Formulario Variabili aleatorie

Variabili aleatorie discrete

$P(X=x)=?$ da calcolare

Funzione di massa di probabilità

1. $0 \le p(x) \le 1, \ \forall k \in \mathbb{R}$
2. $\sum_{x_i \in D_x} p(x_i) = 1$

$$p(x) = P(X=x); x \in \mathbb{R}$$

Se X è una variabile aleatoria discreta di massa di probabilità $p(x)$

- $P(X \in B) = \sum_{x_i \in D_x \cap B} p(x_i)$

Funzione di ripartizione per variabili aleatorie discrete

$$F(k) = P(X \le x); x \in \mathbb{R}$$

- $F(x) = \sum_{x_i \in D_x x_i \le x} p(x_i)$

Proposizioni

F_x è non decrescente (crescente in senso lato)(costante o crescente) ed è continua da destra

1. Per ogni $x \in \mathbb{R}, 0 \le F(x) < 1$

   $\lim_{x \to -\infty} F_x(x) = 0$
   $\lim_{x \to \infty} F_x(x) = 1$

2. $P(a < x \le b) = F(b) - F(a); \quad (a < b)$

Variabili aleatorie continue

$P(X=x)=0$ → sempre

Densità di probabilità

Se soddisfa queste due proprietà, allora è una densità di probabilità

1. $f(x) \ge 0 \qquad \forall x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Si fa l'integrale nell'intervallo dove la funzione è diversa da 0.

▼ Esempio per verificare se $f(x)$ è una densità:

$$f(x) = \begin{cases} 1-|x| & -1 \le x \le 1 \\ 0 & \text{altrimenti} \end{cases}$$

$f(x) > 0$ → sempre (guardare il grafico)

$$\int_{\infty}^{\infty} 1-|x|dx = \int_{-1}^{1} 1-|x|dx = \int_{-1}^{o} (1+x)dx + \int_{0}^{1} (1-x)dx = 1$$

Si fa l'integrale nell'intervallo dove la funzione è diversa da 0. Quindi **è densità**

- Esempio 2 per verificare se $f(x)$ è una densità:

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

$f(x) > 0$

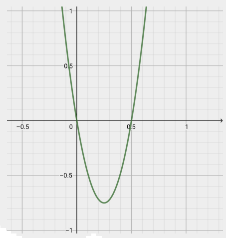
$$\int_{-\infty}^{\infty} f(x) = \int_{-\infty}^{1} 0 + \int_{1}^{\infty} e^{-(x-1)}dx = -e^{-(x-1)}\Big|_{1}^{\infty} = 0 + e^{0} = 1$$

Quindi **è densità**.

- Esempio 3 per verificare se $f(x)$ è densità:

$$f(x) = \begin{cases} 6x \cdot (2x-1) & 0 \leq x \leq 1 \\ 0 \quad x \leq 1 \end{cases}$$

Dal grafico si evince che $f(x) < 0$ tra $0 \leq x \leq \frac{1}{2}$, quindi la funzione **NON è densità**



É la frequenza con cui la variabile aleatoria assume quel valore.

$$P(x \in B) = \int_B f(x)dx = \int_{-\infty}^{\infty} f(x)\mathbf{1}_B(x)dx$$

Se X è una variabile aleatoria continua di densità f allora:

