

1.1 – RIEPILOGO: VARIABILI CASUALI DISCRETE E FUNZIONE DI PROBABILITÀ

Una variabile casuale discreta assume un numero finito o un'infinità numerabile di valori. Definito il supporto $S_x = \{x_1, x_2, x_3\}$, sappiamo per certo che $p(X \in S_x) = 1$, cioè la variabile casuale X può assumere solo i valori del suo supporto.

Definiamo **funzione di probabilità** la funzione tale che:

$$p(X = x_i) = p_x(x_i)$$

Una funzione di probabilità, per essere tale, deve rispettare le seguenti proprietà:

- $p_x(x_i) \geq 0$
- $\sum_{x_i \in S_x} p(x_i) = \sum_{x_i} p(X = x_i) = p(X \in S_x) = 1$

Se conosco le probabilità, la funzione di ripartizione si calcola in questo modo:

$$F(X) = p(X \leq x) = \sum_{x_i \leq x} p_x(x_i)$$

1.2 – VALORE ATTESO E VARIANZA

Il **valore atteso** o media si definisce:

$$E[x] = \sum_i x_i \cdot p(x_i)$$

La **varianza** si definisce:

$$V[x] = \sum_i (x_i - E[x])^2 \cdot p(x_i)$$

oppure si può utilizzare anche la formula:

$$V[x] = E[x^2] - (E[x])^2$$

$$\text{cioè } V[x] = E[x^2] - (E[x])^2 = \sum x_i^2 \cdot p_x(x_i) - \left(\sum x_i \cdot p_x(x_i)\right)^2$$

1.3 – PROPRIETÀ DELLA VARIANZA

La **varianza** rispetta le seguenti proprietà:

$$V(a+bx) = b^2 \cdot V(x)$$

$$V(a+x) = V(x)$$

Se $V(x) = 0$ allora x è degenere, cioè è costante.

Ad esempio quindi avremo $x = a$, cioè $p(x=a) = 1$. È quindi una variabile discreta in cui il supporto è dato da un unico valore.

1.4 – INDIPENDENZA FRA VARIABILI CASUALI

Cosa significa se x e y sono indipendenti? Il fatto di conoscere qualcosa su uno dei due fenomeni non mi dice niente sull'altro. Formalmente significa che, dati gli intervalli A e B , si ha $\forall A, B$:

$$p(x \in A | y \in B) = p(x \in A)$$

Riassumendo, x e y sono indipendenti se un evento qualsiasi relativo a x e un evento qualsiasi relativo a y sono indipendenti, cioè:

$$p(x \in A \cap y \in B) = p(x \in A) \cdot p(y \in B)$$

Analogamente, date n variabili casuali $x_1, x_2, x_3, \dots, x_n$ ed n intervalli vale l'indipendenza (che è quindi generalizzabile), cioè formalmente:

$$p(x_1 \in A_1 \cap x_2 \in A_2 \cap \dots \cap x_n \in A_n) = p(x_1 \in A_1) \cdot p(x_2 \in A_2) \cdot \dots \cdot p(x_n \in A_n)$$

1.5 – VARIABILI CASUALI DISCRETE: DISTRIBUZIONE DI BERNOULLI

Considero un esperimento casuale ε che ha solo due possibili risultati, successo (S) o insuccesso ($\bar{S}$). La variabile casuale X assume quindi i seguenti valori:

$X = 1$ se si è verificato (S)
 $X = 0$ se non si è verificato ($\bar{S}$)

La variabile casuale X ha associate le seguenti funzioni di probabilità:

X	$p_x(x_i)$
1	ϑ
0	$1 - \vartheta$

dove $0 \leq \vartheta \leq 1$.

Quindi:

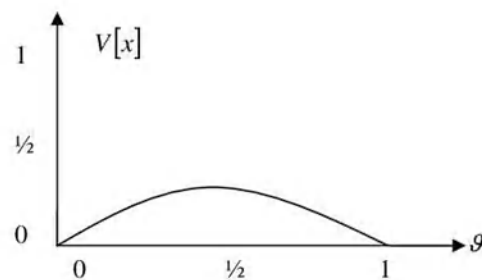
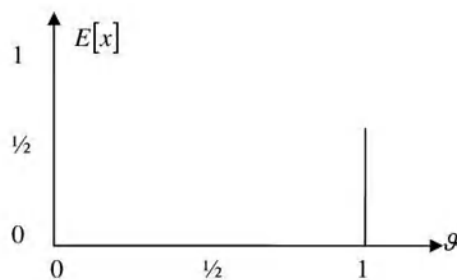
$$p(X = 1) = p_x(1) = p(S)$$

$$p(X = 0) = p_x(0) = p(\bar{S})$$

Nella distribuzione di Bernoulli, andiamo ora a calcolare media e varianza:

$$E[X] = \sum x \cdot p_x(x) = 1 \cdot \vartheta + 0 \cdot (1 - \vartheta) = \vartheta$$

$$V[X] = \sum (x - E[X])^2 \cdot p_x(x) = (1 - \vartheta)^2 \cdot \vartheta + (0 - \vartheta)^2 \cdot (1 - \vartheta) = (1 - \vartheta)^2 \cdot \vartheta + \vartheta^2 \cdot (1 - \vartheta) = \vartheta + \vartheta^3 - 2\vartheta^2 + \vartheta^2 - \vartheta^3 = \vartheta - \vartheta^2 = \vartheta \cdot (1 - \vartheta)$$



1.6 – VARIABILI CASUALI DISCRETE: DISTRIBUZIONE BINOMIALE

Consideriamo il lancio di una moneta regolare, vogliamo che esca testa (T); la variabile casuale X assume i seguenti valori:

$$X_1(\text{Bernulli}) = \begin{cases} 1 = T(\text{lancio1}) \\ 0 = C(\text{lancio1}) \end{cases}$$

$$X_2(\text{Bernulli}) = \begin{cases} 1 = T(\text{lancio2}) \\ 0 = C(\text{lancio2}) \end{cases}$$

...

$$X_n(\text{Bernulli}) = \begin{cases} 1 = T(\text{lancion}) \\ 0 = C(\text{lancion}) \end{cases}$$

La distribuzione binomiale va a studiare quante volte è uscita T (testa).

X è il numero di successi in n prove ed è dato dalla somma degli esiti delle singole prove, cioè:

$$X = X_1 + X_2 + X_3 + \dots + X_n$$

In una distribuzione binomiale è necessario che:

1) **Le prove sono tra loro indipendenti**: ci sono n variabili di Bernoulli X_i tra loro indipendenti

2) $p(X_i = 1) = \vartheta = p(S)$, cioè le X_i hanno la stessa distribuzione:

$$X_i = \begin{cases} 1 = S \\ 0 = \bar{S} \end{cases}$$

In una distribuzione binomiale, quindi, la variabile casuale X dipende da due parametri:

$$X \sim B_i(n, \vartheta)$$

con $0 \leq \vartheta \leq 1$.

Come ricavare la distribuzione di questa variabile e quindi la funzione di probabilità?

Sapendo che $S_x = \{0, 1, 2, \dots, n\}$, la distribuzione binomiale risulta della forma:

$$p(X = x) = p_x(x) = \binom{n}{x} \cdot \vartheta^x \cdot (1 - \vartheta)^{n-x}$$

Dimostrazione

Consideriamo ad esempio $p(X = 1)$, cioè la probabilità di avere un successo. Lo spazio dei possibili risultati Ω risulta composto da 2^n possibili risultati, perché ci può essere un successo o un insuccesso in ogni estrazione ed ogni risultato è una n -pla. Se consideriamo appunto una delle n -ple in cui si è verificato un solo risultato (un solo successo), avremo ad esempio:

$$w_1 = (S, \bar{S}, \bar{S}, \bar{S}, \bar{S}, \dots, \bar{S})$$

cioè:

$$p(w_1) = p((X_1 = 1) \cap (X_2 = 0) \cap (X_3 = 0) \cap \dots \cap (X_n = 0))$$

ma dato che gli eventi sono indipendenti, essa equivale a:

$$p(w_1) = p(X_1 = 1) \cdot p(X_2 = 0) \cdot p(X_3 = 0) \cdot \dots \cdot p(X_n = 0) = \vartheta \cdot (1 - \vartheta)^{n-1}$$

Ma non basta considerare la successione w_1 , devo considerare anche altre successioni, come w_2 :

$$w_2 = (\bar{S}, S, \bar{S}, \bar{S}, \bar{S}, \dots, \bar{S})$$

cioè:

$$p(w_2) = p((X_1 = 0) \cap (X_2 = 1) \cap (X_3 = 0) \cap \dots \cap (X_n = 0))$$

ma dato che gli eventi sono indipendenti, essa equivale a:

$$p(w_2) = p(X_1 = 0) \cdot p(X_2 = 1) \cdot p(X_3 = 0) \cdot \dots \cdot p(X_n = 0) = (1 - \vartheta) \cdot \vartheta \cdot (1 - \vartheta) \cdot \dots \cdot (1 - \vartheta) = \vartheta \cdot (1 - \vartheta)^{n-1}$$

In generale, per tutte le combinazioni possibili, che sono n , avremo:

$$p(X = 1) = n \cdot [\vartheta \cdot (1 - \vartheta)^{n-1}]$$

Ma se voglio calcolare in generale $p(X = x)$?

Avrò ad esempio una successione w_3 del tipo:

$$w_3 = (S, S, S, S, \bar{S}, \bar{S}, \bar{S}, \bar{S}, \dots, \bar{S})$$

dove ci sono x successi e $n - x$ insuccessi, da cui:

$$p(w_3) = \vartheta^x \cdot (1 - \vartheta)^{n-x}$$

e si ottiene un risultato analogo per tutte le n combinazioni:

$$p(w_i) = \vartheta^x \cdot (1 - \vartheta)^{n-x}$$

Ma quante sono queste n -ple che contengono esattamente x successi?

Utilizzo il coefficiente binomiale:

$$\binom{n}{x} = \frac{n!}{(n-x)!x!}$$

Unendo, le relazioni dimostrate, otteniamo quindi:

$$p(X = x) = p_x(x) = \binom{n}{x} \cdot \vartheta^x \cdot (1 - \vartheta)^{n-x}$$

c.v.d.

Media in una distribuzione binomiale

In generale la media è data da:

$$E[X] = \sum x \cdot p_x(x)$$

Nel nostro caso, quindi avremo:

$$E[X] = \sum_{i=0}^n i \cdot \binom{n}{i} \cdot \vartheta^i \cdot (1 - \vartheta)^{n-i}$$

In questo modo la media è scomoda e complicata; è più facile considerare la nostra X , come somma di tutte le n variabili X_i , cioè:

$$X = X_1 + X_2 + \dots + X_n$$

dove $E[X_i] = \vartheta$

Pertanto avremo:

$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = n\vartheta$$

Da notare che vale sempre la seguente relazione:

$$E[X_1 + X_2] = E[X_1] + E[X_2]$$

Varianza in una distribuzione binomiale

Nella distribuzione di Bernoulli, la varianza è pari a:

$$V[X] = E[X^2] - [E[X]]^2 = E[X^2] - \vartheta^2 = \vartheta \cdot (1 - \vartheta)$$

Nel nostro caso avremo:

$$V[X] = V\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n V[X_i] = n(\vartheta \cdot (1 - \vartheta))$$

Da notare che la seguente relazione:

$$V[X_1 + X_2] = V[X_1] + V[X_2]$$

vale solo se le variabili sono incorrelate o indipendenti.

Inoltre se $\vartheta = 1/2$ l'incertezza è massima.

1.7 – ESEMPIO

Consideriamo 4 lanci di una moneta regolare. Vogliamo calcolare la probabilità di avere 2 teste e la probabilità di avere almeno 3 teste. X è il numero totale di teste (T) in 4 lanci.

La probabilità di successo è:

$$p(S) = \vartheta = \frac{1}{2}$$

Pertanto abbiamo a che fare con una distribuzione binomiale:

$$X \sim B_i\left(4, \frac{1}{2}\right)$$

Pertanto avremo:

$$p(X = 2) = \binom{4}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{4-2} = \frac{4!}{2!2!} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{3}{8}$$

$$p(X \geq 3) = p(X = 3) + p(X = 4) = \binom{4}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{4-3} + \binom{4}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^{4-4} = \frac{5}{16}$$

1.8 – ESEMPIO

Supponiamo che la probabilità di avere una figlia femmina sia:

$$p(F) = g = \frac{1}{2}$$

e che il genere dei figli sia indipendente l'uno dall'altro. Quanti figli deve avere una coppia per avere almeno una figlia? X è il numero di figli su n e la coppia vuole essere sicura al 95% di avere una figlia.

Vogliamo quindi calcolare:

$$p(X \geq 1) \geq 0,95$$

Anche qui abbiamo a che fare con una distribuzione binomiale:

$$X \sim B_i\left(n, \frac{1}{2}\right)$$

$$p(X \geq 1) \geq 0,95$$

$$p(X = 1) + p(X = 2) + \dots + p(X = n) \geq 0,95$$

$$1 - p(X = 0) \geq 0,95$$

$$1 - \binom{n}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^{n-0} \geq 0,95$$

$$1 - \frac{n!}{0!n!} \cdot 1 \cdot \left(\frac{1}{2}\right)^n \geq 0,95$$

$$1 - \left(\frac{1}{2}\right)^n \geq 0,95$$

$$-\left(\frac{1}{2}\right)^n \geq -0,05$$

$$\left(\frac{1}{2}\right)^n \leq 0,05$$

$$n \ln\left(\frac{1}{2}\right) \leq \ln(0,05)$$

Poiché $0 \leq \ln \leq 1$ è negativo, il segno cambia nuovamente e quindi si ottiene:

$$n \geq \frac{\ln(0,05)}{\ln(1/2)} = 4,322$$

Pertanto:

$$n \geq 5$$

1.9 – ESEMPIO

Supponiamo che una fabbrica produca 10 viti per ogni scatola. Se un cliente compra una scatola e trova anche un solo pezzo difettoso, la scatola viene cambiata gratuitamente. L'azienda sa che produce normalmente 1% di pezzi difettosi. X è il numero di viti difettose in ogni scatola.

