid a < b\} = \sigma\{(-\infty, a); a \in \mathbb{R}\}$$

$$= \sigma\{(a, +\infty); a \in \mathbb{R}\}$$

$$\mathcal{B}(\mathbb{R}^2) = \sigma\{(a, b) \times (c, d); a, b, c, d \in \mathbb{R} \mid a < b, c < d\}$$

Def. di v.a.

$$X : \Omega \rightarrow \mathbb{R}$$

$$\text{misurante verso } \mathcal{V} \mid \mathcal{I} \in \mathcal{B}(\mathbb{R})$$

$$X^{-1}(\mathcal{I}) \in \mathcal{F}$$

convincente

$$\text{ove } X^{-1}(\mathcal{I}) := \{\omega \in \Omega : X(\omega) \in \mathcal{I}\}$$

Oss:

È equivalente a verificare $X^{-1}(\mathcal{I}) \in \mathcal{F} \quad \forall \mathcal{I} \in \mathcal{B}(\mathbb{R})$, per una v.a. a valori in $\mathbb{R}$ è sufficiente richiedere che:

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \leq x\} = (X \leq x) \in \mathcal{F} \quad \forall x \in \mathbb{R}$$



Distribuzione di una v.a. ^(e $\mathbb{R}^n$) è quella misura ~~definita~~ di probabilità indotta su $\mathbb{R}$ da una v.a.

• $P_X(a,b) := P(X \in (a,b))$ se X è a valori in $\mathbb{R}$

• $P_X((a,b) \times (c,d)) := P(X = (X_1, X_2) \in (a,b) \times (c,d)) = P(X_1 \in (a,b), X_2 \in (c,d)) = P((X_1 \in (a,b)) \cap (X_2 \in (c,d)))$

La distribuzione di una v.a. è completamente descritta dalla funzione di distribuzione o di ripartizione: $F_X(x) = P(X \leq x)$ se X è a valori in $\mathbb{R}$

$F_X(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$ se $X = (X_1, \dots, X_n)$ è n-dimensionale

PROPRIETÀ delle funzioni di distribuzione

1) $F_X(x) \geq 0$ 2) $\lim_{x \rightarrow +\infty} F_X(x) = 1$ 3) $F_X(x)$ continua da destra ed è monotone non decrescente

$x_1 \leq x_2 \implies P(X \leq x_1) \leq P(X \leq x_2)$

Esempio

• Distribuzione binomiale

Def. X si dice continua se F_X è una funzione continua

X si dice assolutamente continua se $\exists f_X(x) \geq 0$: $F_X(x) = \int_{-\infty}^x f_X(t) dt$
in tal caso $f_X(x)$ si dice densità di probabilità di X

Oss: $f_X(x) \geq 0$

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

allora $F_X(x) := \int_{-\infty}^x f_X(t) dt$
è una funzione di ripartizione

Se $F_X(x)$ è derivabile $\forall x \in \mathbb{R}$, eccetto un numero finito di punti

allora $f_X(x) = F'_X(x)$ nei punti in cui F_X è derivabile (è una conseguenza del teorema del Calcolo Integrale)

ESEMP1:

Distribuzione uniforme discreta su $\{x_1, \dots, x_n\}$

$P(X=x_i) = \frac{1}{n}$

$P\left(\bigcup_{i=1}^n (X=x_i)\right) = P(\Omega) = 1$

$\sum_{i=1}^n P(X=x_i)$

gli eventi sono equiprobabili $\Rightarrow P(X=x_i) = \frac{1}{n}$

Esempio: X è il risultato del lancio di un dado: assume valori in $\{1, 2, 3, 4, 5, 6\}$
e $P(X=i) = \frac{1}{6} \quad i=1, \dots, 6$

Schema successi-inuccessi

N prove indipendentemente distribuite ed indipendenti, ciascuna con probabilità di successo p

X = conteo del numero di successi nelle N prove

$P(X=k) = \binom{N}{k} p^k (1-p)^{N-k} \quad k=0, 1, \dots, N$

dist. binomiale

$\binom{N}{k} = \frac{N!}{k! (N-k)!}$

e per la formula del binomio di Newton

$\sum_{k=0}^N \binom{N}{k} p^k (1-p)^{N-k} = (p + 1-p)^N = 1^N = 1$

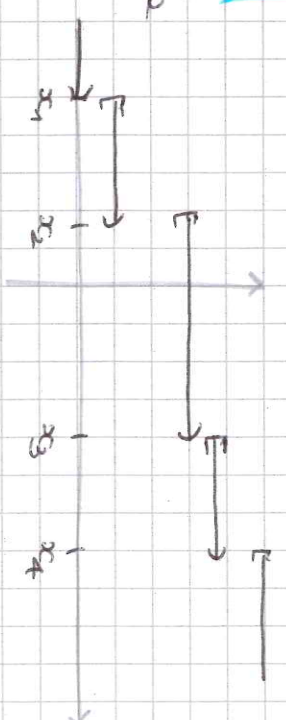
Esempio: Lancio di moneta $p = \frac{1}{2}$