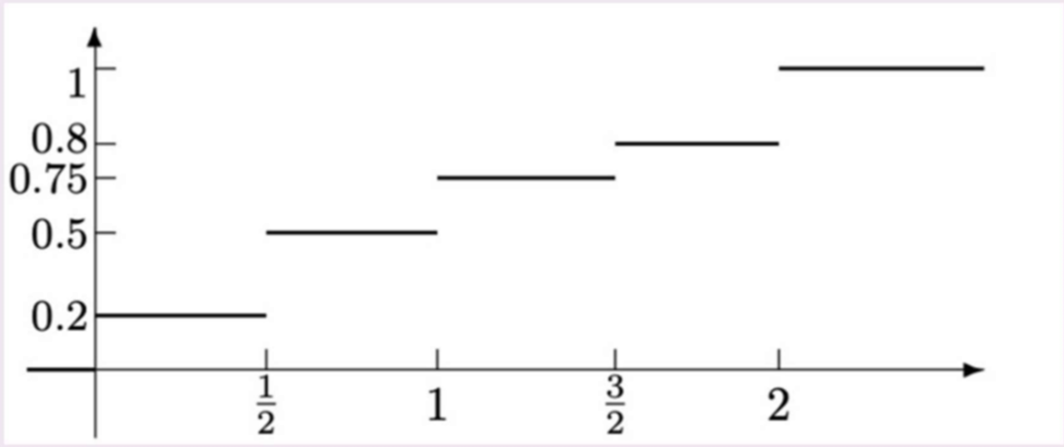
- $P(X \in B) = \int_B f(x)dx$

Funzione di ripartizione per variabili aleatorie continue

- Esempio funzione di ripartizione grafico

Sia X una variabile aleatoria la cui funzione di ripartizione ha il grafico disegnato in figura.
Link: https://www.dropbox.com/s/95fejmwq96s1lc6/figGestA.jpg?dl=0

Indicare le affermazioni vere.
(1 punto)

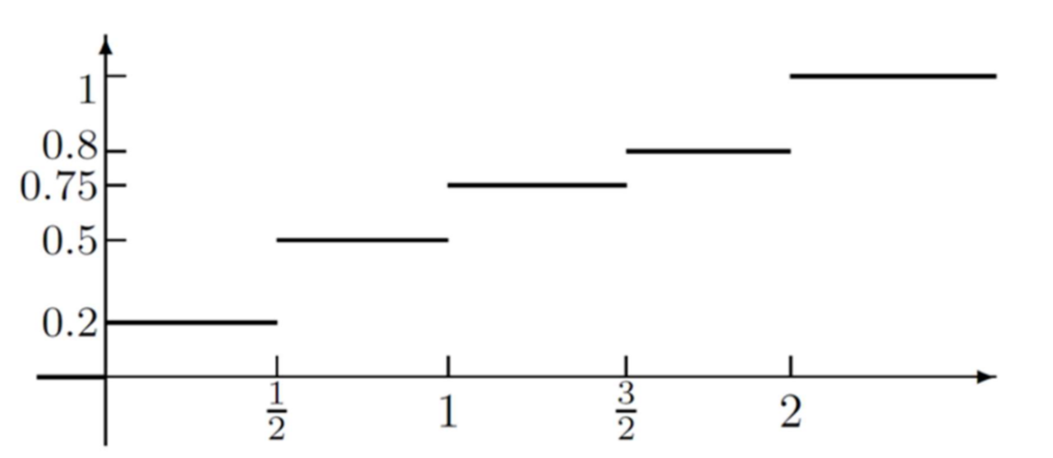


- $P(X = 1) = 0.25$ ✓ 0.75-0.5
- $P\left(X \leq \frac{3}{4}\right) = 0.5$ ✓
- $P(X = 0.5) = 0.2$
- $P(X \geq 2) = 1$

▼ Esempio 2 funzione di ripartizione grafico

5

(1 punto)



Sia X una variabile aleatoria con funzione di ripartizione F_X come in figura. Apporre i corretti valori di v

- [ ] $P\left(X = \frac{3}{2}\right) = 0.05$ ✓
- [ ] $P\left(X = \frac{3}{4}\right) = 0.5$
- [ ] $E[X] = \frac{7}{8}$ ✓

$P(X = x) = 0$ → sempre, per la definizione di variabili aleatorie continue. Si possono calcolare i valori precisi solo nei punti dove c'è un salto (cambio di valore) tipo $P(X = \frac{3}{2}) = 0.8 - 0.75 = 0.05$

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

→ integrale tra il valore più basso dell'intervallo e x

▼ Esempio 1 $x \geq 1$

$$f(x) = \begin{cases} e^{-(x-1)} & x \geq 1 \\ 0 \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{1}^{x} e^{-(x-1)} = 1 - e^{-(x-1)}$$

▼ Esempio 2 $0 \leq x \leq 1$

$$f(x) = \begin{cases} \frac{3}{5} + \frac{6}{5}x^2 & 0 \leq x \leq 1 \\ 0 \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{0}^{x} \frac{3}{5} + \frac{6}{5}x^2 = \frac{3}{5}x + \frac{6}{15}x^3$$

F_x è non decrescente (crescente in senso lato)(costante o crescente) ed è continua