La probabilità di successo (in questo caso il successo è trovare una vite difettosa) è:

$$p(S) = 0,01 = 1\%$$

Vogliamo calcolare:

$$p(X \geq 1) = 1 - p(X = 0)$$

e sappiamo che la distribuzione è binomiale:

$$X \sim B_i(10, 0.01)$$

Pertanto avremo:

$$p(X \geq 1) = 1 - p(X = 0) = 1 - \binom{10}{0} \cdot (0,01)^0 \cdot (1 - 0,01)^{10} = 1 - \frac{10!}{0!10!} \cdot 1 \cdot (0,99)^{10} = 1 - (0,99)^{10} = 0,096$$

1.1 – ESERCIZIO

Un'urna contiene N palline, delle quali K sono verdi ed $N - K$ sono rosse. Si estraggono n palline con reinserimento. Calcolare la probabilità che la seconda pallina estratta sia verde.

Definito V_i , i -esima pallina estratta verde, sappiamo che:

$$p(V_2) = \frac{K}{N}$$

Ad esempio, se $N = 10$, $K = 7$, $n = 3$, ne segue che:

$$p(V_2) = \frac{7}{10}$$

Calcolare inoltre la probabilità che fra le n palline estratte ve ne siano esattamente 2 verdi.

$$\begin{aligned} p(2V) &= p(V_1 \cap V_2 \cap \bar{V}_3) + p(V_1 \cap \bar{V}_2 \cap V_3) + p(\bar{V}_1 \cap V_2 \cap V_3) = p(V_1) \cdot p(V_2) \cdot p(\bar{V}_3) + p(V_1) \cdot p(\bar{V}_2) \cdot p(V_3) \\ &+ p(\bar{V}_1) \cdot p(V_2) \cdot p(V_3) = \frac{7}{10} \cdot \frac{7}{10} \cdot \frac{3}{10} + \frac{7}{10} \cdot \frac{3}{10} \cdot \frac{7}{10} + \frac{3}{10} \cdot \frac{7}{10} \cdot \frac{7}{10} = 3 \cdot \left(\frac{7}{10}\right)^2 \cdot \frac{3}{10} \cong 0,441 \end{aligned}$$

dove $p(\bar{V}_i) = 1 - p(V_i)$

Ma è più conveniente utilizzare la distribuzione Bernoulliana:

$$X_i = \begin{cases} 1 = V = S \\ 0 = \bar{V} = R = \bar{S} \end{cases}$$

con V verde, R rossa, S successo.

$$\vartheta = p(S) = p(X_i = 1) = \frac{K}{N} = 0,7$$

$$\sum_{i=1}^n X_i = n^{\circ} \text{verdi} = X$$

$p(X = 2) = \text{Binomiale}(n, \vartheta) = (3, 0,7)$ quindi:

$$p(X = 2) = \binom{3}{2} \frac{7}{10}^2 \frac{3}{10} = 0,441$$

$$p(X \geq 2) = p(X = 2) + p(X = 3) = 0,441 + \left(\frac{7}{10}\right)^3 = 0,441 + 0,343 = 0,784$$

Qual è il numero atteso di palline V ? E la varianza?

$$E[X] = n\vartheta = \frac{21}{10} = 2,1$$

$$V[X] = 3 \cdot \frac{7}{10} \cdot \frac{3}{10} = \frac{63}{100} = 0,63$$

Se abbiamo anche la variabile casuale Y che conta il numero di palline rosse nel campione, consideriamo sempre una binomiale, ed in particolare si ha che:

$$X + Y = n$$

da cui si ottiene:

$$Y = n - X$$

$$Y = \text{Binomiale}(n, 1 - \vartheta)$$

$$E[Y] = n(1 - \vartheta) = \frac{9}{10} = 0,9$$

$$V[Y] = n(1 - \vartheta) \cdot \vartheta = V[X] = 0,63$$

1.2 – ESERCIZIO

Un'urna contiene $N = 10$ palline, delle quali $K = 7$ sono verdi ed $N - K = 3$ sono rosse. Si estraggono n palline con reinserimento. Calcolare la probabilità di estrarre r palline rosse prima di estrarre 1 pallina verde.

$$X_i = \begin{cases} 1 & \text{se l'iesima pallina estratta è Verde} \\ 0 & \text{se l'iesima pallina estratta è Rossa} \end{cases}$$

X_i è distr. Bernulli(ϑ)

$$\vartheta = \frac{k}{N} = \frac{7}{10}$$

Vogliamo quindi calcolare:

$$p(X_1 = 0, X_2 = 0, X_3 = 0, \dots, X_r = 0, X_{r+1} = 1)$$

ma dato che gli *eventi sono indipendenti*, essa equivale a:

$$p(X_1 = 0) + p(X_2 = 0) + p(X_3 = 0) + \dots p(X_r = 0) + p(X_{r+1} = 1) = (1 - \vartheta)^r \cdot \vartheta$$

Se ad esempio $r = 2$, otteniamo:

$$p(X_1 = 0, X_2 = 0, X_3 = 0, \dots, X_r = 0, X_{r+1} = 1) = \left(\frac{3}{10}\right)^2 \cdot \frac{7}{10} = 0,063$$

2.1 – VARIABILI CASUALI DISCRETE: DISTRIBUZIONE IPERGEOMETRICA

La distribuzione ipergeometrica ha uno schema simile a quello di una distribuzione binomiale. Ma mentre in una distribuzione binomiale, le estrazioni erano con reinserimento e le prove erano quindi indipendenti l'una dall'altra, nella distribuzione ipergeometrica le estrazioni sono *senza reinserimento* e le singole *prove risultano quindi dipendenti*, e non più indipendenti.

Consideriamo un'urna contenente N palline, di cui h sono bianche (è il nostro successo) e $N - h$ sono nere (è il nostro insuccesso). Supponiamo che si facciano n estrazioni senza reinserimento. Le singole prove, come detto sopra, risultano dipendenti l'una dall'altra. X è il numero totale di successi in n prove.

$$X_i = \begin{cases} 1 = S \\ 0 = \bar{S} \end{cases}$$

all' i -esima estrazione.

Sappiamo inoltre che vale la relazione:

$$X = X_1 + X_2 + X_3 + \dots + X_n$$

In una distribuzione binomiale, in una situazione come quella descritta nell'esempio:

$$p(S) = \frac{h}{N} = \vartheta = p(X_i = 1)$$

inoltre, visto che le singole prove sono indipendenti

$$p(X_2 = 1 | X_1 = 1) = \vartheta$$

In una distribuzione ipergeometrica, invece, abbiamo visto che le prove sono dipendenti. Se ad esempio $N = 10$, $h = 5$, $\vartheta = 1/2$, la probabilità sopra scritta è pari a:

$$p(X_2 = 1 | X_1 = 1) \neq p(X_2 = 1)$$

$$p(X_2 = 1 | X_1 = 1) = \frac{4}{9}$$

$$p(X_2 = 1) = \frac{5}{10} = \frac{1}{2}$$

Vogliamo calcolare, nel caso della distribuzione ipergeometrica, $p(X = x)$ secondo la definizione classica:

$$p(X = x) = \frac{\text{casi favorevoli}}{\text{casi possibili}}$$

Casi possibili

I° estrazione	N	} n prove
II° estrazione	$N - 1$	
III° estrazione	$N - 2$	
...	...	
n ° estrazione	$N - n + 1$	

I casi possibili sono quindi:

$$N \cdot (N - 1) \cdot (N - 2) \cdot \dots \cdot (N - n + 1) = \frac{N!}{(N - n)!} = \frac{N!}{(N - n)!} \cdot \frac{n!}{n!} = \binom{N}{n} \cdot n!$$

infatti:

$$\frac{N!}{(N - n)!} = \frac{N \cdot (N - 1) \cdot (N - 2) \cdot \dots \cdot (N - n + 1) \cdot (N - n) \cdot \dots \cdot 3 \cdot 2 \cdot 1}{(N - n) \cdot \dots \cdot 3 \cdot 2 \cdot 1} = N \cdot (N - 1) \cdot (N - 2) \cdot \dots \cdot (N - n + 1)$$

Casi favorevoli

Consideriamo ad esempio la sequenza:

$$w_1 = (S, S, S, \bar{S}, \bar{S}, \bar{S}, \bar{S}, \dots, \bar{S})$$

con x successi e $n - x$ insuccessi.

In generale ci saranno $\binom{n}{x}$ modi di distribuire gli x successi sulle n posizioni.

Ma qual è la probabilità della sequenza w_1 ?

Modi in cui posso scegliere le x estrazioni: successi

I° estrazione (palline bianche)	h	} x prove
II° estrazione	$h - 1$	
III° estrazione	$h - 2$	
...	...	
n° estrazione	$h - x + 1$	

cioè:

$$h \cdot (h - 1) \cdot (h - 2) \cdot \dots \cdot (h - x + 1) = \frac{h!}{(h - x)!} = \frac{h!}{(h - x)!} \cdot \frac{x!}{x!} = \binom{h}{x} \cdot x!$$

Modi in cui posso scegliere le $n - x$ estrazioni: insuccessi

I° estrazione	$N - h$	} $n - x$ prove
II° estrazione	$N - h - 1$	
III° estrazione	$N - h - 2$	
...	...	
n° estrazione	$N - h - (n - x) + 1$	

cioè:

$$(N - h) \cdot \dots \cdot (N - h - (n - x) + 1) = \frac{(N - h)!}{(N - h - (n - x))!} = \frac{(N - h)!}{(N - h - (n - x))!} \cdot \frac{(n - x)!}{(n - x)!} = \binom{N - h}{n - x} \cdot (n - x)!$$

Pertanto la distribuzione ipergeometrica assume la forma:

$$p(X = x) = \frac{\text{casi favorevoli}}{\text{casi possibili}} = \frac{\binom{n}{x} \cdot \binom{h}{x} \cdot x! \cdot \binom{N - h}{n - x} \cdot (n - x)!}{\binom{N}{n} \cdot n!}$$

ma, ricordando che:

$$\binom{n}{x} = \frac{n!}{(n - x)! \cdot x!}$$

la distribuzione ipergeometrica assume la forma:

$$p(X = x) = \frac{\binom{n}{x} \cdot \binom{h}{x} \cdot \binom{N - h}{n - x}}{\binom{N}{n} \cdot \binom{n}{x}} = \frac{\binom{h}{x} \cdot \binom{N - h}{n - x}}{\binom{N}{n}}$$

In una distribuzione ipergeometrica, quindi, la variabile casuale X dipende da tre parametri:

$$X \sim \text{Ig}(N, n, h)$$

Ma quali sono i valori di x , cioè i possibili valori per il numero totale di successi?

In una distribuzione binomiale si ha che:

$$0 \leq x \leq n$$

In una distribuzione ipergeometrica:

$$x_{\max} = \begin{cases} x_{\max} \leq n \\ x_{\max} \leq h \end{cases}$$

da cui:

$$x_{\max} = \min(n, h)$$

$$x_{\min} = \begin{cases} x_{\min} \geq 0 \\ x_{\min} \geq n - (N - h) \end{cases}$$

da cui:

$$x_{\min} = \max(0, n - (N - h))$$

Pertanto i possibili valori di x sono:

$$\max(0, n - (N - h)) \leq x \leq \min(n, h)$$

Esempio

Su 100 ragazzi, il 20% possiede un cellulare di ultima generazione (successo). Estratti a caso 10 ragazzi, quanti possiedono il cellulare di ultima generazione? Qual è la probabilità che a possedere il cellulare di ultima generazione siano 4 ragazzi?

$$N = 100 \quad n = 10 \quad h = 20 \quad x = 4 \quad \Rightarrow \quad 0 \leq x \leq 10$$

La probabilità di successo è pari a:

$$p(S) = \frac{h}{N} = \frac{20}{100} = 0,2$$

$$p(X = 4) = \frac{\binom{20}{4} \cdot \binom{80}{6}}{\binom{100}{10}} = 0,0841 \Rightarrow 8\% \text{ circa}$$

Media in una distribuzione ipergeometrica

$$E[X] = n \cdot \frac{h}{N}$$

Dimostrazione

Consideriamo la nostra X , come somma di tutte le n variabili X_i , cioè:

$$X = X_1 + X_2 + \dots + X_n$$

da cui:

$$E[X] = E[X_1 + X_2 + \dots + X_n]$$

$$E[X] = E[X_1] + E[X_2] + \dots + E[X_n]$$

$$\text{dove } E[X_i] = p(X_i = 1) = \mathcal{G}$$

Pertanto avremo:

$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = n\mathcal{G}$$

che è la media di una distribuzione binomiale.

Nel nostro caso abbiamo che:

$$p(S) = g = \frac{h}{N}$$

quindi avremo:

$$E[X] = n \cdot \frac{h}{N}$$

c.v.d.

Varianza in una distribuzione ipergeometrica

In una distribuzione ipergeometrica, dato che le prove non sono indipendenti:

$$V[X] \neq \sum_{i=1}^n V[X_i]$$

Devo pertanto tenere conto della *covarianza*.

La varianza in una distribuzione ipergeometrica assume la forma:

$$V[X] = \left(\sum_{i=1}^n V[X_i] \right) \cdot \frac{N-n}{N-1} = n \cdot g \cdot (1-g) \cdot \frac{N-n}{N-1} = n \cdot \frac{h}{N} \cdot \left(1 - \frac{h}{N} \right) \cdot \frac{N-n}{N-1}$$

Pertanto si ha che:

$$V[X]_{IG} < V[X]_B$$

Se $N = n$, allora ne segue che $X = h$ e quindi $V[X] = 0$: la variabile è degenere, cioè è costante.