La dist. uniforme su $x_1, \dots, x_n$ e la binomiale sono esempi di distribuzioni discrete, in generale

$$P(X=x_i) = p_i \quad 0 \leq p_i \leq 1 \quad \text{e} \quad \sum_{i=1}^n p_i = 1$$

$F_X(x)$ funzione costante a tratti, continua a destra

$$F_X(x) = \begin{cases} 0 & x < x_1 \\ p_1 & x \in [x_1, x_2) \\ p_1 + p_2 & x \in [x_2, x_3) \\ \vdots & \vdots \end{cases}$$



$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X=x_i) = \sum_{x_i \leq x} p_i$$

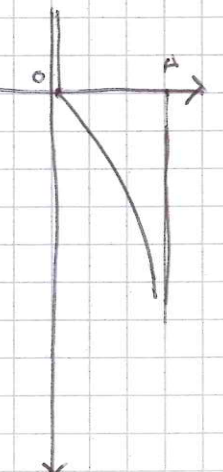
$$\lim_{t \rightarrow x^-} F_X(x) - F_X(x^-) = P(X=x)$$

Oss: Se X è continua allora $P(X=x) = 0 \quad \forall x \in \mathbb{R}$

Distribuzione esponenziale di parametro $\lambda > 0$

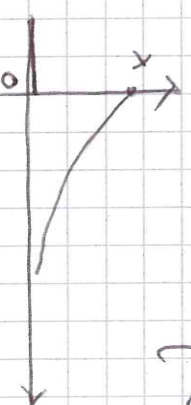
$$F_T(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

In $\exp(\lambda)$



T è una v.a. assolutamente continua (in $t=0$ F_T non è derivabile)

$$f'_T(t) = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t < 0 \end{cases} \Rightarrow \text{densità } f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$



• Distribuzione gaussiana

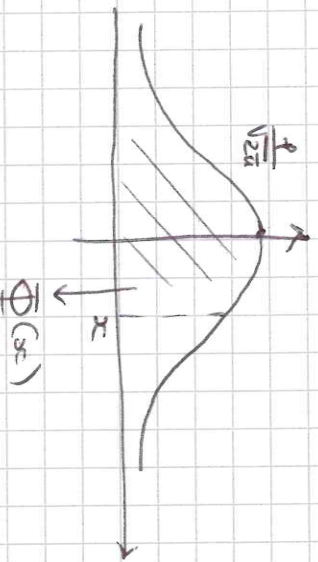
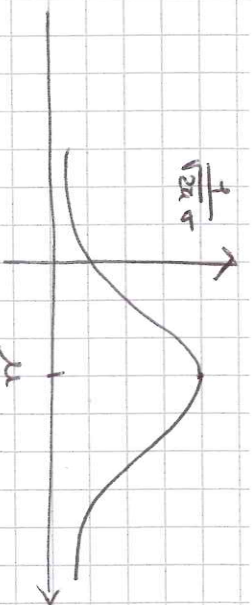
$$X \sim N(\mu, \sigma^2)$$

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

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

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

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



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

area sotto al grafico

$\Phi(x)$ non è nota esplicitamente ma i suoi valori sono tabulati (TABELLE)

$$P(X \in (a, b]) = F(b) - F(a) = P(X \leq b) - P(X \leq a)$$

è la funzione densità

$$P(X \in (a, b]) = \int_a^b f_X(x) dx$$

PROPRIETÀ
GENERALI

Proprietà delle gaussiane

$$X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

$$N \sim \mathcal{N}(0, 1) \Rightarrow \mu + \sigma N \sim \mathcal{N}(\mu, \sigma^2)$$

Esercizio per
Cassa N.1

VALORE ATTESO DI UNA V.A.