1. Per ogni $x \in \mathbb{R}, 0 \leq F(x) < 1$

   $\lim_{x \to -\infty} F_x(x) = 0 \qquad \lim_{x \to \infty} F_x(x) = 1$

- $P(a \leq X \leq b) = P(X \in [a; b]) = \int_{a}^{b} f(x)dx$

- $P(X = a) = \int_a^a f(x)dx = 0$
- $P(a < X \leq b) = P(a \leq X \leq b) - P(X = a)$ (che è uguale a 0 per la proprietà 2)
  $P(a \leq X \leq b) = F_x(b) - F_x(a)$
- $F(x) = P(X \leq a) = P(X \in (-\infty; x]) = \int_{-\infty}^{x} f(t)dt$
- $P(X > x) = 1 - F(x)$
- $P(X < |a|) = P(-a < X < a)$
- $P(X > |a|) = P(x \leq -a) \vee P(x \geq a)$

$$\text{densità} \rightarrow f(x) = \frac{d}{dx}F(x) \leftarrow \text{Funzione di ripartizione}$$

Coppie e vettori di variabili aleatorie

Funzione di ripartizione congiunta per variabili discrete

$$F(x, y) := P(X \leq x, Y \leq y)$$
$$(X \leq x, Y \leq y) = (X \leq x) \cap (Y \leq y) \subset S$$

Distribuzione congiunta per variabili aleatorie discrete

Massa di probabilità congiunta

$$p(x_i, y_j) := P(X = x_i, Y = y_j);$$
$$i = 1, 2, \dots \quad j = 1, 2, \dots$$

è la loro funzione di massa di probabilità congiunta

$$(X = x_i; Y = y_j) =$$
$$= (X = x_i) \cap (Y = y_j)$$

▼ Esempio coppia di variabili aleatorie (tabella)

X\Y	-1	0	1	P(X = x)
0	1/4	1/6 [4]	1/3 [1]	3/4
1	1/12 [3]	0	1/6	1/4
P(Y = y)	1/3 [2]	1/6	1/2	1

$E(Y) = \frac{1}{6}$

$x + \frac{1}{6} = \frac{1}{2} \rightarrow x = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$ [1]

$E(Y) = \sum_{y \in D_y} y \cdot P(Y = y) = -1 \cdot P(Y = y) + 0 \cdot P(Y = 0) + 1 \cdot P(Y = 1)$

$\frac{1}{6} = -1 \cdot x + \frac{1}{2} \rightarrow +x = \frac{1}{2} - \frac{1}{6} \rightarrow x = \frac{1}{3}$ [2]

$\frac{1}{4} + x = \frac{1}{3} \Rightarrow x = \frac{1}{12}$ [3]

$1 = \frac{1}{4} + x + \frac{1}{3} + \frac{1}{12} + 0 + \frac{1}{6} \Rightarrow x = 1 - \frac{1}{4} - \frac{1}{3} - \frac{1}{12} - \frac{1}{6} \Rightarrow x = \frac{1}{6}$ [4]

$P(X = x) \rightarrow$ Somma dei valori nelle righe

$P(Y = y) \rightarrow$ Somma dei valori delle colonne

b) $E(x) = \sum_{x \in D_x} x \cdot P(X=x) \to 0 \cdot P(X=0) + 1 \cdot P(X=1) \to 1 \cdot \frac{1}{4} = \frac{1}{4}$

$E(x^2) = \sum_{x \in D_x} x^2 \cdot P(X=x) \to 0 \cdot P(X=0) + 1^2 \cdot P(X=1) \to 1 \cdot \frac{1}{4} = \frac{1}{4}$

$Var(x) = E(x^2) - (E(x))^2 \to \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$

$E(Y^2) = \sum_{y \in D_y} y^2 \cdot P(Y=y) \to (-1)^2 \cdot P(Y=-1) + 0 \cdot P(Y=0) + 1^2 \cdot P(Y=1) \to \frac{1}{3} + \frac{1}{2} = \frac{5}{6}$

$Var(Y) = E(Y^2) - (E(Y))^2 = \frac{5}{6} - \left(\frac{1}{6}\right)^2 = \frac{5}{6} - \frac{1}{36} = \frac{29}{36}$

$E(XY) = \sum_{x \in D_x} \sum_{y \in D_y} x \cdot y \cdot P(X=x, Y=y) \to 0 \cdot (-1) P(X=0, Y=-1) + 0 \cdot 0 \cdot P(X=0, Y=0) + 0 \cdot 1 \cdot P(X=0, Y=1)$
$+ 1 \cdot (-1) P(X=1, Y=-1) + 1 \cdot 0 \cdot P(X=1, Y=0) + 1 \cdot 1 \, P(X=1, Y=1)$
$= -\frac{1}{12} + \frac{1}{6} = \frac{1}{12}$

$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{12} - \frac{1}{4} \cdot \frac{1}{6} = \frac{1}{12} - \frac{1}{24} = \frac{1}{24}$

Nella tabella i valori non possono essere negativi.