Se invece il rapporto n/N è molto basso (cioè se è bassa la percentuale che vado ad estrarre) la distribuzione ipergeometrica si avvicina alla binomiale.

2.2 – VARIABILI CASUALI DISCRETE: DISTRIBUZIONE DI POISSON

Nella distribuzione di Poisson X è il numero totale di arrivi in un certo periodo di tempo:

$$X \in \{0, 1, 2, 3, \dots, n\} = S_x = N$$

cioè studia i cosiddetti **fenomeni di arrivi casuali**.

In una distribuzione di Poisson:

- 1) Il numero di arrivi in un intervallo di tempo X_1 è indipendente dal numero di arrivi nell'intervallo di tempo X_2
- 2) Il numero di arrivi in un certo intervallo di tempo deve dipendere solo dalla lunghezza dell'intervallo considerato, cioè:

$$p(X_1 = k) = p(X_2 = k)$$

$$E[X_1] = E[X_2] = g$$

$$\text{ma } X_1 \neq X_2$$

Si noti che g è il numero medio di arrivi in un'unità di tempo.

La distribuzione di Poisson è della forma:

$$p(X = x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

con $x = 0, 1, 2, 3, \dots, n$; $\lambda > 0$.

È una funzione di probabilità?

- 1) $p_x(x) \geq 0$? sì infatti $x = 0, 1, 2, 3, \dots, n$

- 2) $\sum_{x=0}^{\infty} p_x(x) = 1$? sì (vedi dimostrazione)

Dimostrazione

Per definizione la funzione esponenziale e^x si può vedere anche come:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Dobbiamo dimostrare che:

$$\sum_{x=0}^{\infty} p_x(x) = 1$$

$$\sum_{x=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^x}{x!} = 1$$

$$e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = 1$$

$$e^{-\lambda} \cdot e^{\lambda} = 1$$

$$e^{-\lambda+\lambda} = e^0 = 1$$

Quindi è verificata la seconda proprietà delle funzioni di probabilità.

In una distribuzione di Poisson, quindi, la variabile casuale X dipende solo dal parametro λ :

$$X \sim Po(\lambda)$$

Media in una distribuzione di Poisson

$$E[X] = \lambda$$

Dimostrazione

In generale vale che:

$$E[x] = \sum_i x_i \cdot p(x_i)$$

ma nel nostro caso:

$$p_x(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

da cui si ottiene che:

$$E[x] = \sum_{x=0}^{\infty} x \cdot e^{-\lambda} \cdot \frac{\lambda^x}{x!} = \lambda$$

Varianza in una distribuzione di Poisson

$$V[X] = \lambda$$

Si conclude che nella distribuzione di Poisson, vale sempre che:

$$E[X] = V[X] = \lambda$$

Inoltre il parametro λ è tale che:

$$\lambda = T \cdot \vartheta$$

dove ϑ è il numero medio di arrivi in un'unità di tempo, T è l'unità di tempo considerata.

La distribuzione di Poisson si utilizza anche negli *eventi rari* come approssimazione della distribuzione binomiale:

$$Po(\lambda) \approx Bi(n, \vartheta)$$

Quindi:

$$n\vartheta \rightarrow \lambda$$

con $n \rightarrow \infty$, ϑ molto basso.

Esempio

X è il numero di telefonate che arrivano al centralino tra le 9.00 e le 10.00 del mattino. In media, inoltre, tra le 9.00 e le 10.00, arrivano 30 telefonate, cioè $\vartheta = 30$. Calcolare la probabilità che in tre minuti non arrivino telefonate (1), e calcolare la probabilità che in cinque minuti arrivino più di cinque telefonate (2).

(1) Per calcolare T nel nostro caso, faccio una semplice proporzione: $1:60 = T:3$ da cui si ottiene che:

$$T = \frac{3}{60} = 0,05$$

Inoltre sappiamo che $\vartheta = 30$ da cui si ottiene che:

$$\lambda = 30 \cdot 0,05 = 1,5$$

Quindi:

$$p(X = 0) = e^{-1,5} \cdot \frac{(1,5)^0}{0!} = e^{-1,5} = 0,22 \rightarrow \approx 22\%$$

(2) Per calcolare T nel nostro caso, faccio una semplice proporzione: $1:60 = T:5$ da cui si ottiene che:

$$T = \frac{5}{60} = \frac{1}{12} = 0,0833$$

Inoltre sappiamo che $\vartheta = 30$ da cui si ottiene che:

$$\lambda = 30 \cdot \frac{1}{12} = \frac{5}{2}$$

Voglio calcolare $p(X > 5)$:

$$p(X > 5) = p(X = 6) + p(X = 7) + p(X = 8) + \dots$$

che equivale a:

$$p(X > 5) = 1 - p(X \leq 5) = 1 - p(X = 1) - p(X = 2) - p(X = 3) - p(X = 4) - p(X = 5)$$

$$p(X > 5) = 1 - e^{-5/2} \cdot \frac{(5/2)^1}{1!} - e^{-5/2} \cdot \frac{(5/2)^2}{2!} - e^{-5/2} \cdot \frac{(5/2)^3}{3!} - e^{-5/2} \cdot \frac{(5/2)^4}{4!} - e^{-5/2} \cdot \frac{(5/2)^5}{5!} = 0,124$$

Esempio

Supponiamo che $\vartheta = 0,001$, $n = 2000$ e X sia il numero di persone malate. Per quanto detto in precedenza, dato che $n \rightarrow \infty$ e ϑ è molto basso, possiamo utilizzare la distribuzione binomiale, quindi $X \sim Bi(n, \vartheta) = Bi(2000, 0,001)$.

Calcolo la probabilità che su 2000 estratti, 3 persone sole siano malate:

$$p(X = 3) = \binom{2000}{3} \cdot (0,001)^3 \cdot (1 - 0,001)^{2000-3} = 0,1805$$

$$p(X = 3) = e^{-2} \cdot \frac{2^3}{3!} = 0,1804$$

infatti $n\vartheta \rightarrow \lambda$ e quindi $\lambda = n\vartheta = 2$.

2.3 – VARIABILI CASUALI CONTINUE E FUNZIONE DI DENSITÀ

Nelle variabili discrete i possibili valori erano elencabili (si poteva cioè fare una lista finita o infinita di valori possibili): $X \in \{0,1,2,\dots\}$. Ma non tutti i fenomeni sono discreti e assumono questa caratteristica: ad esempio il peso, l'altezza, la temperatura sono variabili continue (variano cioè in modo continuo, **non sono elencabili**): si parlerà ad esempio della variabile casuale $X \in (0,\infty)$, cioè tutti i possibili valori tra 0 e $+\infty$.

Nel caso discreto, abbiamo finora utilizzato la funzione di probabilità, andando a calcolare $p(X = x)$. Nel caso continuo, invece, non è possibile utilizzare la funzione di probabilità perché non ha senso, non mi dice niente:

$$p(X = x) = 0$$

$\forall x$, cioè è sempre 0. Le probabilità nel caso continuo vanno associate quindi non ai singoli valori, ma agli intervalli: avremo ad esempio $p(a < X < b)$ oppure $p(X < k)$.

Nel caso continuo parleremo di **funzione di densità** $f_x(X)$: la variabile casuale X è continua se esiste una certa funzione detta funzione di densità, tale che la funzione di ripartizione di X , $F_x(X) = p(X \leq x)$ si calcola come:

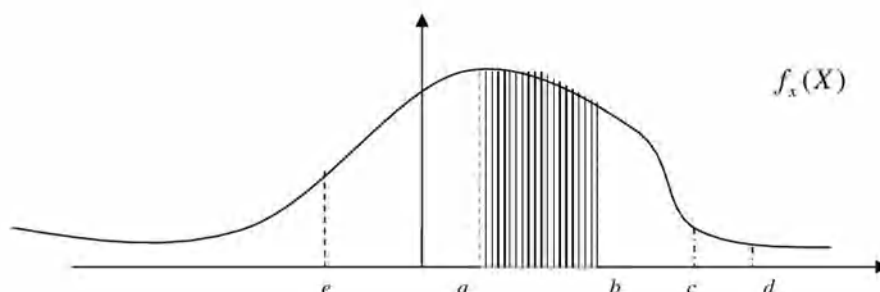
$$F_x(X) = \int_{-\infty}^x f_x(X) \cdot dx$$

cioè **la funzione di ripartizione è l'integrale della funzione di densità**.

Pertanto si avrà che:

$$p(a < X < b) = \int_a^b f_x(X) \cdot dx$$

Come interpreto la funzione di densità?



La probabilità è l'area sotto la funzione di densità

La probabilità non è quindi la funzione di densità.

Inoltre si ha che:

$$p(a < X < b) > p(c < X < d)$$

e vale che $p(X = e) = 0$, dato che il punto e ha area 0.

Se conosco la funzione di ripartizione, come calcolo la funzione di densità?

$$F'_x(X) = \frac{dF_x(X)}{dX} = f_x(X)$$

cioè, **la funzione di densità è la derivata della funzione di ripartizione**.

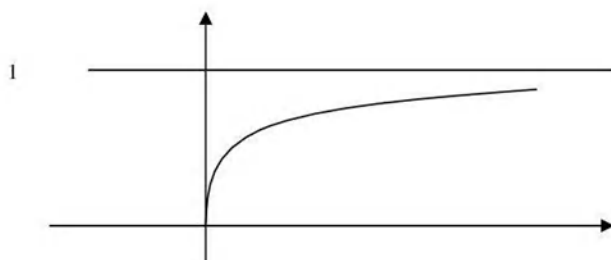
Esempio

È data la seguente funzione di ripartizione:

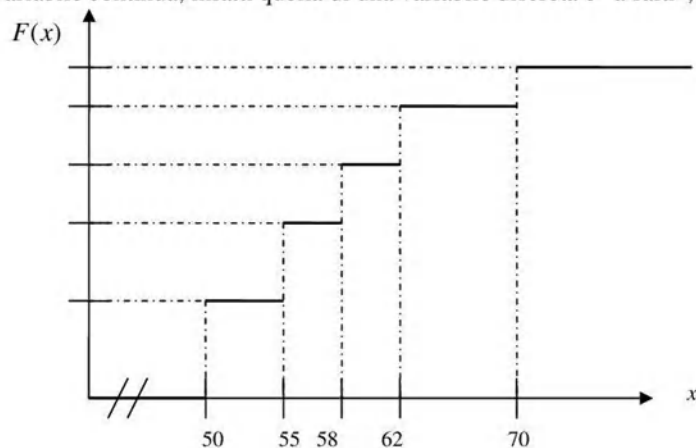
$$F_x(X) = \begin{cases} 1 - e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Verifichiamo innanzitutto se è una funzione di ripartizione, controllando se la funzione rispetta le proprietà:

- $\lim_{x \rightarrow -\infty} F_x(X) = 0$ sì
- $\lim_{x \rightarrow +\infty} F_x(X) = 1$ sì
- $F_x(X)$ è non decrescente, cioè se $x_1 \leq x_2$ allora $F(x_1) \leq F(x_2)$ sì
- $F_x(X)$ ammette limite da sinistra $\forall x$ sì
- $F_x(X)$ è continua da destra $\forall x$ sì



È la funzione di ripartizione di una variabile discreta o continua? È la funzione di ripartizione di una variabile continua, infatti quella di una variabile discreta è “a salti”, cioè del tipo:



In generale si ha che:

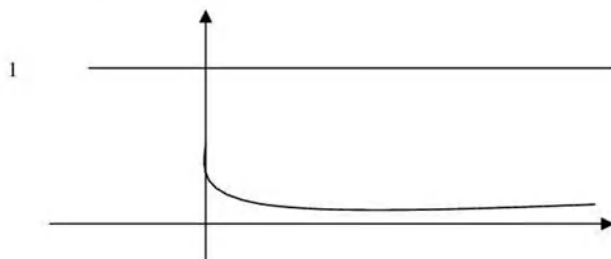
La funzione di ripartizione di una variabile continua è sempre derivabile e quindi anche continua.

Calcoliamo la funzione di densità:

$$F'_x(X) = \frac{dF_x(X)}{dX} = f_x(X)$$

da cui si ottiene:

$$F_x(X) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$



Ad esempio quindi si avrà che:

$$p(1 < X < 2) = \int_1^2 e^{-x} \cdot dx$$

2.4 – PROPRIETÀ DELLA FUNZIONE DI DENSITÀ

Una funzione di densità per essere tale deve rispettare le seguenti proprietà:

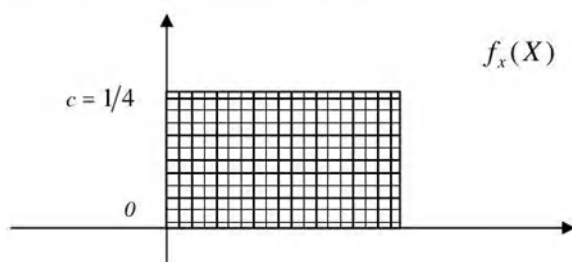
- 1) $f_x(X) \geq 0$ altrimenti le probabilità che sono aree e per cui necessariamente vale $0 \leq p \leq 1$, sarebbero negative

2) $p(X \in R) = p(-\infty < X < \infty) = 1$ cioè $\int_{-\infty}^{+\infty} f(x) \cdot dx = 1$

ovvero tutta l'area sottostante la funzione di densità deve essere pari a 1.

Esempio

Dobbiamo pesare un oggetto che può avere un qualsiasi peso X , sappiamo solo che $X \leq 4$, quindi, dato che un peso non può essere negativo, sappiamo che $x \in [0, 4]$. Come si sceglie la funzione di densità se tutti i valori dell'intervallo sono equipossibili? Se considero gli intervalli di ugual grandezza, ad esempio $0 < X < 0,1$ e $0,1 < X < 0,2$, essi hanno la stessa probabilità (a tutti i valori è associata la stessa probabilità di verificarsi).