$(\Omega, \mathcal{F}, P)$

X v.a. a valori in $\mathbb{R}$ (o $\mathbb{R}^n$)

P_X dist. di probabilità su $\mathbb{R}$ (o $\mathbb{R}^n$)

$$\text{Def. } E(X) = \int_{\mathbb{R}} x P_X(dx) = \int_{\mathbb{R}} x f_X(x) dx$$

è assolutamente continuo
X ammette densità
 $f_X(x)$

$$\left[\sum_{i=1}^n x_i p_i \right]$$

se
X è discreta e ammette
densità $p_i = P(X=x_i)$

PROPRIETÀ

$$\bullet E(c) = c$$

$c \in \mathbb{R}$

$$\bullet E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i) \quad \text{Linearietà}$$

Varianza

$$\text{Var}(X) = E((X - E(X))^2)$$

Def.

$$E((X - E(X))^2) = E(X^2 + E(X)^2 - 2XE(X)) = E(X^2) + E(X)^2 - 2E(X)^2 = E(X^2) - E(X)^2$$

formula
alternativa

Def.

X e Y v.o.
si definisce la covarianza

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

Oss:

$$\text{Cov}(X, X) = \text{Var}(X), \text{ formula alternativa } \text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Proprietà della Varianza

- $\text{Var}(C) = 0$
- $\text{Var}(a + X) = \text{Var}(X)$
- $\text{Var}(aX) = a^2 \text{Var}(X)$
- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$

Esercizio per
casa N.2

Proprietà della Covarianza

- $\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$
- $\text{Cov}(aX + b, Y) = a \text{Cov}(X, Y)$

Esercizio per
casa N.3

Def.

σ -algebra generata da una v.o. X (l'insieme di informazioni che ci fornisce la conoscenza di X)
 $\sigma(X) = \{X \in \mathcal{I}\}; \mathcal{I} \in \mathcal{B}(\mathbb{R})\}$

Esempio: se X è discreta a valori in $\{x_1, \dots, x_n, \dots\}$

$$\sigma(X) = \sigma\{X = x_i\}; i \geq 1\}$$

Esempio:

$X = \begin{cases} 1 & \text{essa tira al lancio di un dado} \\ -1 & \text{essa dispare} \end{cases}$

$$P(X=1) = \frac{1}{2}$$

$$P(X=-1) = \frac{1}{2}$$

$$\sigma(X) = \sigma\{X=1\}, \{X=-1\}$$

$$\sigma(X=1) = \{\text{è uscito 2, 3, 4, 5, 6}\} = \{2, 4, 6\}$$

PROPOSIZIONE

R funzione

$$\mathbb{E}[R(X)] = \int_{\mathbb{R}} R(x) f_X(x) dx$$

Dim.

$$\text{se } R(x) = \mathbb{1}_I(x) = \begin{cases} 1 & x \in I \\ 0 & x \notin I \end{cases}$$

$$\mathbb{E}[R(X)] = \mathbb{E}\left[\sum_I \mathbb{1}_I(X)\right] = P(X \in I) = \int_I f_X(x) dx = \int_{\mathbb{R}} \mathbb{1}_I(x) f_X(x) dx = \int_{\mathbb{R}} R(x) f_X(x) dx$$

$$\text{se } R(x) = \sum_{i \geq 1} x_i \mathbb{1}_{I_i}(x)$$

$$\mathbb{E}[R(X)] = \mathbb{E}\left[\sum_{i \geq 1} x_i \mathbb{1}_{I_i}(X)\right] = \sum_{i \geq 1} x_i P(X \in I_i) = \sum_{i \geq 1} \sum_{\mathbb{R}} x_i \mathbb{1}_{I_i}(x) f_X(x) dx = \sum_{\mathbb{R}} R(x) f_X(x) dx$$

La proprietà è vera nelle funzioni semplici $\Rightarrow$ è vera per ogni funzione.

□

Esempi:

$$\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 f_X(x) dx \quad \mathbb{E}(X^R) = \int_{\mathbb{R}} x^R f_X(x) dx \quad \text{momento di ordine } R \text{ della v.a.}$$

$$\mathbb{E}(XY) = \int_{\mathbb{R}} \int_{\mathbb{R}} xy f_{(X,Y)}(x,y) dx dy$$

PROPOSIZIONE

Se X è $\sigma(Y)$ -misurabile $\Rightarrow \exists$ g funzione : $X = g(Y)$

Esempio: Y è discreta

$$\sigma(Y) = \{ (Y = y_i) ; i \geq 1 \}$$

$$X = \sum_i c_i \mathbb{1}_{(Y=y_i)}$$

infatti se X assume solo valori diversi

~~non~~ su

$$(Y = y_i) :$$

$$X^{-1}(c) \notin \sigma(Y) \text{ perché } X^{-1}(c) \subset (Y = y_i)$$

Ricordo che X è $\sigma(Y)$ -misurabile se $\forall I \in \mathcal{B}(\mathbb{R})$