▼ Esempio 2 coppia di variabili aleatorie

Sia X, Y una coppia di variabili aleatorie con la seguente densità congiunta dipendente da un parametro $\theta \in \mathbb{R}$

	$X=0$	$X=1$	$X=2$
$Y=0$	1/8	1/4	θ
$Y=1$	1/8	1/4	$1/4-\theta$

Domanda B.1 Quali valori può assumere il parametro θ?

Soluzione: I valori in tabella non possono essere negativi quindi

$$0 \le \theta \le 1/4$$

Domanda B.2 Quali sono le densità marginali di X e Y?

Soluzione:

$$X \sim B(2, 1/2) \qquad Y \sim B(1, 5/8 - \theta)$$

Domanda B.3 Trovare tutti i valori di θ per i quali X e Y non sono correlate.

Soluzione: Risulta

$$\mathbb{E}[X] = 1, \quad \mathbb{E}[Y] = \frac{5}{8} - \theta, \quad \mathbb{E}[XY] = \mathbb{P}\{X=1, Y=1\} + 2\mathbb{P}\{X=1, Y=2\} = \frac{3}{4} - 2\theta$$

X e Y non sono correlate se e solo se

$$\frac{5}{8} - \theta = \frac{3}{4} - 2\theta \qquad \Leftrightarrow \qquad \theta = \frac{1}{8}$$

Funzioni di massa marginali

Se X e Y sono due variabili aleatorie discrete su S. $(X : S \to \mathbb{R}, Y : S \to \mathbb{R})$

$$\boxed{P(X=x_i)=\sum_j P(X=x_i,Y=y_j)}$$
$$\boxed{P(Y=y_j)=\sum_i P(X=x_i,Y=y_j)}$$

- $P(X=x_i)$ → massa di probabilità (marginale) di X
- $\sum_i P(X=x_i,Y=y_j)$ → massa di probabilità congiunta di X e Y

$$\text{La funzione: }\begin{cases} p_x(x_i):=P(X=x_i); \\ i=1,2,\ldots \\ p_y(y_j):=P(Y=y_j); \\ j=1,2,\ldots \end{cases}$$

$p_x(x_i)$ e $p_y(y_j)$ si dicono funzioni di massa marginali (o individuali).

La funzione di ripartizione congiunta per variabili continue

$$F(x,y)=P((X,Y)\in(-\infty;x]*(-\infty;y])$$
$$=P((X\le x),(Y\le y))$$
$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(t,z)dtdz \qquad \forall(x,y)\in\mathbb{R}^2$$

Distribuzione congiunta per variabili aleatorie continue

Densità congiunta

Due variabili aleatorie X e Y sono congiuntamente continue, se esiste una funzione $f:\mathbb{R}^2\to\mathbb{R}$ tale che:

1. $f(x,y)>0 \quad \forall(x,y)\in\mathbb{R}^2$
2. $\int\int_{\mathbb{R}} f(x,y)dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)dxdy=1$
3. Per ogni $C\subset\mathbb{R}^2$ (sottoinsieme di $\mathbb{R}^2$), si ha:

$$P((X,Y)\in C)=\int\int_C f(x,y)dxdy$$

f si dice la densità congiunta delle variabili aleatorie X e Y.

Densità marginali

Se X e Y sono due variabili congiuntamente continue, di densità congiunta $f(x,y)$. $Y:S\to\mathbb{R}$
$D_y\subset\mathbb{R} \quad (Y\in D_y)=S \quad D_y\subset\mathbb{R}\Rightarrow(Y\in D_y)\subset(Y\in\mathbb{R})\subset S\Rightarrow(Y\in\mathbb{R})=S$

$$\boxed{f_x(x)=\int_{-\infty}^{\infty}f(x,y)dy}$$
$$\boxed{f_y(y)=\int_{-\infty}^{\infty}f(x,y)dx}$$

Le densità f_x e f_y si dicono densità marginali di X e Y.