Verifichiamo se sono rispettate le proprietà della funzione di densità

- 1) $f_x(X) \geq 0$ sì

2) $4 \cdot c = 1 \rightarrow c = \frac{1}{4}$ sì

Infatti:

$$\int_0^4 c \cdot dx = [cx]_0^4 = 4c$$

$$\int_0^4 c \cdot dx = 1 \rightarrow 4c = 1 \rightarrow c = \frac{1}{4}$$

Pertanto la funzione di densità assume la forma seguente:

$$f_x(X) = \begin{cases} 0 & x < 0 \\ c = \frac{1}{4} & 0 \leq x \leq 4 \\ 0 & x > 4 \end{cases}$$

E la funzione di ripartizione assume la forma seguente:

$$F_x(X) = \begin{cases} 0 & x < 0 \\ \int_0^x \frac{1}{4} = \frac{1}{4}x & 0 \leq x \leq 4 \\ 1 & x > 4 \end{cases}$$

Infine si definisce supporto della variabile casuale continua X:

$$S_x = \{X : f_x(X) > 0\}$$

2.1 – ESERCIZIO SULLE DISTRIBUZIONI BINOMIALE ED IPERGEOMETRICA

È data una popolazione di 100 individui (quindi $N = 100$), di cui 40 sono fumatori ($k = 40$) e i restanti non sono fumatori ($N - k = 60$). Viene estratto un campione di 4 individui ($n = 4$).

Calcolare la probabilità che nel campione estratto vi sia almeno 1 fumatore su 4.

a) Estrazione con reinserimento

X è il numero di fumatori nel campione. Dobbiamo trovare il numero di successi delle prove Bernoulliane. La variabile casuale X segue la distribuzione **binomiale**. Calcoliamo quindi $p(X \geq 1)$.

Possiamo calcolare questa probabilità come:

$$p(X \geq 1) = p(X = 1) + p(X = 2) + p(X = 3) + p(X = 4)$$

ma è più comodo calcolare la suddetta probabilità come:

$$p(X \geq 1) = 1 - p(X < 1)$$

ma $p(X < 1) = p(X = 0)$, quindi:

$$p(X \geq 1) = 1 - p(X = 0)$$

Abbiamo detto che $X \sim Bi(n, \theta)$. In questo caso abbiamo $n = 4$, $\theta = 40/100 = 0,4$. Quindi:

$$X \sim Bi(4, 0,4)$$

Calcoliamo quindi $p(X \geq 1) = 1 - p(X = 0)$:

$$p(X \geq 1) = 1 - p(X = 0) = 1 - \binom{n}{x} \cdot \theta^x \cdot (1 - \theta)^{n-x} = 1 - \binom{4}{0} \cdot (0,4)^0 \cdot (1 - 0,4)^4 = 1 - 0,1296 = 0,8704 \rightarrow 87,04\%$$

b) Estrazione senza reinserimento

Se la distribuzione è senza reinserimento, cambia la distribuzione della variabile casuale X , che è il numero di fumatori nel campione. In questo caso si ha che $X \sim Ig(N, k, n)$, ovvero una distribuzione **ipergeometrica**, dove N è la numerosità complessiva (nel nostro caso è 100), n è il numero di fumatori nel campione. Calcoliamo quindi $p(X \geq 1)$.

Abbiamo detto che $X \sim Ig(N, k, n)$. In questo caso abbiamo $n = 4$, $N = 100$, $k = 40$. Quindi:

$$X \sim Ig(100, 40, 4)$$

Calcoliamo quindi $p(X \geq 1) = 1 - p(X = 0)$:

$$\begin{aligned} p(X \geq 1) &= 1 - \frac{\binom{k}{x} \cdot \binom{N-k}{n-x}}{\binom{N}{n}} = 1 - \frac{\binom{40}{0} \cdot \binom{60}{4}}{\binom{100}{4}} = 1 - \frac{\frac{40!}{0!} \cdot \frac{60!}{4!}}{\frac{100!}{96!4!}} = 1 - \frac{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56!}{100 \cdot 99 \cdot 98 \cdot 97 \cdot 96!} = \\ &= 1 - \frac{60 \cdot 59 \cdot 58 \cdot 57}{100 \cdot 99 \cdot 98 \cdot 97} = 1 - 0,1244 = 0,8756 \rightarrow 87,56\% \end{aligned}$$

c) C'è qualche ragione per cui la binomiale è circa l'ipergeometrica? Se sì, a quale aspetto del problema è legato?

La risposta è affermativa, anche se non succede comunque sempre! Il motivo è che N è grande rispetto a n .

d) Qual è il numero medio di fumatori nel campione?

$$E[X] = n\theta = 0,4 \cdot 4 = 1,6 \text{ sia nel caso della distribuzione binomiale, che dell'ipergeometrica.}$$

e) Qual è la varianza della variabile casuale X ?

Nel caso di estrazione con reinserimento, abbiamo:

$$V[X] = n \cdot g \cdot (1 - g) = 4 \cdot 0.4 \cdot 0.6 = 0.96$$

Nel caso di estrazione senza reinserimento, abbiamo:

$$V[X] = n \cdot g \cdot (1 - g) \cdot \frac{N - n}{n - 1} = 4 \cdot 0.4 \cdot 0.6 \cdot \frac{100 - 4}{100 - 1} = 0.93$$

2.2 – ESERCIZIO SULLA DISTRIBUZIONE DI POISSON

Tra le 9.00 e le 10.00 di un giorno feriale, il numero medio di telefonate ad un centralino è di 2,5 telefonate al minuto (è la nostra λ).

a) Qual è la probabilità che in un minuto arrivino 0 telefonate?

$$p(X = 0) = e^{-\lambda} \cdot \frac{\lambda^x}{x!} = e^{-2.5} \cdot \frac{(2.5)^0}{0!} = 0.0821 \rightarrow 8.21\%$$

b) Qual è la probabilità che in un minuto arrivino 2 telefonate?

$$p(X = 2) = e^{-\lambda} \cdot \frac{\lambda^x}{x!} = e^{-2.5} \cdot \frac{(2.5)^2}{2!} = 0.2565 \rightarrow 25.65\%$$

c) Qual è la probabilità che in due minuti arrivino 0 telefonate?

$$\lambda = \frac{2.5}{\text{min}} \cdot 2 \text{ min} = 5$$

$$p(X = 0) = e^{-\lambda} \cdot \frac{\lambda^x}{x!} = e^{-5} \cdot \frac{(5)^0}{0!} = 0.67\%$$

d) Qual è la probabilità che in un minuto arrivino al più 4 telefonate?

$$p(X \leq 4) = p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) + p(X = 4)$$

si procede in modo analogo.

e) Qual è la probabilità che in un minuto arrivino più di 6 telefonate?

$$p(X > 6) = 1 - [p(X \leq 4) + p(X = 5)]$$

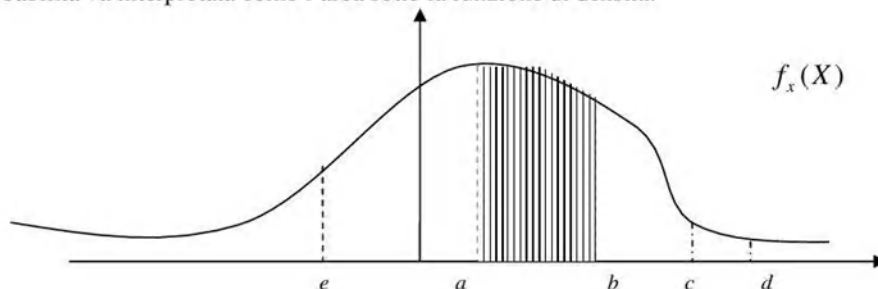
si procede in modo analogo.

3.1 – RIEPILOGO VARIABILI CONTINUE, FUNZIONE DI DENSITÀ ED ESEMPIO

Nelle variabili continue i possibili valori non sono elencabili, quindi la variabile casuale X avrà come possibili valori, tutti i valori di un intervallo. Quindi in generale $X \in [a, b]$. Nel caso continuo parliamo di **funzione di densità** $f_x(X)$. La probabilità che $X \in [a, b]$ è data da:

$$p(a < X < b) = \int_a^b f_x(X) \cdot dx$$

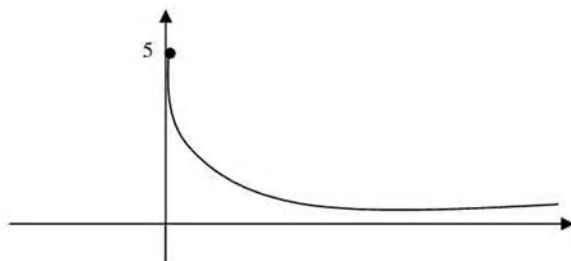
La probabilità va interpretata come l'area sotto la funzione di densità.



Esempio

È data la seguente funzione di densità:

$$f_x(X) = \begin{cases} 5e^{-5x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$



Verifichiamo innanzitutto se la funzione è una funzione di densità: per essere tale deve rispettare le seguenti proprietà:

1) $f_x(X) \geq 0$

SI

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

Verifichiamo la seconda proprietà di una funzione di densità:

$$\int_{-\infty}^{+\infty} f(x) \cdot dx = \int_{-\infty}^0 0 \cdot dx + \int_0^{+\infty} 5e^{-5x} \cdot dx = \left[-e^{-5x} \right]_0^{+\infty} = 1$$

SI

Si conclude pertanto che la funzione è una funzione di densità.

Calcolare $p(X > 5)$ e $p(2 < X < 4)$:

$$p(X > 5) = \int_5^{+\infty} 5e^{-5x} \cdot dx = \left[-e^{-5x} \right]_5^{+\infty} = e^{-25}$$

$$p(2 < X < 4) = \int_2^4 5e^{-5x} \cdot dx = -e^{-20} + e^{-10}$$

3.2 – CONFRONTO TRA VARIABILI DISCRETE E VARIABILI CONTINUE

Variabili discrete

- Possibili valori sono elencabili, cioè $S_x = \{x_1, x_2, x_3, \dots\}$
- Si utilizza la *funzione di probabilità*, cioè $p(X = x_i) = p_x(x_i)$ che per essere tale deve rispettare le seguenti proprietà:
 1. $p_x(x_i) \geq 0$
 2. $\sum_i p_x(x_i) = 1$
- Vale che, con $x_i \in [a, b]$
$$p(a < X < b) = \sum_i p_x(x_i)$$
- Il valore atteso o media è dato da: $\mu = E[X] = \sum_i x_i \cdot p_x(x_i)$
- La varianza è data da: $V[X] = V[(x - \mu)^2] = E[X^2] - (E[X])^2$

Variabili continue

- Possibili valori non sono elencabili, ma si considera un intervallo di valori, cioè $S_x = [a, b]$
- Si utilizza la *funzione di densità* $f_x(X)$, cioè:

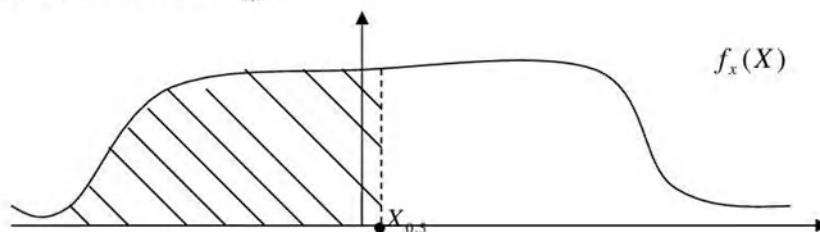
$$F_x(X) = p(X \leq x) = \int_{-\infty}^x f_x(X) \cdot dx \text{ che per essere tale deve rispettare le seguenti proprietà:}$$

1. $f_x(X) \geq 0$
 2. $\int_{-\infty}^{+\infty} f(x) \cdot dx = 1$
- Vale che:
$$p(a < X < b) = \int_a^b f_x(X) \cdot dx$$
 - Il valore atteso o media è dato da: $\mu = E[X] = \int_{-\infty}^{+\infty} x \cdot f_x(X) \cdot dx$
 - La varianza è data da: $V[X] = \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot f_x(X) \cdot dx = E[X^2] - (E[X])^2$

Si noti che il significato della media e della varianza è lo stesso sia per le variabili discrete che per le variabili continue.

3.3 – MEDIANA

Intuitivamente la mediana è il punto che taglia a metà la nostra distribuzione, il punto cioè che sta nel mezzo, che indicheremo con $X_{0,5}$.