$$X^{-1}(I) = \{X \in I\} \in \sigma(Y),$$

INDIPENDENZA

Def. $A, B \in \mathcal{F}$ si dicono indipendenti se $P(A \cap B) = P(A)P(B)$

Def. $\mathcal{G}_1$ e $\mathcal{G}_2$ sono σ -algebra di $\mathcal{F}$ si dicono indipendenti se
 $\forall A \in \mathcal{G}_1 \quad \forall B \in \mathcal{G}_2 \quad P(A \cap B) = P(A)P(B)$

Def. X, Y sono v.a. indipendenti se $\sigma(X)$ e $\sigma(Y)$ lo sono:
 $\forall A \in \sigma(X) \quad \forall B \in \sigma(Y) \quad P(A \cap B) = P(A)P(B)$

$$A = \{X \in I\} \quad B = \{Y \in J\}$$

$$I \in \mathcal{B}(\mathbb{R}) \quad J \in \mathcal{B}(\mathbb{R}) \quad P(X \in I, Y \in J) = P(X \in I)P(Y \in J)$$

PROPOSIZIONE 1

$$X, Y \text{ sono indipendenti} \stackrel{(1)}{\iff} F_{(X,Y)}(x,y) = F_X(x)F_Y(y) \stackrel{(2)}{\iff} f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$$

Dim.

Dalla definizione di v.a. indipendenti e poiché gli intervalli $(-\infty, x]$ sono generati da $\mathcal{B}(\mathbb{R})$

$$\text{si ha che } X, Y \text{ indipendenti} \iff P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

$$\overset{F_{(X,Y)}}{F_{(X,Y)}}(x,y) \quad \overset{F_X}{F_X}(x) \quad \overset{F_Y}{F_Y}(y)$$

e quindi $G_A(\tau)$

Ora ricordiamo che

$$F_{(X,Y)}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(u,v) du dv$$

$$F_X(x) = \int_{-\infty}^x f_X(u) du$$

$$F_Y(y) = \int_{-\infty}^y f_Y(v) dv$$

La (2) segue dal fatto che $\int_{-\infty}^x f_X(u) du \cdot \int_{-\infty}^y f_Y(v) dv = \int_{-\infty}^x \int_{-\infty}^y f_X(u) \cdot f_Y(v) dv$

PROPOSIZIONE 2

Se X e Y sono indipendenti allora $E(XY) = E(X)E(Y)$

Dim.

$$\begin{aligned} E(XY) &= \iint_{\mathbb{R} \times \mathbb{R}} xy f_{(X,Y)}(x,y) dx dy \quad \text{dalla Prop. 1} \Rightarrow E(XY) = \iint_{\mathbb{R} \times \mathbb{R}} xy f_X(x) f_Y(y) dx dy = \\ &= \int_{\mathbb{R}} x f_X(x) dx \cdot \int_{\mathbb{R}} y f_Y(y) dy \\ &= E(X)E(Y) \quad \square \end{aligned}$$

~~oss. Non vale se X e Y sono correlate e se X e Y sono indipendenti.~~

Def. X e Y si dicono scorrelate se $Cov(X,Y) = 0$

Osserviamo che $Cov(X,Y) = E(XY) - E(X)E(Y)$; quindi X e Y sono scorrelate $\Leftrightarrow E(XY) = E(X)E(Y)$.
La Prop. 2 dice che l'indipendenza $\Rightarrow$ scorrelatezza, in generale non vale il viceversa.

Def. Coefficiente di correlazione

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)} \sqrt{Var(Y)}}$$

Si può dimostrare che $\rho \in [-1, 1]$, e $\rho = 0$ X e Y sono scorrelate, e $\rho > 0$ sono correlate positivamente e $\rho < 0$ correlate negativamente.