Variabili indipendenti

Se X e Y (discrete o continue) sono indipendenti se e solo se:

$$\boxed{F(x,y) = F_x(x) * F_y(y)}$$
$$\text{Funzione di ripartizione congiunta} = \text{funzione di ripartizione di } x * y$$

Variabili discrete

Se X e Y sono due variabili aleatorie discrete di funzione di massa p_x e p_y, e di funzione di massa congiunta p, allora X e Y sono indipendenti se e solo se

$$\Leftrightarrow \boxed{p(x,y) = p_x(x) * p_y(y)} \qquad \forall x, y \in \mathbb{R}$$

Variabili continue

Se X e Y sono due variabili aleatorie continue di densità f_x e f_y, e di densità congiunta f, allora X e Y sono indipendenti se e solo se

$$\Leftrightarrow \boxed{f(x,y) = f_x(x) * f_y(y)} \qquad \forall x, y \in \mathbb{R}$$

Distribuzione condizionata

Variabili discrete

Siano X e Y due variabili aleatorie discrete con **funzioni di massa congiunta** $p(x,y)$

Si dice funzione di massa di probabilità condizionata di X dato Y e si denota con:

$$p_{X|Y}(x|y) := P(X = x | Y = y)$$

$$= \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x,y)}{p_y(y)} \quad \text{con } p_y(y) > 0 \quad \forall x, y$$

$P(Y = y)$ → massa marginale.

Dove $P(X = x, Y = y)$ → è la probabilità che si verifichino entrambi

Variabili continue

Siano X e Y due variabili aleatorie continue con funzione di densità $f(x,y)$. Si dice densità di X dato Y, la funzione:

$$f_{X|Y}(x|y) := \frac{f(x,y)}{f_y(y)}$$
$$\forall x, y \text{ con } f_y(y) > 0$$

Dove $f(x,y)$ → è la probabilità che si verifichino entrambi

▼ Esempio funzione di ripartizione condizionata per variabili continue

Il tempo di rottura di un macchinario (in migliaia di ore) viene descritto da una variabile aleatoria T con funzione di ripartizione

$$F_T(t) = \begin{cases} 0 & t < 1 \\ 1 - \frac{1}{t^4} & t \geq 1 \end{cases}.$$

2.2 Condizionatamente all'ipotesi che il macchinario non si sia ancora rotto dopo 2000 ore, calcolare la probabilità (condizionale) che si rompa non prima di 4000 ore.

Soluzione:

$$P\{T > 4 | T > 2\} = \frac{1 - F_T(4)}{1 - F_T(2)} = \frac{2^4}{4^4} = \frac{1}{16}$$

Generalizzazione a più di due variabili aleatorie:

Tutti risultati della sezione precedente si possono estendere in modo analogo ad un numero arbitrario n di variabili aleatorie:

1. La funzione di ripartizione congiunta di $x_1, \ldots, x_n$ è data da:

$$F(x_1, x_2, \ldots, x_n) := P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$$

2. Se $X_1, X_2, \ldots, X_n$ sono n variabili aleatorie discrete, la funzione di massa congiunta è:

$$p(x_1, x_2, \ldots, x_n) := P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

   - La funzione di massa marginale di X_iè data da:

$$p_{xi}(x_i) = \sum_{x_1} \cdots \sum_{x_{i-1}} \cdots \sum_{x_{i+1}} \cdots \sum_{x_n} p(x_1, \ldots, x_{i-1}, \ldots, x_{i+1}, \ldots, x_n)$$

   Non compare la i nella funzione di massa (se calcoliamo la funzione di massa di 2, non comparirà la $p_{xi}(2)$ nel calcolo).

     - $p_{x1}(x_1) = \sum_{x_2} p(x_1, x_2)$
     - $p_{x2}(x_2) = \sum_{x_1} p(x_1, x_2)$

3. Se $X_1, X_2, \ldots, X_n$ sono n variabili aleatorie continue, allora esiste una funzione $f : \mathbb{R}^n \to \mathbb{R}$ tale che:
   1. $f(x_1, \ldots, x_n) \geq 0 \qquad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n$
   2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1$
   3. $\int \cdots \int_{(x_1, \ldots, x_n) \in C} f(x_1, \ldots, x_n) dx_1, \ldots, dx_n = P((X_1, \ldots, X_n) \in C)$

4. n variabili aleatorie (discrete o continue) si dicono indipendenti se per ogni sottoinsieme

$$A_1, \ldots, A_n \qquad \text{di } \mathbb{R} \text{ si ha:}$$
$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1) * \cdots * P(X_n \in A_n)$$

   - $X_1, \ldots, X_n$ sono indipendenti se e solo se

$$F(x_1, \ldots, x_n) = F_{x_1}(x_1) * \cdots * F_{x_n} x_n$$
$$\Leftrightarrow \begin{cases} p(x_1, \ldots, x_n) = p_{x_1}(x_1) * \cdots * p_{x_n}(x_n) \to \text{ nel caso discreto} \\ f(x_1, \ldots, x_n) = f_{x_1}(x_1) * \cdots * f_{x_n}(x_n) \to \text{ nel caso continuo} \end{cases}$$

5. Le collezioni infinite di variabili aleatorie si dicono indipendenti se ogni loro sottogruppo finito è formato da variabili aleatorie tutte indipendenti.