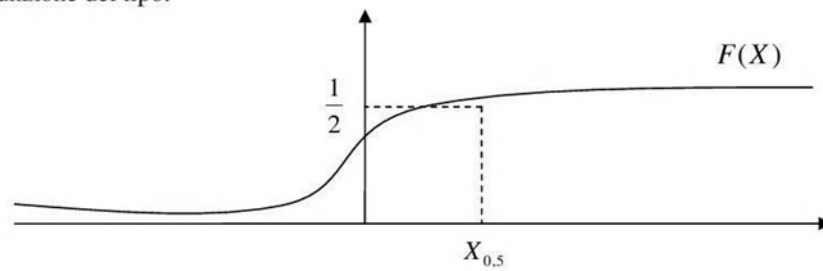
$$F(X) = p(X \leq x) = \frac{1}{2}$$

ovvero:

$$F(X_{0,5}) = \frac{1}{2}$$

Esempio

Se ho una funzione del tipo:



vado a cercare quel valore di x tale che:

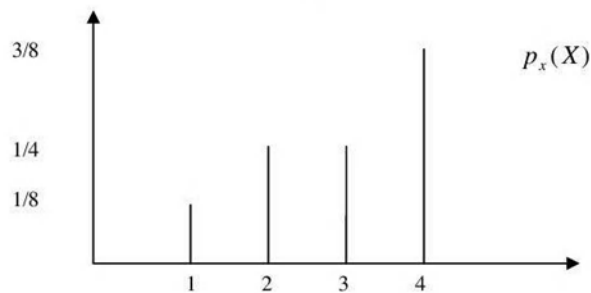
$$F(X) = \frac{1}{2}$$

Esempio

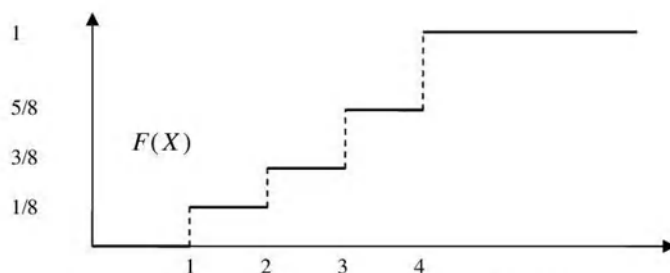
Se però ho una variabile casuale discreta tale che:

X	1	2	3	4
$p(X)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{3}{8}$

che pertanto ha un grafico del tipo:



non riusciamo a trovare il valore della mediana con la definizione utilizzata, infatti la sua funzione di ripartizione ha grafico:



Nel grafico non c'è nessun valore di x tale che $F(X) = 0,5$.

Si conclude che la definizione di mediana come:

$$F(X) = p(X \leq x) = \frac{1}{2}$$

è **valida solo per le variabili casuali continue**.

Definisco allora la mediana in modo più generale, che funziona sia per le **variabili casuali continue** che le **variabili casuali discrete**:

$$p(X < X_{0,5}) \leq \frac{1}{2} \leq p(X \leq X_{0,5})$$

Nel caso continuo, in particolare, vale sempre che:

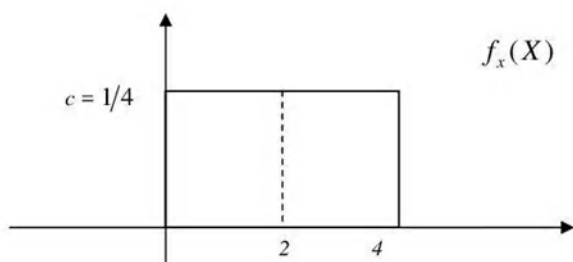
$$p(X < X_{0,5}) = p(X \leq X_{0,5}) = \frac{1}{2}$$

Nel nostro esempio, pertanto, la mediana risulta pari a 3.

Esempio

Consideriamo la variabile casuale continua $X \in [0,4]$. La sua funzione di densità $f_x(X)$ è tale che:

$$f_x(X) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leq x \leq 4 \\ 0 & x > 4 \end{cases}$$



$$X_{0,5} = 2$$

che si ottiene risolvendo la seguente equazione in funzione di x , sostituendo ad $F(X)$ la relazione con la funzione di densità:

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

Esempio

Considero la seguente funzione di densità:

$$f_x(X) = \begin{cases} 5e^{-5x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Voglio calcolare la mediana e quindi risolvo la seguente equazione in funzione di x :

$$F(X) = 0,5$$

Calcoliamo innanzitutto $F(X)$, sfruttando la relazione con la funzione di densità:

$$F(X) = p(X \leq x) = \int_{-\infty}^x f_u(U) \cdot du = \int_0^x 5e^{-5u} \cdot du = \left[-e^{-5u} \right]_0^x = 1 - e^{-5x} \quad \text{con } x > 0$$

A questo punto troviamo la x , sfruttando l'equazione scritta sopra:

$$1 - e^{-5x} = 0,5$$

$$-e^{-5x} = 0,5 - 1$$

$$-e^{-5x} = -0,5$$

$$e^{-5x} = 0,5$$

$$\ln e^{-5x} = \ln(0,5)$$

$$-5x = \ln(0,5)$$

da cui si ottiene il valore della mediana:

$$x = \frac{1}{5} \cdot \ln 2$$

3.4 – QUANTILE

Vogliamo calcolare x_q , dove $0 \leq q \leq 1$. In generale un quantile si definisce come:

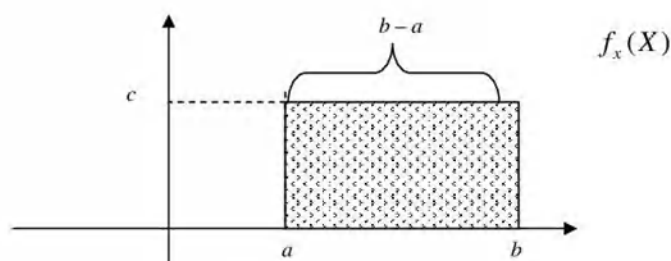
$$p(X < x_q) \leq q \leq p(X \leq x_q)$$

La definizione è valida sia per le *variabili casuali continue* che le *variabili casuali discrete*.

3.5 – VARIABILI CASUALI CONTINUE: DISTRIBUZIONE UNIFORME (O RETTANGOLARE)

Nella distribuzione uniforme (o rettangolare) la variabile $X \in (a, b)$. La funzione di densità $f_x(X)$ della distribuzione è della forma:

$$f_x(X) = \begin{cases} 0 & x \leq a \\ c & a < x < b \\ 0 & x \geq b \end{cases}$$



Per essere una funzione di densità devono essere rispettate le seguenti proprietà:

$$1) f_x(X) \geq 0$$

SI

$$2) \int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

Perché valga la seconda proprietà, si conclude che:

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

Valore atteso in una distribuzione uniforme

$$E[X] = \int_a^b x \cdot \frac{1}{b-a} \cdot dx = \frac{1}{b-a} \cdot \left[\frac{x^2}{2} \right]_a^b = \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{1}{b-a} \cdot \frac{(b-a) \cdot (b+a)}{2} = \frac{(b+a)}{2}$$

Pertanto il valore atteso in una distribuzione uniforme è:

$$E[X] = \frac{(b+a)}{2}$$

Varianza in una distribuzione uniforme

$$V[X] = \int_a^b (x - E[X])^2 \cdot \frac{1}{b-a} \cdot dx = E[X^2] - (E[X])^2$$

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

Pertanto la varianza in una distribuzione uniforme è:

$$V[X] = \frac{(b-a)^2}{12}$$

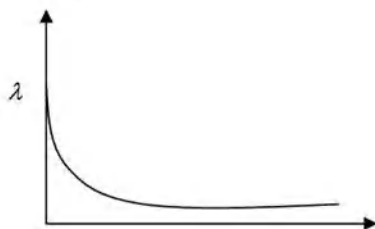
3.6 – VARIABILI CASUALI CONTINUE: DISTRIBUZIONE ESPONENZIALE

La distribuzione esponenziale si usa per **calcolare i tempi di attesa**. Supponiamo di avere un fenomeno di *arrivi casuali*, che è descritto, come abbiamo visto, dalla distribuzione di Poisson. Il numero di arrivi in un'unità di tempo $N(0,1)$ è una distribuzione $Po(\lambda)$, dove λ è il numero medio di arrivi in un intervallo di lunghezza unitaria.

Ma *quanto tempo passa da un arrivo all'altro*? Il tempo è una variabile continua ed ha una distribuzione esponenziale. Diremo che $X \sim Esp(\lambda)$.

La funzione di densità $f_x(X)$ della distribuzione è della forma:

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



Calcoliamo ora la funzione di ripartizione $F(X)$ della distribuzione per $x > 0$ (se $x \leq 0$ la funzione infatti è sempre 0):

$$F(X) = p(X \leq x) = \int_0^x \lambda \cdot e^{-\lambda u} \cdot du = \left[-e^{-\lambda u} \right]_0^x = 1 - e^{-\lambda x}$$

Valore atteso in una distribuzione esponenziale (Integro per parti)

$$E[X] = \int_0^{+\infty} x \cdot \lambda \cdot e^{-\lambda x} \cdot dx = \left[-x \cdot e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} \cdot dx = 0 + 0 + \left[-\frac{1}{\lambda} \cdot e^{-\lambda x} \right]_0^{+\infty} = 0 + \frac{1}{\lambda} = \frac{1}{\lambda}$$

Pertanto il valore atteso in una distribuzione esponenziale è:

$$E[X] = \frac{1}{\lambda}$$

Varianza in una distribuzione esponenziale (Integro per parti)

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

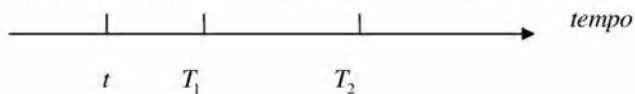
$$V[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Pertanto la varianza in una distribuzione esponenziale è:

$$V[X] = \frac{1}{\lambda^2}$$

Legame Poisson – Esponenziale

Consideriamo la linea del tempo e su di essa segniamo i tempi di alcuni arrivi:



$$p(T_1 \leq t) = F(T_1)$$

$$p(T_1 > t) = 1 - F(T_1)$$

Sappiamo che il numero di arrivi in un'unità di tempo $N(0,1) \sim Po(\lambda)$, dove λ è il numero medio di arrivi in un intervallo di lunghezza unitaria. In generale, quindi, varrà:

$$N(0,t) \sim Po(\lambda \cdot t)$$

dove λ è il numero medio di arrivi in un intervallo di lunghezza t .

In generale, quindi, avremo:

$$p(N(0,t) = x) = e^{-\lambda t} \cdot \frac{(\lambda \cdot t)^x}{x!}$$

se $x = 0$, allora diventa:

$$p(N(0,t) = 0) = e^{-\lambda t}$$

da cui:

$$F(T_1) = 1 - p(T_1 > t) = p(T_1 \leq t) = 1 - p(N(0,t) = 0) = 1 - e^{-\lambda t}$$

$$p(T_1 > t) = e^{-\lambda t}$$

che coincide con la funzione di ripartizione dell'esponenziale. Si conclude che T_1 è una distribuzione esponenziale di parametro λ , e anche T_2 e in generale tutti i tempi di attesa. Si dimostra inoltre che tutti i tempi di attesa sono tra loro indipendenti.

Proprietà della distribuzione esponenziale

La distribuzione esponenziale è l'unica variabile continua **senza memoria**

$$p(X > s+t | X > t) = \frac{p((X > s+t) \cap (X > t))}{p(X > t)} = \frac{p(X > s+t)}{p(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s}$$

ovvero la probabilità di aspettare un ulteriore tempo s dato che si è aspettato il tempo t , non dipende da t . Il tempo che devo aspettare ancora ha la forma di una distribuzione esponenziale.

$$p(X > x) = \int_x^{+\infty} \lambda \cdot e^{-\lambda u} \cdot du = \left[-e^{-\lambda u} \right]_x^{+\infty} = e^{-\lambda x}$$

3.7 – FUNZIONE GAMMA

La funzione Gamma (Γ) si definisce:

$$\Gamma(\alpha) = \int_0^{+\infty} y^{\alpha-1} \cdot e^{-y} \cdot dy$$

con $\alpha > 0$.

Se $\alpha = 1$, si ottiene:

$$\Gamma(1) = \int_0^{+\infty} e^{-y} \cdot dy = 1$$

Quindi vale sempre che:

$$\Gamma(1) = 1$$

Supponiamo che $\alpha > 1$ e applichiamo la definizione ed effettuiamo una integrazione per parti.

$$\Gamma(\alpha) = \left[-e^{-y} \cdot y^{\alpha-1} \right]_0^{+\infty} + (\alpha-1) \cdot \int_0^{+\infty} y^{\alpha-2} \cdot e^{-y} \cdot dy = 0 + (\alpha-1) \cdot \Gamma(\alpha-1) = (\alpha-1) \cdot \Gamma(\alpha-1)$$

Per calcolare la funzione Gamma, si ricorre spesso quindi alla seguente formula ricorsiva, se $\alpha > 1$:

$$\Gamma(\alpha) = (\alpha-1) \cdot \Gamma(\alpha-1)$$

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

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

$$\Gamma(4) = 3 \cdot \Gamma(3) = 6$$

....

Se α è un numero intero, vale che:

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

Se α non è un numero intero, vale che:

$$\Gamma(\alpha) = (\alpha-1) \cdot (\alpha-2) \cdot (\alpha-3) \cdot \dots \cdot \Gamma(\varepsilon)$$

dove $\varepsilon \leq 1$. Ad esempio $\Gamma(3,6) = 2,6 \cdot 1,6 \cdot 0,6 \cdot \Gamma(0,6)$.

Si dimostra inoltre che vale:

$$\Gamma(0,5) = \sqrt{\pi}$$

3.8 – VARIABILI CASUALI CONTINUE: DISTRIBUZIONE GAMMA

La variabile casuale continua X , in una distribuzione Gamma, dipende dai parametri α e β , cioè scriveremo $X \sim Ga(\alpha, \beta)$. La funzione di densità $f_x(X)$ della distribuzione è della forma:

$$f_x(X, \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha-1} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

con $\alpha > 0$. Verifichiamo innanzitutto se la funzione è una funzione di densità: per essere tale deve rispettare le seguenti proprietà:

$$1) f_x(X) \geq 0$$

SI

$$2) \int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

Verifichiamo la seconda proprietà di una funzione di densità:

$$\int_{-\infty}^0 0 \cdot dx + \int_0^{+\infty} \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha-1} \cdot dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \int_0^{+\infty} e^{-y} \cdot \left(\frac{y}{\beta}\right)^{\alpha-1} \cdot \frac{1}{\beta} \cdot dy = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{1}{\beta^\alpha} \cdot \int_0^{+\infty} e^{-y} \cdot y^{\alpha-1} \cdot \frac{1}{\beta} \cdot dy = 1$$

Per risolvere l'integrale cambio la variabile da x a y , ponendo:

$$x = \frac{y}{\beta} \rightarrow \frac{dx}{dy} = \frac{1}{\beta}$$

SI

Si conclude pertanto che la funzione è una funzione di densità.