La funzione $|f| \leq 1$ segue dalle disuguaglianze di Cauchy-Schwarz

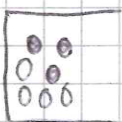
$$E|X+1| \leq \sqrt{E(X^2)} \sqrt{E(1^2)} \quad \left(\int f \cdot g dx \leq \sqrt{\int |f|^2 dx} \cdot \sqrt{\int |g|^2 dx} \right)$$

Def $A, B \in \mathcal{F}$ con $P(B) > 0$ $P(A|B) := \frac{P(A \cap B)}{P(B)}$

Osserviamo che A e B sono indipendenti $\Leftrightarrow P(A|B) = P(A)$

Da conoscenza di B
non sapere da conoscenza
di A

Esempio: estrazioni con rimpiazzo e senza rimpiazzo



N palline b bianche r rosse $N = b + r$

$$P(r) = \frac{r}{N} \quad P(b) = \frac{b}{N} = 1 - P(r)$$

2 estrazioni con rimpiazzo: $P(2^{\circ}b | 1^{\circ}r) = P(b)$
(rimettiamo dentro la pallina)

2 estrazioni senza rimpiazzo: $P(2^{\circ}b | 1^{\circ}r) = \frac{b}{N-1}$ $P(2^{\circ}r | 1^{\circ}r) = \frac{r-1}{N-1}$

$$P(2^{\circ}b | 1^{\circ}b) = \frac{b-1}{N-1}$$

$$P(2^{\circ}b) = P(2^{\circ}b, 1^{\circ}r) + P(2^{\circ}b, 1^{\circ}b) = P(2^{\circ}b | 1^{\circ}r) P(1^{\circ}r) + P(2^{\circ}b | 1^{\circ}b) P(1^{\circ}b)$$

$$= \frac{b}{N-1} \cdot \frac{r}{N} + \frac{b-1}{N-1} \cdot \frac{b}{N} = \frac{b}{N} \cdot \frac{(r+b-1)}{N-1} = \frac{b}{N}$$

ricaviamo che $P(2^{\circ}b) = P(1^{\circ}b)$ ma le estrazioni non sono indipendenti: $P(2^{\circ}b, 1^{\circ}b) = \frac{b-1}{N-1} \cdot \frac{b}{N} \neq P(2^{\circ}b) P(1^{\circ}b)$

Otteniamo anche che: $P(A \cap B) = P(A|B) P(B)$

$$\frac{b}{N} \cdot \frac{b}{N}$$

Formule di Bayes:

if any partitione di Ω :

$$\left\{ \begin{array}{l} A_i \cap A_j = \emptyset \\ \bigcup_{i=1}^n A_i = \Omega \end{array} \right.$$



$$P(A) = P(A \cap \Omega) = P\left(A \cap \bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A \cap A_i) = \sum_{i=1}^n P(A|A_i)P(A_i)$$

$$P(A) = \sum_{i=1}^n P(A|A_i)P(A_i)$$

Osserviamo infine che : $P(A|B)P(B) = P(A \cap B) = P(B|A)P(A) \Rightarrow P(B|A) = \frac{P(A|B)P(B)}{P(A)}$

$$ES: P(1^o | 2^o b) = \frac{P(2^o b | 1^o r)P(1^o r)}{P(2^o b)} = \frac{P(1^o r | 2^o b)P(2^o b)}{P(2^o b)}$$

$$\frac{b}{N-1} \cdot \frac{r}{N} = \frac{b \cdot r}{N(N-1)}$$

$$= \frac{b \cdot r}{b \cdot r + b(b-1)} = \frac{b \cdot r}{b(r+b-1)}$$

OSS: Una v.o. discreta induce una partizione $A_i = (X = x_i)$

$$\frac{b \cdot r}{b(N-1)} = \frac{r}{N} \neq P(1^o r) = \frac{r}{N}$$

ritroviamo che le estrazioni non sono indipendenti

$$P(1^o r | 2^o b) \neq P(1^o r) \neq \frac{r}{N} \neq \frac{r}{N-1}$$

$$F_{(X,Y)}(x,y) = P(X \leq x, Y \leq y) \quad f \text{ di distribuzione congiunta delle 2 v.a.}$$

$$= \sum_{-\infty}^x \sum_{-\infty}^y f_{(X,Y)}(u,v) du dv \quad \text{e } f_{(X,Y)} \text{ è densità congiunta}$$

PROPOSIZIONE

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) dv$$

calcolo delle marginali a partire dalla densità congiunta

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) du$$

Dalla densità congiunta possiamo ricavare le densità marginali, non vale il viceversa.