Quantili e quartili

Il **quantile** è il valore per cui $P(X \leq x) = \frac{\text{quantile}}{100}$ → $F(x) = \frac{\text{quantile}}{100}$

Il 1° **quartile** è pari al 25-esimo **quantile.**

Esempio

C.1 Sia X una variabile aleatoria con funzione di ripartizione:

$$F(x) = \begin{cases} 0 & \text{per } x \leq 0 \\ x/3 & \text{per } 0 < x \leq 3 \\ 1 & \text{per } x > 3 \end{cases}$$

Calcola il primo quartile di X, ossia il 25-esimo quantile:

$$P(X \leq x) = 0.25 \rightarrow \frac{x}{3} = \frac{1}{4} \rightarrow x = \frac{3}{4}$$

$x = \frac{3}{4}$ → è il **primo quartile** di X, nonchè il 25-esimo **quantile**

Media o valore atteso

Il risultato sarà il valore medio che assume la variabile aleatoria X. Se si sommano i valori di ogni esperimento e si fa una media, questo numero si avvicinerà al risultato di $E[x]$.

Valore atteso per una variabile discreta

$$E[X] = \sum_{x \in D_x} x_i p_{x_i} \rightarrow \sum_{x \in D_x} x_i P(X = x_i)$$

$$E[X^2] = \sum_{x \in D_x} x_i^2 p_{x_i} \rightarrow \sum_{x \in D_x} x_i^2 P(X = x_i)$$

Funzione e variabile aleatoria discreta

$$E(g(X)) = \sum_{x_i \in D_x} g(x_i) p_x(x_i)$$

Momento n-esimo per variabile aleatoria discreta

$$E(X^n) = \sum_{x_i \in D_x} x_i^n p(x_i)$$

Valore atteso per una coppia di variabili aleatorie discrete

$$E(g(X, Y)) = \sum_x \sum_y g(X, Y) p(x, y)$$

$p(x, y)$ → massa congiunta

Valore atteso per una variabile continua

L'integrale è definito nell'intervallo. Al posto di $-\infty$ e ∞ ci vanno i valori dell'intervallo nel quale è definita $f(x)$ e si vuole calcolare il valore atteso:

▼ Esempio

$$f(x) = \begin{cases} 0 & \text{per } x \leq 1 \\ \frac{3}{x^4} & \text{per } x > 1 \end{cases}$$

$E(X^2) = \int_1^\infty x^2 \cdot \frac{3}{x^4} = 3$ → l'intervallo è $(1; \infty)$ e quindi integriamo tra $(1$ e $\infty)$

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Funzione e variabile aleatoria continua

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f_x(x)dx$$

Momento n-esimo per variabile aleatoria continua

$$E(X^n) = \int_{-\infty}^{\infty} x^n f(x)dx$$

Valore atteso per una coppia di variabili continue

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(X,Y)f(x,y)dxdy$$

$f(x,y)$ → densità congiunta

Variabile aleatoria e costanti

$$E(aX+b) = aE(x)+b \quad \text{con } g(x) = aX+b$$

Se $g(X,Y) = X+Y$ abbiamo:

$$E(X+Y) = E(g(X,Y)) = E(X)+E(Y)$$

Generalizzazioni a più variabili

$$\begin{gathered} E(X_1,X_2,X_3) = E(X_1+(X_2+X_3)) \\ \Rightarrow E(X_1+Y) = E(X_1)+E(Y) \\ \Rightarrow E(X_1)+E(X_2+X_3) \\ \Rightarrow E(X_1)+E(X_2)+E(X_3) \end{gathered}$$

Varianza

Essa fornisce una **misura** della **variabilità** dei valori assunti dalla variabile stessa; nello specifico, la misura di quanto essi si discostino quadraticamente rispettivamente dalla media aritmetica o dal valore atteso $E(X)$.