Valore atteso in una distribuzione Gamma

Considero $X \sim Ga(\alpha+1, \beta)$ e ottengo:

$$f_x(X, \alpha+1, \beta) = \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} \cdot e^{-\beta \cdot x} \cdot x^\alpha \quad x > 0$$

che sappiamo essere una funzione di densità e che pertanto rispetta le due proprietà fondamentali. In particolare, se sfruttiamo la seconda proprietà, sappiamo che:

$$\int_0^{+\infty} \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} \cdot e^{-\beta \cdot x} \cdot x^\alpha \cdot dx = 1$$

Calcoliamo ora il valore atteso:

$$\begin{aligned} E[X] &= \int_0^{+\infty} x \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha-1} \cdot dx = \int_0^{+\infty} \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^\alpha \cdot dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \cdot \int_0^{+\infty} \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} \cdot e^{-\beta \cdot x} \cdot x^\alpha \cdot dx = \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \cdot 1 = \frac{1}{\beta} \cdot \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \end{aligned}$$

ma sappiamo che $\Gamma(\alpha) = (\alpha-1) \cdot \Gamma(\alpha-1)$, quindi varrà anche che:

$$\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$$

Se sostituiamo, si ottiene:

$$E[X] = \frac{1}{\beta} \cdot \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{1}{\beta} \cdot \frac{\alpha \cdot \Gamma(\alpha)}{\Gamma(\alpha)} = \frac{\alpha}{\beta}$$

Pertanto il valore atteso in una distribuzione Gamma è:

$$E[X] = \frac{\alpha}{\beta}$$

Varianza in una distribuzione Gamma

Considero $X \sim Ga(\alpha+2, \beta)$ e ottengo:

$$f_x(X, \alpha + 2, \beta) = \frac{\beta^{\alpha+2}}{\Gamma(\alpha+2)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha+1} \quad x > 0$$

che sappiamo essere una funzione di densità e che pertanto rispetta le due proprietà fondamentali. In particolare, se sfruttiamo la seconda proprietà, sappiamo che:

$$\int_0^{+\infty} \frac{\beta^{\alpha+2}}{\Gamma(\alpha+2)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha+1} \cdot dx = 1$$

Calcoliamo ora il valore atteso:

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

$$\begin{aligned} V[X] &= \left(\int_0^{+\infty} x^2 \cdot \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha-1} \cdot dx \right) - \left(\frac{\alpha}{\beta} \right)^2 = \left(\int_0^{+\infty} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha+1} \cdot dx \right) - \left(\frac{\alpha}{\beta} \right)^2 = \\ &= \left(\frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} \cdot \int_0^{+\infty} \frac{\beta^{\alpha+2}}{\Gamma(\alpha+2)} \cdot e^{-\beta \cdot x} \cdot x^{\alpha+1} \cdot dx \right) - \left(\frac{\alpha}{\beta} \right)^2 = \left(\frac{1}{\beta^2} \cdot \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \cdot 1 \right) - \left(\frac{\alpha}{\beta} \right)^2 = \\ &= \frac{1}{\beta^2} \cdot \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} - \frac{\alpha^2}{\beta^2} \end{aligned}$$

ma sappiamo che $\Gamma(\alpha) = (\alpha-1) \cdot \Gamma(\alpha-1)$, quindi varrà anche che:

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

ma sappiamo anche che:

$$\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$$

Se sostituiamo, si ottiene:

$$V[X] = \frac{1}{\beta^2} \cdot \frac{(\alpha+1) \cdot \Gamma(\alpha+1)}{\Gamma(\alpha)} - \frac{\alpha^2}{\beta^2} = \frac{1}{\beta^2} \cdot \frac{(\alpha+1) \cdot \alpha \cdot \Gamma(\alpha)}{\Gamma(\alpha)} - \frac{\alpha^2}{\beta^2} = \frac{\alpha^2 + \alpha - \alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$$

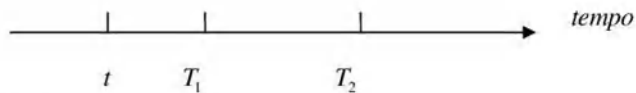
Pertanto la varianza in una distribuzione Gamma è:

$$V[X] = \frac{\alpha}{\beta^2}$$

Legame Esponenziale – Gamma

Se $\alpha = 1$, abbiamo che:

$$Ga(1, \beta) \sim Esp(\beta)$$



Quindi avremo:

$$T_1 \sim Esp(\lambda) \sim Ga(1, \lambda)$$

ma vale anche:

$$T_1 + T_2 + T_3 \sim Ga(3, \lambda)$$

da cui, generalizzando:

$$T_1 + T_2 + T_3 + \dots + T_n \sim Ga(n, \lambda)$$

La distribuzione Gamma non misura il tempo di attesa fra un arrivo e l'altro ma tra n arrivi.

Esempio

A partire dalle ore 9.00 iniziano ad arrivare i clienti. Sappiamo inoltre che i clienti arrivano con un'intensità costante. L'intensità media è di 5 clienti per minuto (cioè $\lambda = 5$). Calcolare la probabilità che il decimo cliente arrivi dopo le 9.30. X è il tempo di arrivo del decimo cliente. T_i è il tempo di attesa tra un cliente e l'altro.

Sappiamo che $T_i \sim \text{Esp}(5)$ ma vogliamo trovare la probabilità che il decimo cliente arrivi dopo le 9.30, quindi $T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8 + T_9 + T_{10} \sim \text{Ga}(10, 5)$.

Vogliamo quindi calcolare $p(X > 30)$, quindi:

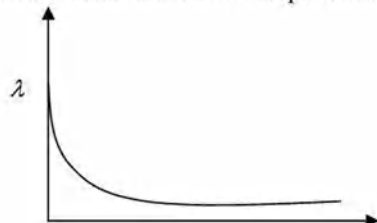
$$p(X > 30) = \int_{30}^{+\infty} \frac{5^{10}}{\Gamma(10)} \cdot e^{-5x} \cdot x^9 \cdot dx$$

Rappresentazione della funzione di densità di Γ

Vogliamo capire come è fatta, dal punto di vista grafico, la funzione di densità:

$$f_x(X, \alpha + 2, \beta) = \frac{\beta^{\alpha+2}}{\Gamma(\alpha+2)} \cdot e^{-\beta x} \cdot x^{\alpha+1}$$

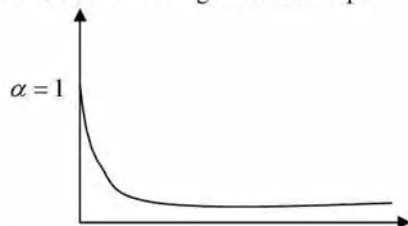
Nel caso della distribuzione esponenziale, avevamo un grafico del tipo:



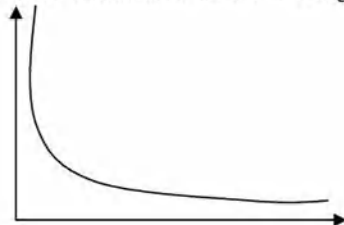
Nel caso della distribuzione Gamma, la faccenda è più complicata, in quanto intervengono anche i parametri α e β ad influenzare il grafico. In particolare abbiamo che:

- α è un **parametro di scala**, perché la sua variazione semplicemente allunga o accorcia l'asse delle x (cambia quindi solo la scala e non la forma della funzione di densità)
- β è un **parametro di forma**, perché la sua variazione cambia la forma della funzione di densità. In particolare abbiamo:

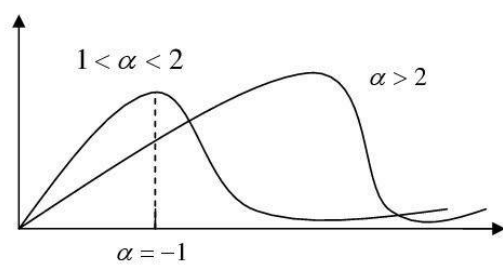
1. se $0 < \alpha < 1$ il grafico è del tipo:



2. se $\alpha < 1$ la funzione va a $+\infty$ e il grafico è del tipo:



3. se $1 < \alpha < 2$ e se $\alpha > 2$ abbiamo grafici del tipo:



3.1 – ESERCIZIO SULLE VARIABILI CASUALI CONTINUE

È data la funzione di densità $f_x(X)$, così definita:

$$f_x(X) = \begin{cases} 0 & x \leq 1 \\ \frac{1}{x} & 1 < x < k \\ 0 & x \geq k \end{cases}$$

N.B. Se viene indicato che una funzione assume un determinato valore in un intervallo, automaticamente (se non è indicato altrimenti) negli altri intervalli la funzione ha valore 0.

1) Determinare k affinché $f_x(X)$ sia una funzione di densità:

a) $f_x(X) \geq 0$ sì

b) $\int_1^k f_x(X) \cdot dx = 1$ sì

infatti $\int_1^k \frac{1}{x} \cdot dx = [\ln|x|]_1^k = \ln k - \ln 1 = \ln k \Rightarrow \ln k = 1 \Rightarrow e^{\ln k} = e^1 \rightarrow k = e \cong 2,72$

2) Determinare la funzione di ripartizione della variabile casuale X , che ha la funzione di densità $f_x(X)$ sopra definita:

$$F(X) = p(X \leq x) = \int_{-\infty}^x f_t(t) \cdot dt = \int_1^x \frac{1}{t} \cdot dt = [\ln|t|]_1^x = \ln x \quad \text{con } x > 1$$

Cosa succede se $x < 1$? Qual è, ad esempio, il valore che assume la funzione di ripartizione in $1/2$?

$$F(1/2) = \int_{-\infty}^{1/2} f_t(t) \cdot dt = \int_{-\infty}^{1/2} 0 \cdot dt = 0$$

In generale, per qualsiasi valore $x < 1$, otteniamo che la funzione di ripartizione vale 0.

Cosa succede se $x > 1$ e $x > e$? Qual è, ad esempio, il valore che assume la funzione di ripartizione in $e + 1$?

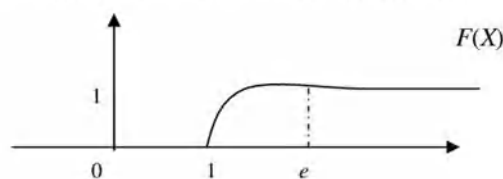
$$F(e+1) = \int_{-\infty}^{e+1} f_t(t) \cdot dt = \int_{-\infty}^1 0 \cdot dt + \int_1^e \frac{1}{t} \cdot dt + \int_e^{e+1} 0 \cdot dt = 1$$

In generale, per qualsiasi valore $x > e$, otteniamo che la funzione di ripartizione vale 1.

Pertanto la funzione di ripartizione $F(X)$ assume la forma:

$$F(X) = \begin{cases} 0 & x \leq 1 \\ \ln x & 1 < x < e \\ 1 & x \geq e \end{cases}$$

N.B. Nelle variabili continue è indifferente mettere l'uguale in un intervallo o un altro.



3) Calcolare $p(X \leq 2)$:

$$p(X \leq 2) = F(2) = \ln 2$$

4) Calcolare $p(1/2 < X < 2)$:

$$p(1/2 < X < 2) = \int_{1/2}^2 f(t) \cdot dt = \int_{1/2}^1 0 \cdot dt + \int_1^2 \frac{1}{t} \cdot dt = \ln 2 - \ln 1 = \ln 2$$

5) Calcolare $p(e/2 < X < 2)$:

$$p(e/2 < X < 2) = F(2) - F\left(\frac{e}{2}\right) = \ln 2 - \ln\left(\frac{e}{2}\right) = \ln\left(2 \cdot \frac{2}{e}\right) = \ln\left(\frac{4}{e}\right) = \ln 4 - \ln e = \ln 4 - 1 = 0,386$$

6) Calcolare la mediana, cioè quel valore di x che bipartisce in due la distribuzione, tale che:

$$F(X) = 1/2$$

quindi si ottiene:

$$\ln x = 1/2$$

$$x = e^{1/2} = 1,648$$

7) Calcolare media e varianza

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f_x(X) \cdot dx = \int_1^e x \cdot \frac{1}{x} \cdot dx = [x]_1^e = e - 1 = 1,718$$

$$V[X] = E[X^2] - (E[X])^2 = \left[\frac{x^2}{2} \right]_1^e - (e-1)^2 = \frac{e^2}{2} - \frac{1}{2} - e^2 + 1 - 2e = 0,243$$

3.2 – ESERCIZIO SULLE VARIABILI CASUALI CONTINUE

È data la funzione di densità $f_x(X)$, così definita:

$$f_x(X) = \begin{cases} 0 & x \leq -1 \\ \frac{k}{1+x^2} & -1 < x < 1 \\ 0 & x \geq 1 \end{cases}$$

1) Determinare k affinché $f_x(X)$ sia una funzione di densità:

a) $f_x(X) > 0$ sì

b) $\int_{-\infty}^{+\infty} f_x(X) \cdot dx = 1$ sì

$$\text{infatti } \int_{-1}^1 \frac{k}{1+x^2} \cdot dx = k \cdot \int_{-1}^1 \frac{1}{1+t^2} \cdot dt = k \cdot [\text{artg}(t)]_{-1}^1 = k \cdot (\text{artg}(1) - \text{artg}(-1)) = k \cdot \left(\frac{\pi}{4} + \frac{\pi}{4} \right) = k \cdot \frac{\pi}{2}$$

$$\Rightarrow k \cdot \frac{\pi}{2} = 1 \Rightarrow k = \frac{2}{\pi} = 0,637$$

2) Determinare la funzione di ripartizione della variabile casuale X , che ha la funzione di densità $f_x(X)$ sopra definita:

$$F(X) = p(X \leq x) = \frac{2}{\pi} \cdot \int_1^x \frac{1}{1+t^2} \cdot dt = \frac{2}{\pi} \cdot [\text{artg}(t)]_1^x = \frac{2}{\pi} \cdot (\text{artg}(x) - \text{artg}(1)) = \frac{2}{\pi} \cdot \left(\text{artg}(x) + \frac{\pi}{4} \right)$$

con $-1 < x < 1$

In generale, per qualsiasi valore $x < -1$, otteniamo che la funzione di ripartizione vale 0.