Dim.

$$F_X(x) = P(X \leq x) = P(X \leq x, Y \in \mathbb{R}) = \sum_{-\infty}^x \underbrace{\left(\sum_{-\infty}^{+\infty} f_{(X,Y)}(u,v) dv \right)}_{f_X(u)} du = \sum_{-\infty}^x f_X(u) du$$

Se (X,Y) sono discrete

$$P(X=x_i) = P(X=x_i, Y=y_j) = P\left(\bigcup_j (X=x_i, Y=y_j)\right) = \sum_j P(X=x_i, Y=y_j)$$

$$\Rightarrow P_X(x_i) = \sum_j P_{(X,Y)}(x_i, y_j)$$

□

Densità condizionata

X, Y discrete a valori in $\{x_1, x_2, \dots\}$ e $\{y_1, y_2, \dots\}$

$$P(X=x_i | Y=y_j) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)}$$

$$P_{X|Y}(x_i | y_j) = \frac{P_{X,Y}(x_i, y_j)}{P_Y(y_j)}$$

In modo analogo si definisce:

$$f_{X|Y}(x|y) := \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$\forall y: f_Y(y) \neq 0$$

Oss: Fissato $y: f_Y(y) \neq 0$ $f_{X|Y}(\cdot | y)$ è una densità di probabilità
infatti $f_{X|Y}(x|y) \geq 0 \quad \forall x$

$$\int_{\mathbb{R}} f_{X|Y}(x|y) dx = \frac{\int_{\mathbb{R}} f_{X,Y}(x,y) dx}{f_Y(y)} = \frac{f_Y(y)}{f_Y(y)} = 1$$

Media condizionata

$$E(X | Y=y) := \int_{\mathbb{R}} x f_{X|Y}(x|y) dx$$

media rispetto alla densità
condizionata

• $E(X|Y=y) = H(y)$ è una funzione di y

Def.

$E(X|Y)$ quella v.a. $\sigma(Y)$ -misurabile data da $H(Y)$

PROPOSIZIONE

$$E(E(X|Y)) = E(X)$$

Dim.

$$E(E(X|Y)) = E(H(Y)) = \int_{\mathbb{R}} H(y) f_Y(y) dy = \int_{\mathbb{R}} \int_{\mathbb{R}} x \underbrace{f_{X|Y}(x|y)}_{f_{X,Y}(x,y)} \cdot f_Y(y) dx dy =$$

$$= \int_{\mathbb{R}} x \underbrace{\left\{ \int_{\mathbb{R}} f_{X,Y}(x,y) dy \right\}}_{f_X(x)} dx = \int_{\mathbb{R}} x f_X(x) dx = E(X) \quad \square$$

PROPOSIZIONE $H(Y) = E(X|Y)$ rappresenta la migliore stima di X a partire dalla conoscenza di Y

$$\min_{\substack{Z \text{ v.a.} \\ \sigma(Y)\text{-mis.}}} E((Z - X)^2) = E((H(Y) - X)^2)$$

ovvero minimizza lo scarto quadratico medio

$$\text{Dim.} \quad Z = F(Y) \quad E((F(Y) - X)^2) = E((F(Y) - H(Y) + H(Y) - X)^2) = E((F(Y) - H(Y))^2) + E((H(Y) - X)^2) + 2 E((F(Y) - H(Y))(H(Y) - X))$$

$$\text{Se mostriamo che } E((F(Y) - H(Y))(H(Y) - X)) = 0 \Rightarrow E((F(Y) - X)^2) \geq E((H(Y) - X)^2) \text{ e quindi da tesi}$$

$$E((F(Y) - H(Y))(H(Y) - X)) = E(E((F(Y) - H(Y))(H(Y) - X)) | Y)) = E(\underbrace{(F(Y) - H(Y))}_{\text{misurabile}} E(H(Y) - X | Y))$$

$$E(H(Y) - X | Y) = E(\overbrace{H(Y)}^{\text{costante}} | Y) - E(X | Y) = H(Y) - E(X | Y) = 0$$

è $\sigma(Y)$ -misurabile *
e dunque è nota se
mi condizione su Y

ESERCIZI

$$1) \cdot P\left(\frac{X-\mu}{\sigma} \leq x\right) = P(X \leq \sigma x + \mu) = \int_{-\infty}^{\sigma x + \mu} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = P(N \leq x) \quad N \sim N(0,1)$$

Cambio di variabile

$$\begin{cases} y = \frac{t-\mu}{\sigma} \\ dy = \frac{1}{\sigma} dt \Rightarrow dt = \sigma dy \end{cases} \quad \text{per } t = \sigma x + \mu \rightarrow y = \frac{\sigma x + \mu - \mu}{\sigma} = \frac{\sigma x}{\sigma} = x$$