La varianza di una variabile aleatoria X con media $\mu = E(X)$, se esiste è data da:

$$\sigma^2 = Var(X) = E(X^2) - (E(X))^2 = E[(X-E(X))^2]$$

Proprietà:

1. $Var(aX+b) = a^2Var(X)$ → La b (costante) sparisce.
2. $Var(b) = 0 \quad \forall b \text{ costante}$
3. $Var(X+Y) \neq Var(X)+Var(Y)$ →
    1. Se sono **indipendenti** → $Var(X+Y) = Var(X)+Var(Y)$
        1. ATTENZIONE, se sono **indipendenti** → $Var(2X-2Y) = 2^2Var(X)+Var(-2Y) = 4Var(X)+(-2)^2Var(Y) = Var(X)+4Var(Y)$

$$Var(X-Y) = Var(X) + Var(-Y) = Var(X) + Var(Y)$$

2. Se sono **dipendenti** → $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

$$Var(X-Y) = Var(X) + (-1)^2 Var(Y) - 2Cov(X,Y)$$

4. Se gli eventi sono indipendenti, $Var(X) = Var(-X)$ → è simmetrica

$$Var(X-Y) = Var(X) + Var(-Y) = Var(X) + Var(Y)$$

5. Se X e Y sono indipendenti

$$Var(XY) = (E(X))^2 \cdot Var(Y) + (E(Y))^2 \cdot Var(X) + Var(X) \cdot Var(Y)$$

Ricordiamo che per $|x| < 1$:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

$$\text{derivata prima} \rightarrow \sum_{n=0}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1} \tag{7}$$

$$\text{derivata } 2^\circ \rightarrow \sum_{n=0}^{\infty} n(n-1) x^{n-2} = \frac{2}{(1-x)^3} = \sum_{n=2}^{\infty} n(n-1) x^{n-2} \tag{8}$$

Deviazione standard

É un **indice di dispersione statistico**, vale a dire una **stima della variabilità** di una popolazione di dati o di una variabile casuale. È uno dei modi per esprimere la **dispersione dei dati** intorno ad un **indice di posizione.**

$$\sqrt{Var(X)}$$

Covarianza

La **covarianza** di due variabili aleatorie X e Y con $\mu_x = E(X)$ e $\mu_y = E(Y)$, è un numero che fornisce una misura di quanto le due varino assieme, ovvero della loro dipendenza. Se esiste è data da:

$$Cov(X,Y) = E(XY) - E(X) \cdot E(Y)$$
$$Cov(X,Y) = E[(X - E(X)) \cdot (Y - E(Y))]$$

Proprietà

1. $Cov(X,Y) = Cov(Y,X)$

   ## Teorema

   Se X e Y sono due **variabili indipendenti**, allora $Cov(X,Y) = 0$

   Se la $Cov(X,Y) = 0 \not\Rightarrow X$ e Y siano indipendenti, ma implica che esse sono **scorrelate**

2. $Cov(X,X) = Var(X)$
3. $Cov(aX,Y) = aCov(X,Y)$

4. $Cov(X, a) = 0$ con $a \in \mathbb{R}$ → La costante e la variabile aleatoria non si influenzano tra di loro. Quindi sono sempre **indipendenti**.

   La a può essere vista come una variabile aleatoria costante che ha sempre lo stesso valore.

Generalizzazione a più variabili

Lemma → vale sempre, anche se non sono indipendenti

$$Cov(X+Y, Z) = Cov(X, Z) + Cov(Y, Z)$$

$$Cov(X+Y, X-Y) = Cov(X, X) - \cancel{Cov(X, Y)} + \cancel{Cov(Y, X)} - Cov(Y, Y) = Var(X) - Var(Y)$$

$$Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$$

$$Cov(\sum_{i=1}^{n} X_i, Y) = \sum_{i=1}^{n} Cov(X_i, Y)$$

E in modo analogo:

$$Cov(X, \sum_{j=1}^{m} Y_j) = \sum_{j=1}^{m} Cov(X, Y_j)$$

Proprietà:

1. $Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} Cov(X_i, Y_j)$
2. $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j) \text{ con } i \neq j$

- Se $X_1, \ldots, X_n$ sono indipendenti, abbiamo visto che

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

Coefficiente di correlazione di X e Y

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

Funzione generatrice dei momenti

Di una variabile aleatoria X, si indica con ϕ ed è definita per ogni numero reale t tale che $E(e^{tx})$ esiste:

$g(x) = e^{tx}$

Nel caso discreto

$$\phi(t) = E(g(x)) = E(e^{tx}) = \sum_{i} e^{tx_i} \cdot p_x(x_i)$$

Nel caso continuo

$$\phi(t) = E(g(x)) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} \cdot f_x(x)dx$$

Osservazione

$$\text{derivata} \rightarrow \phi'(t) = \frac{d}{dt}E(e^{tx}) = E(\frac{d}{dt}e^{tx}) = E(Xe^{tx})$$
$$\text{derivata } 2^\circ \rightarrow \phi''(t) = \frac{d^2}{dt^2}E(e^{tx}) = E(\frac{d^2}{dt^2}e^{tx}) = E(X^2e^{tx})$$
$$\text{derivata n}^\circ \rightarrow \phi^n(t) = \frac{d^n}{dt^n}E(e^{tx}) = E(\frac{d^n}{dt^n}e^{tx}) = E(X^ne^{tx})$$

In particolare:

$$\phi'(o) = E(X)$$
$$\phi''(o) = E(X^2)$$
$$\phi^n(o) = E(X^n)$$

$$Var(X) = \phi''(0) - (\phi'(0))^2$$

Funzione generatrice dei momenti di due variabili aleatorie indipendenti

Se X e Y sono indipendenti (discrete o continue), con funzione generatrice ϕ_x e ϕ_y e se ϕ_{x+y} è la funzione generatrice di $X+Y$, allora:

$$\phi_{x+y}(t) = \phi_x(t) \cdot \phi_y(t)$$

La legge debole dei grandi numeri

Disuguaglianza di Markov

Per ogni variabile aleatoria X non negativa, si ha:

$$P(X \geq a) \leq \frac{E(X)}{a} \quad \forall a > 0$$

$P(X \geq a)$ → limite superiore

Disuguaglianza di Chebyshev

Probabilità che una variabile assuma un valore compreso tra altri due valori.

Sia X una variabile aleatoria con media μ, varianza σ^2, e deviazione standard σ. Allora per ogni $r > 0$

$r \in \mathbb{R}$ → è un valore simmetrico rispetto alla media (es: se la media è 24, 21 e 27 sono valori simmetrici rispetto alla media con $r = 24 - 21 = 3$)

Limite superiore della probabilità

$$P(-r \leq X - \mu \leq r) \leq \frac{\sigma^2}{r^2}$$

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

$$P(\mu - r \leq X \leq \mu + r) = 1 - P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2} = P(|X_\mu| \geq r) \geq 1 - \frac{\sigma^2}{r^2}$$

Limite inferiore della probabilità

$$P(|X - \mu| \leq r) \geq 1 - \frac{\sigma^2}{r^2}$$

Teorema (Legge debole dei grandi numeri)

Siano $X_1, \dots, X_n$ n variabili aleatorie **i.i.d** (indipendenti e identicamente distribuite (stessa distribuzione) $P(X = x)$), tutte con media $E(X_i) = \mu$. Allora per ogni $\epsilon > 0$

$$P\Big(\Big|\frac{X_1 + \cdots + X_n}{n} - \mu\Big| > \epsilon\Big) \to 0 \qquad \text{quando } n \to \infty$$

Mediana

Variabile aleatorie continue

La mediana z è il valore tale che $F(z) = \frac{1}{2}$ → è il punto medio della probabilità

Dove $F(z) \Rightarrow P(X \leq z)$

$$f(x) = \begin{cases} e^{-(x-1)} & x > 1 \\ 0 \quad x \leq 1 \end{cases}$$
$$F(x) = \int_{-\infty}^{t} f(x)dx = \begin{cases} 1 - e^{-(t-1)} & t > 1 \\ 0 \quad t \leq 1 \end{cases}$$

$F(z) = \frac{1}{2}$ se $1 - e^{-(z-1)} = \frac{1}{2} \Rightarrow -e^{-z+1} = \frac{1}{2} - 1 \Rightarrow e^{-z+1} = \frac{1}{2} \Rightarrow z = \ln 2 + 1 \simeq 1,69$

Variabili aleatorie discrete

É la media aritmetica della somma dei numeri x che assume una variabile aleatoria discreta X, divisi per quanti numeri x ci sono.(la quantità di x)

- Esempio

  Un urna contiene 3 palline → i valori che x può assumere sono **3** e sono $1, 2, 3$

$$\text{Mediana } \to z = \frac{1+2+3}{3} = 6$$