In generale, per qualsiasi valore $x > 1$, otteniamo che la funzione di ripartizione vale 1.

Pertanto la funzione di ripartizione $F(X)$ assume la forma:

$$F(X) = \begin{cases} 0 & x \leq -1 \\ \frac{2}{\pi} \cdot (\text{artg}(x)) + \frac{1}{2} & -1 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

- 3) Calcolare la moda, cioè il valore massimo che assume la nostra funzione.
Calcoliamo pertanto la derivata della funzione di densità e poniamola uguale a zero.

$$f'(X) = \frac{2}{\pi} \cdot \left(-\frac{2x}{(1+x^2)^2} \right) = 0$$

Studiamo il segno:

$$f'(X) = 0 \Rightarrow x = 0$$

$$f'(X) > 0 \Rightarrow x < 0$$

$$f'(X) < 0 \Rightarrow x > 0$$

Pertanto $x = 0$ è punto di massimo della funzione, cioè la moda (che è un massimo assoluto).

- 4) Calcolare la mediana:

$$F(X) = 1/2$$

quindi si ottiene:

$$\frac{2}{\pi} \cdot \text{artg}(x) + \frac{1}{2} = \frac{1}{2}$$

$$\frac{2}{\pi} \cdot \text{artg}(x) = 0$$

$$\text{artg}(x) = 0$$

$$x = 0$$

N.B. Si nota che nell'esercizio i valori di moda, mediana e media coincidono.

3.3 – ESERCIZIO SULLA DISTRIBUZIONE ESPONENZIALE NEGATIVA

È data la variabile casuale X , tale che $X \sim \text{En}(\lambda)$, Sappiamo inoltre che $E[X] = 2$.

Dato che $E[X] = \frac{1}{\lambda}$ in una distribuzione esponenziale, si ricava che $\lambda = \frac{1}{2}$.

- 1) Calcolare $p(X \leq 1,5)$:

In una distribuzione esponenziale si ha che: $f_x(X) = \lambda \cdot e^{-\lambda \cdot x}$ per $x, \lambda > 0$. Nel nostro caso, dato che $\lambda = 1/2$, si ottiene: $f_x(X) = (1/2) \cdot e^{-(1/2)(1,5)}$. In una distribuzione esponenziale,

poi, si ha che $F(X) = p(X \leq x) = \int_0^x \lambda \cdot e^{-\lambda \cdot u} \cdot du = \left[-e^{-\lambda \cdot u} \right]_0^x = 1 - e^{-\lambda \cdot x}$. Pertanto, nel nostro

caso, otteniamo che $F(1,5) = -e^{-(3/4)} + 1 = 0,5276$

- 2) Un transistor funziona ininterrottamente da due anni, qual è la probabilità che restista per un altro anno (occorre ricordare che l'esponenziale è una distribuzione senza memoria oppure $r=2, s=1$). Devo quindi calcolare:

$$p(X \geq 2+1 | X \geq 2) = p(X \geq 3 | X \geq 2) = p(X \geq 1) = (e^{-\lambda})^1 = e^{-\lambda} = 1/e^{(1/2)} = 0,606$$

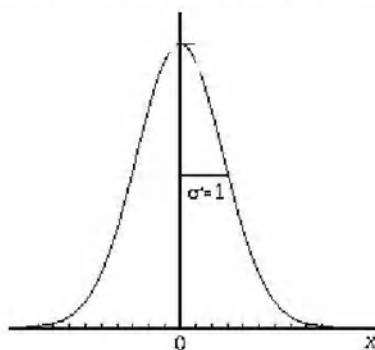
4.1 – VARIABILI CASUALI CONTINUE: DISTRIBUZIONE DI GAUSS O NORMALE STANDARD

La distribuzione di Gauss è in generale la distribuzione di variabili casuali continue più utilizzata. Nella distribuzione **normale standard**, la variabile casuale $Z \in (-\infty, +\infty)$, ha una funzione di densità così definita:

$$f_z(Z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} = \varphi(Z)$$

La curva descritta da tale funzione ha una forma caratteristica “a campana”:

Distribuzione normale “standardizzata”



Questa funzione di densità ha la seguente caratteristica:

$$\varphi(Z) = \varphi(-Z)$$

cioè un valore positivo di Z ha la stessa densità del valore negativo corrispondente. Infatti la curva di Gauss normale è **simmetrica** attorno allo 0.

Verifichiamo se $f_z(Z) = \varphi(Z)$ è una funzione di densità:

$$1) f_z(Z) > 0 \quad SI'$$

$$2) \int_{-\infty}^{+\infty} f(z) \cdot dz = 1 \quad SI'$$

infatti si dimostra che:

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} \cdot dz = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} \cdot dz = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1$$

Pertanto si conclude che $f_z(Z) = \varphi(Z)$ è una funzione di densità.

Come è fatta la funzione di ripartizione $F(Z)$ di una distribuzione di Gauss standardizzata?

$$F(Z) = p(Z \leq z) = \int_{-\infty}^z f(z) \cdot dz = \Phi(z)$$

infatti non riesco a trovare la primitiva della funzione di ripartizione, quindi l'integrale della funzione di densità viene denotato come $\Phi(z)$. Per calcolare la funzione di ripartizione in un qualsiasi punto z , utilizzerò le tavole apposite.

Valore atteso o media in una distribuzione di Gauss standardizzata

$$E[Z] = \int_{-\infty}^{+\infty} z \cdot f(z) \cdot dz$$

Questo integrale non è semplice da calcolare direttamente, allora spezzo l'integrale in più intervalli:

$$E[Z] = \int_{-\infty}^0 z \cdot f(z) \cdot dz + \int_0^{+\infty} z \cdot f(z) \cdot dz$$

se nel primo intervallo pongo $z = -z$, otteniamo:

$$E[Z] = \int_0^{+\infty} -z \cdot f(-z) \cdot dz + \int_0^{+\infty} z \cdot f(z) \cdot dz$$

da cui si ottiene, dato che $f(z) = f(-z)$, che il valore atteso è pari a:

$$E[Z] = - \int_0^{+\infty} z \cdot f(z) \cdot dz + \int_0^{+\infty} z \cdot f(z) \cdot dz = 0$$

In una distribuzione di Gauss standard, quindi, il valore atteso è pari a:

$$E[Z] = 0$$

In una distribuzione di Gauss standard, in particolare, vale sempre che:

$$Mo(Z) = Me(Z) = E[Z] = 0$$

Varianza in una distribuzione di Gauss standardizzata

$$V[Z] = E[Z^2] - (E[Z])^2 = 1 - 0 = 1$$

infatti si dimostra, integrando per parti, che:

$$E[Z^2] = \int_{-\infty}^{+\infty} z^2 \cdot f(z) \cdot dz = 1$$

In una distribuzione di Gauss standard, quindi, la varianza è pari a:

$$V[Z] = 1$$

Possiamo quindi concludere che in una distribuzione di Gauss standardizzata, la variabile casuale Z è così definita:

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

4.2 – VARIABILI CASUALI CONTINUE: DISTRIBUZIONE DI GAUSS O NORMALE

La distribuzione di Gauss **normale** è una generalizzazione della normale standard, infatti, la media e la varianza sono arbitrarie. Nella distribuzione normale, la variabile casuale X , che verrà tra poco definita è tale che:

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

Per trovare una variabile generale di questo tipo, faccio una trasformazione lineare di Z in questo modo:

$$X = \mu + Z\delta$$

da cui si ottiene che:

$$\begin{aligned} E[X] &= \mu + E[Z] \cdot \delta = \mu \\ V[X] &= \mu + V[Z] \cdot \delta = \delta^2 \end{aligned}$$

Nella formula sopra definita, risulta quindi che μ e δ^2 sono rispettivamente la media e la varianza della distribuzione.

Ho quindi costruito una nuova variabile X , che ha varianza e media arbitrarie. Come faccio a calcolare la distribuzione?

$$F(X) = p(X \leq x) = p(\mu + Z\delta \leq x) = p\left(Z \leq \frac{x-\mu}{\delta}\right) = \Phi\left(\frac{x-\mu}{\delta}\right)$$

Vogliamo a questo punto conoscere anche $f_x(X)$, ovvero la funzione di densità della variabile X .

Per calcolare la funzione di densità, basta fare la derivata della funzione di ripartizione di X sopra definita:

$$f(X) = \frac{d\Phi\left(\frac{x-\mu}{\delta}\right)}{dx} = \varphi\left(\frac{x-\mu}{\delta}\right) \cdot \frac{1}{\delta}$$

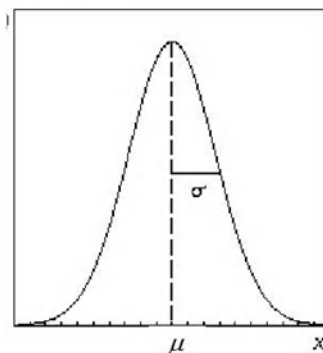
Pertanto, se consideriamo la definizione di prima, concludiamo che:

$$f_x(X) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\delta} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} = \varphi(X)$$

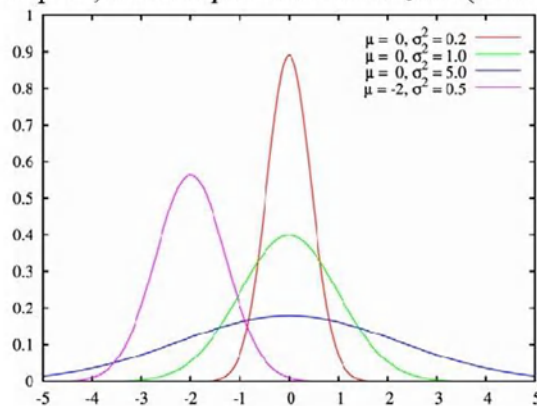
definita per $-\infty < x < +\infty$.

La forma della distribuzione di Gauss normale è analoga a quella della distribuzione normale standard; l'unica differenza è che anziché essere simmetrica rispetto allo zero, è simmetrica rispetto alla media μ .

Esempio di funzione di Gauss



In particolare, δ determina la **variabilità** della mia distribuzione ed è detto **parametro di scala 1** (cioè più è piccola, più la curva è stretta, e più è grande, più la curva è larga), mentre se vario μ , la mia funzione di densità si sposta, ed è detto **parametro di locazione** (viene traslata a dx o a sinistra).



In una distribuzione normale, dato che la funzione di densità è simmetrica, vale sempre che:

$$Mo(X) = Me(X) = E[X] = \mu$$

Cosa accade se prendo la distribuzione di Gauss normale X e faccio una trasformazione lineare?

Ad esempio, se considero una nuova variabile casuale W tale che:

$$W = aX + b$$

si ottiene, sostituendo quando detto prima, che:

$$W = a(\mu + Z\delta) + b = a\mu + aZ\delta + b$$

Pertanto la nuova variabile sarà tale che:

$$W \sim N(a\mu + b, a^2\delta^2)$$

infatti si ha che:

$$E[W] = aE[X] + b = a\mu + b$$

$$V[X] = aV[X] + b = a^2\delta^2$$

Per svolgere gli esercizi pratici, si riconduce sempre una distribuzione di Gauss normale ad una normale standard:

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

ma sappiamo anche che:

$$X = \mu + Z\delta$$

da cui si ottiene che:

$$Z = \frac{X - \mu}{\delta}$$

Pertanto, unendo le relazioni ottenute, ricaviamo che:

$$\frac{X - \mu}{\delta} \sim N(0,1)$$

da cui si ricava:

$$\begin{aligned} E\left[\frac{X - \mu}{\delta}\right] &= 0 \\ V\left[\frac{X - \mu}{\delta}\right] &= V\left[\frac{X}{\delta}\right] = \frac{V[X]}{\delta^2} = 1 \\ F(Z) = p(X \leq x) &= p\left(\frac{X - \mu}{\delta} \leq \frac{x - \mu}{\delta}\right) = p\left(Z \leq \frac{x - \mu}{\delta}\right) = \Phi\left(\frac{x - \mu}{\delta}\right) \end{aligned}$$

Esempi sulla distribuzione di Gauss standard

$$F(Z) = p(Z \leq 1,03) = \Phi(1,03) = 0,8485$$

Basta guardare il valore corrispondente a 1,03 sulle tavole

$$F(Z) = p(Z \leq -0,38) = 1 - p(Z \leq 0,38) = 1 - 0,648$$

infatti in generale vale sempre che:

$$p(Z \leq -z) = p(Z \geq z) = 1 - p(Z \leq z) = 1 - \Phi(Z) = \Phi(-Z)$$

$$p(-0,35 < Z < 1,96) = F(1,96) - F(-0,35) = 0,9750 - (1 - 0,6368) = 0,9750 - 1 + 0,6368 = 0,6118$$

$$p(Z \leq z) = \Phi(Z) = 0,9948$$

Prendo la tavola e cerco il valore di z tale che $\Phi(Z) = 0,9948$.