• $N \sim N(0,1)$

$$P(\mu + \sigma N \leq x) = P(N \leq \frac{x-\mu}{\sigma}) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$\text{per } t = \frac{y-\mu}{\sigma} \quad y = \mu + \sigma t \quad \text{per } t = \frac{x-\mu}{\sigma} \Rightarrow y = \mu + \sigma \frac{(x-\mu)}{\sigma} = x$$

$$dt = \frac{1}{\sigma} dy$$

• oppure $X \sim N(\mu, \sigma^2)$

$$F_N(x) = F_X(\sigma x + \mu)$$

• sia $N = \frac{X-\mu}{\sigma}$

$$F_N(x) = F'_X(\sigma x + \mu) \cdot \sigma = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\sigma x + \mu - \mu)^2}{2\sigma^2}} \cdot \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

• $N \sim N(0,1)$

• sia $X = \mu + \sigma N$

$$F'_X(x) = F'_N\left(\frac{x-\mu}{\sigma}\right)$$

$$F'_X(x) = F'_N\left(\frac{x-\mu}{\sigma}\right) \cdot \frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Valori attesi di distribuzioni note

1) $X \sim B(n, p)$
Binomiale

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$k=0, 1, \dots, n$

$$E(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}$$

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

Modo più semplice

$X_i = \begin{cases} 1 & \text{successo alla } i\text{-esima prova} \\ 0 & \text{insuccesso} \end{cases}$

$$E(X_i) = 1 \cdot p + 0 \cdot (1-p) = p$$

$$X = \sum_{i=1}^n X_i \quad E(X) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n p = np$$

2) $X \sim \text{Exp}(\lambda)$ Esponenziale

$$E(X) = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = -x e^{-\lambda x} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx = -x e^{-\lambda x} \Big|_0^{+\infty} - \frac{e^{-\lambda x}}{\lambda} \Big|_0^{+\infty} =$$

$$f = \infty \quad f' = 1$$

$$g' = \lambda e^{-\lambda x} \quad g = -e^{-\lambda x}$$

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

$$= \frac{1}{\lambda}$$

Esercizio per casa Calcolare $\text{Var}(X) = E(X^2) - E(X)^2$

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

Risultato $\text{Var}(X) = \frac{1}{\lambda^2}$

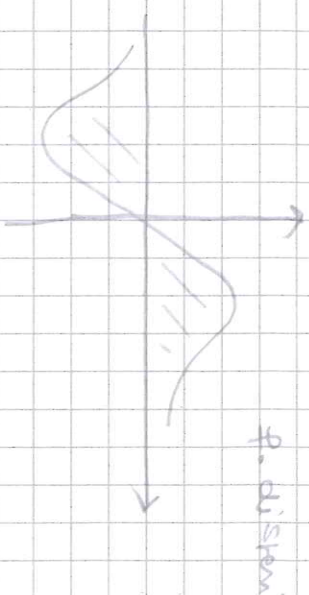
3) $X \sim N(0, 1)$ Normale standard

$$E(X) = 0$$

$$\text{Var}(X) = 1$$

$$E(X) = \int_{-\infty}^{+\infty} x \underbrace{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}_{\text{f. densità}} dx$$

$$= \int_{-\infty}^0 + \int_0^{+\infty}$$



opporre $\int_{-\infty}^{+\infty} x e^{-x^2/2} dx = -e^{-x^2/2} \Big|_{-\infty}^{+\infty} = 0$

$$D(e^{-x^2/2}) = e^{-x^2/2} \left(-\frac{2x}{2} \right) = -x e^{-x^2/2}$$

$$\text{Var}(X) = E(X^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \left(\underbrace{\int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx}_{=0} + \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx \right) = \frac{\sqrt{2\pi}}{\sqrt{2\pi}} = 1$$

$$f = x \quad f' = x$$

$$g' = x e^{-x^2/2} \quad g = -e^{-x^2/2}$$

$X \sim N(\mu, \sigma^2)$ poiché $X = \mu + \sigma N$ con $N \sim N(0, 1)$

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

$$\Rightarrow E(X) = \mu + \sigma E(N) = \mu$$

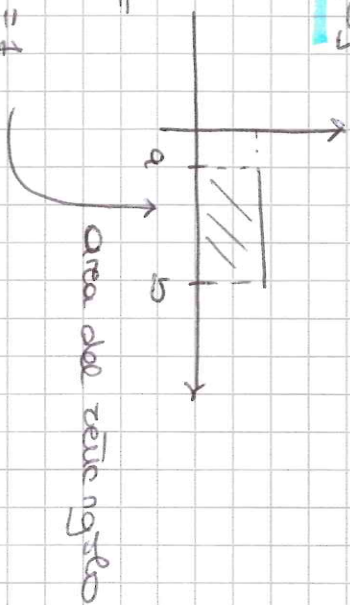
$$\text{Var}(X) = \text{Var}(\mu + \sigma N) = \text{Var}(\sigma N) = \sigma^2 \text{Var}(N) = \sigma^2$$

Esempio: distribuzione uniforme in $[a, b]$

Densità $f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}$ $a < b$

$$f_X \geq 0 \text{ e } \int_{-\infty}^{+\infty} f_X(x) dx = \int_a^b \frac{1}{b-a} dx = \frac{x}{b-a} \Big|_a^b =$$

$$= \frac{b-a}{b-a} = 1$$



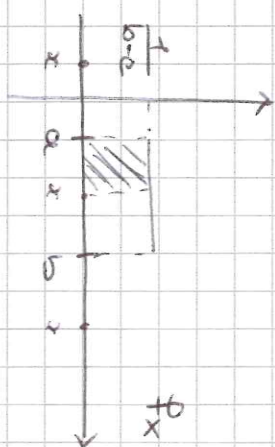
Funzione di distribuzione:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

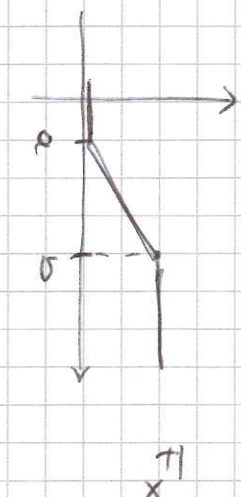
$$x < a$$

$$a \leq x < b$$

$$x \geq b$$

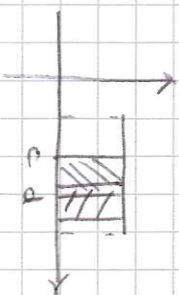


F_X è una funzione continua e $F_X'(x) = f_X(x)$ $\forall x \in \mathbb{R}, x \neq a, x \neq b$



Osserviamo che $\forall [c, d] \subseteq [a, b]$

$$P(c \leq X \leq d) = F_X(d) - F_X(c) = \int_c^d f_X(x) dx = \frac{d-c}{b-a}$$



e ora prendiamo $[c', d'] \subseteq [a, b]$

con $d' - c' = d - c$ stessa lunghezza

$$P(c' \leq X \leq d') = P(c \leq X \leq d) \text{ non dipende}$$

dai valori di c', d', c, d ma solo dalla lunghezza

di questo intervallo $\Rightarrow X$ assume in modo

equivalente i valori di $[a, b]$

Valore medio di $X \sim \text{Unif}(a, b)$

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

punto medio
dell'intervallo $[a, b]$

Proprietà di mancanza di memoria della distribuzione esponenziale $\tau \sim \exp(\lambda)$

$\forall t, s > 0$

$$P(\tau \leq t+s \mid \tau > s) \stackrel{?}{=} P(\tau \leq t+s) \quad \text{se } \tau = \text{tempo di default} \quad P(\text{avere default nell'intervallo } [0, t+s] \mid$$

se τ è il tempo di vita di una apparecchiatura priva di usura
che non c'è stato default in $[0, s]$)
= $P(\text{avere default in } [s, s+t])$
= $P(\text{avere default in } [0, t])$

$$P(\tau \leq t+s \mid \tau > s) = \frac{P(\tau \leq t+s, \tau > s)}{P(\tau > s)} = \frac{P(s < \tau \leq t+s)}{P(\tau > s)} = \frac{F_\tau(t+s) - F_\tau(s)}{1 - F_\tau(s)}$$

$$F_\tau(s) = P(\tau \leq s) = 1 - e^{-\lambda s} \quad \text{se } s \geq 0$$

$$\Rightarrow P(\tau \leq t+s \mid \tau > s) = \frac{1 - e^{-\lambda(t+s)} - (1 - e^{-\lambda s})}{e^{-\lambda s}} = \frac{e^{-\lambda s} - e^{-\lambda(t+s)}}{e^{-\lambda s}} = 1 - e^{-\lambda t} = P(\tau \leq t)$$