In questo caso si ha che $z = 2,56$

$$p(Z > z) = 0,0015$$

$$1 - p(Z > z) = 1 - 0,0015$$

$$p(Z \leq z) = 1 - 0,0015$$

$$p(-z \leq Z \leq z) = 0,6826$$

Per simmetria vale che:

$$p(0 \leq Z \leq z) = 0,3413$$

ma sappiamo che:

$$p(-\infty \leq Z \leq 0) = 0,5$$

quindi si ottiene che:

$$p(-\infty \leq Z \leq z) = 0,5 + 0,3413 = 0,8413$$

Quindi si ha che $z=1$, $-z=-1$

Esempi sulla distribuzione di Gauss normale

Consideriamo $X \sim N(2,1)$. Vogliamo calcolare $p(-1 \leq X \leq 1)$. Riconduciamo la distribuzione normale a quella standard:

$$Z = \frac{X - 2}{\sqrt{1}} \sim N(0,1)$$

Quindi in questo caso abbiamo:

$$Z = X - 2$$

Pertanto avremo:

$$p(-1 \leq X \leq 1) = p(-1 - 2 \leq X - 2 \leq 1 - 2) = p(-3 \leq Z \leq -1) = F(-1) - F(-3) = 0,1574$$

Consideriamo delle viti. La loro lunghezza (che chiamiamo X) è diversa. La variabile casuale X è una normale tale che $X \sim N(1,0.25)$. Se $X < 0,9$ oppure $X > 1,2$ la vite è scartata. Calcolare la percentuale di viti scartate. Devo quindi calcolare:

$$p((X < 0,9) \cup (X > 1,2)) = p(X < 0,9) + p(X > 1,2)$$

dato che gli eventi sono tra loro incompatibili

Riconduciamo la distribuzione normale a quella standard:

$$Z = \frac{X - 1}{\sqrt{0,25}} = \frac{X - 1}{0,5} \sim N(0,1)$$

Quindi in questo caso abbiamo:

$$Z = \frac{X - 1}{0,5}$$

Pertanto avremo:

$$p(X < 0,9) + p(X > 1,2) = p\left(Z < \frac{0,9 - 1}{0,5}\right) + p\left(Z > \frac{1,2 - 1}{0,5}\right) = p(Z < -0,2) + p(Z > 0,4) = 0,7653$$

Consideriamo $X \sim N(\mu, (6,5)^2)$. Sappiamo che l'intera popolazione ha un'altezza media $\mu = 173$.

Qual è l'intervallo di valori che contiene il 95% della popolazione? Vogliamo cioè calcolare: $p(x_1 \leq X \leq x_2) = 0,95$.

Riconduciamo la distribuzione normale a quella standard:

$$Z = \frac{X - 173}{6,5} \sim N(0,1)$$

Quindi in questo caso abbiamo:

$$Z = \frac{X - 173}{6,5}$$

Pertanto avremo:

$$p(X \leq x_2) = 0,5 + \frac{0,95}{2} = 0,975$$

(infatti la distribuzione è simmetrica rispetto alla media)

$$p\left(\frac{X - 173}{6,5} \leq \frac{x_2 - 173}{6,5}\right) = 0,975$$

Andiamo pertanto sulle tavole a cercare il valore corrispondente a 0,975:

$$\frac{x_2 - 173}{6,5} = 1,96$$

$$x_2 = 1,96 \cdot 6,5 + 173 = 186$$

da cui otteniamo:

$$x_2 - \mu = 186 - 173 = 13$$

$$x_1 = 173 - 13 = 160$$

4.3 – DISTRIBUZIONE DI UNA FUNZIONE DI UNA VARIABILE CASUALE

Data una variabile casuale X , di cui conosco la funzione di ripartizione $F_x(X)$ e la funzione di densità $f_x(X)$, come faccio a calcolare la distribuzione di una sua trasformata $Y = g(X)$?

Trasformazione lineare

Se $Y = a + bX$, otteniamo, come abbiamo visto per la distribuzione di Gauss, che:

$$F_y(Y) = p(Y \leq y) = p(a + bX \leq y) = p\left(X \leq \frac{Y - a}{b}\right) = F_x\left(\frac{Y - a}{b}\right) \quad \text{con } b > 0$$

$$f_y(Y) = \frac{dF_x\left(\frac{Y - a}{b}\right)}{dy} = f_x\left(\frac{Y - a}{b}\right) \cdot \frac{1}{b}$$

Trasformazioni biunivoche e derivabili

Possiamo generalizzare quanto detto sopra. È data una variabile casuale continua $X \in (a, b)$, che in quanto continua ammette una funzione di ripartizione e una funzione di densità. Supponiamo inoltre che $Y = g(X)$ sia una **trasformazione monotona** (quindi o crescente o decrescente) e **derivabile**, ma con derivata diversa da zero, ovvero:

$$\frac{dg(X)}{dX} \neq 0$$

Vogliamo calcolare la funzione di densità della nuova variabile. Si ottiene che:

$$f_y(Y) = f_x(g^{-1}(y)) \cdot \left| \frac{dg^{-1}(y)}{dy} \right|$$

Dimostrazione

Supponiamo che la funzione g sia crescente. Avremo una situazione tale che:

$$F_y(Y) = p(Y \leq y) = p(g(X) \leq y) = p(X \leq g^{-1}(y)) = F_x(g^{-1}(y))$$

$$f_y(Y) = \frac{dF_y(Y)}{dy} = \frac{dF_x(g^{-1}(y))}{dy} = f_x(g^{-1}(y)) \cdot \frac{d(g^{-1}(y))}{dy}$$

Dove vale la funzione di densità? Quali sono i possibili valori di Y ?

Se $X \in (a, b) \rightarrow Y \in (g(a), g(b))$

Possiamo dimostrare in maniera analoga se la funzione g è decrescente, usando però il valore assoluto, perché la quantità ottenuta sarebbe negativa. Se g è decrescente si ha che:
 $X \in (a, b) \rightarrow Y \in (g(b), g(a))$ Al di fuori dell'intervallo specificato, la funzione di densità vale 0.

Esempio

Consideriamo una variabile casuale $X \sim \text{Exp}(1)$ tale che:

$$f_x(X) = \begin{cases} e^{-x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

Si ha quindi che $X \in (0, +\infty)$.

Consideriamo una variabile casuale $Y = g(X) = \ln x$.

Dato che $X \in (0, +\infty)$, ne deriva che $Y \in (\ln 0, \ln(+\infty)) = (-\infty, +\infty)$.

Ricaviamo inoltre, visto che $Y = g(X) = \ln x$, che:

$$X = g^{-1}(y) = e^y$$

La trasformazione è monotona e derivabile, con derivata diversa da zero.

Calcoliamo ora la funzione di densità della nuova variabile Y :

$$f_y(Y) = f_x(g^{-1}(y)) \cdot \frac{d(g^{-1}(y))}{dy} = e^{-e^y} \cdot e^y$$

Trasformazioni non biunivoche

Consideriamo ora trasformazioni non biunivoche, cioè non monotone e di cui non si può quindi calcolare la funzione inversa. Ad esempio consideriamo $Y = X^2 = g(X)$. Vogliamo calcolare la funzione di ripartizione della nuova variabile casuale Y , data la variabile X e la sue funzioni di ripartizione e di densità:

$$F_y(Y) = p(Y \leq y) = p(g(X) \leq y) = p(X^2 \leq y) = p(-\sqrt{y} \leq X \leq \sqrt{y}) = F_x(\sqrt{y}) - F_x(-\sqrt{y})$$

Se a questo punto faccio la derivata della funzione di ripartizione di Y , ottengo anche la funzione di densità.

Esempio

Consideriamo $Y = \ln X$. Vogliamo calcolare la media della nuova variabile Y . In generale vale che $E[Y] = E[g(X)]$, quindi nel nostro esempio abbiamo che $E[Y] = E[\ln X]$. Utilizzando il procedimento prima utilizzato, calcolo la funzione di densità della variabile casuale Y , $f_y(Y)$. A questo punto posso calcolare $E[Y]$:

$$E[Y] = \int_{-\infty}^{+\infty} y \cdot f_y(y) \cdot dy$$

ma è più comodo sfruttare la funzione di densità della variabile X , che conosciamo già e calcolare la media di Y come:

$$E[Y] = \int_{-\infty}^{+\infty} \ln x \cdot f_x(x) \cdot dx$$

Infatti in generale vale che:

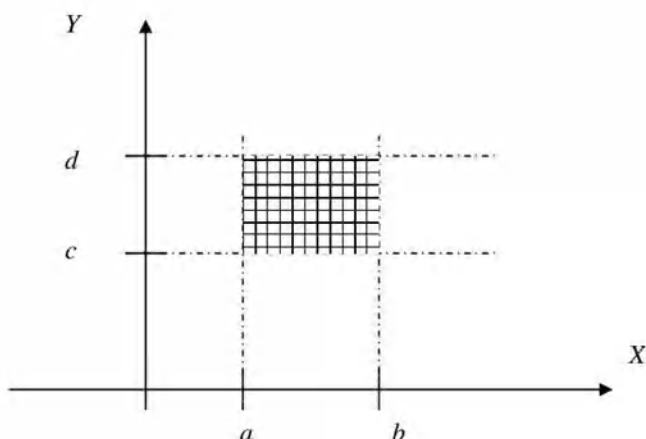
$$E[Y] = \int_{-\infty}^{+\infty} g(x) \cdot f_x(x) \cdot dx$$

4.4 – DISTRIBUZIONI CASUALI MULTIDIMENSIONALI

Date k variabili casuali $X_1, X_2, \dots, X_k$, vogliamo sapere qual è la relazione che le lega. In particolare ci soffermeremo sulle **distribuzioni casuali bidimensionali**, quindi troveremo le possibili relazioni tra due variabili casuali X e Y .

Fino ad ora, abbiamo considerato un'unica variabile X e l'abbiamo descritta tramite la funzione di ripartizione. Avevamo $F_x(X) = p(X \leq x)$, e abbiamo trovato probabilità del tipo $p(a \leq X \leq b)$.

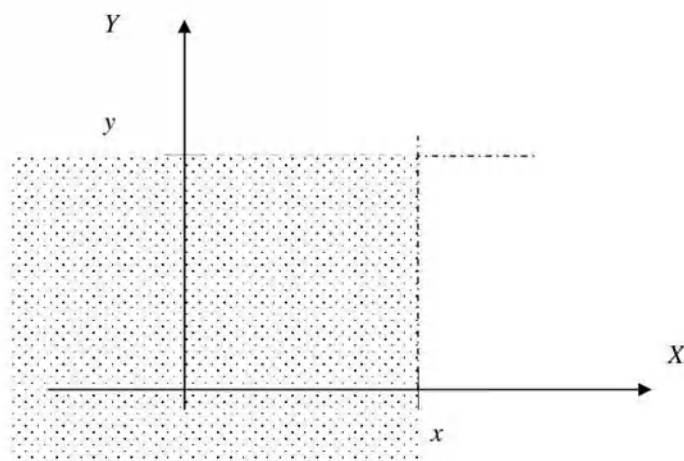
Adesso vogliamo calcolare congiuntamente $p(a \leq X \leq b)$ e $p(c \leq Y \leq d)$, cioè vogliamo calcolare $p(a \leq X \leq b, c \leq Y \leq d)$, che è rappresentata in figura (infatti sappiamo che la probabilità è l'area della figura).



La funzione di ripartizione di una distribuzione casuali dimensionale è così definita:

$$F_{x,y}(x, y) = p(X \leq x, Y \leq y)$$

tale che $R^2 \rightarrow [0,1]$ (infatti la probabilità è compresa tra 0 e 1).



Se conosco $F_{x,y}(x, y)$ posso ricavare ad esempio $p(a \leq X \leq b, c \leq Y \leq d)$.

4.5 – PROPRIETÀ DELLA DISTRIBUZIONI CASUALI BIDIMENSIONALI

La funzione di ripartizione $F_{x,y}(x, y)$ rispetta le seguenti proprietà:

$$\begin{aligned} \lim_{x \rightarrow -\infty} F_{x,y}(x, y) &= 0 \\ \lim_{y \rightarrow -\infty} F_{x,y}(x, y) &= 0 \\ \lim_{x \rightarrow +\infty} F_{x,y}(x, y) &= 1 \\ \lim_{y \rightarrow +\infty} F_{x,y}(x, y) &= 1 \end{aligned}$$

Se invece consideriamo $X \leq +\infty$ o $Y \leq +\infty$, (che sono sempre verificate) si ha un modo per passare dalla funzione di ripartizione congiunta a quella marginale:

$$\begin{aligned} p(X \leq +\infty, Y \leq y) &= p(Y \leq y) = F_y(Y) \\ p(X \leq x, Y \leq +\infty) &= p(X \leq x) = F_x(X) \end{aligned}$$

È quindi sempre possibile passare dalla funzione di ripartizione congiunta alle funzioni marginali:

$$F_{x,y}(x, y) \Rightarrow \begin{pmatrix} F_x(X) \\ F_y(y) \end{pmatrix} \quad \text{SI'}$$

mentre non è possibile il contrario:

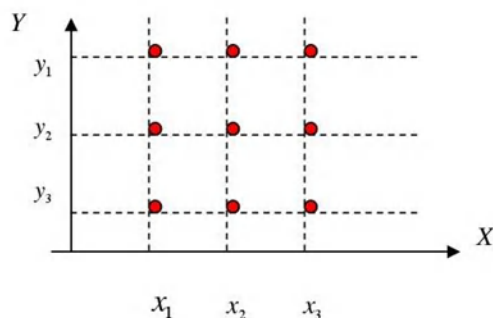
$$\begin{pmatrix} F_x(X) \\ F_y(y) \end{pmatrix} \Rightarrow F_{x,y}(x, y) \quad \text{NO}$$

4.6 – VARIABILI CASUALI DISCRETE BIDIMENSIONALI

Consideriamo due variabili casuali discrete X e Y . Dato che sono discrete, i loro possibili valori sono elencabili. Avremo quindi:

$$X = \{x_1, x_2, x_3, \dots, x_n\}$$

$$Y = \{y_1, y_2, y_3, \dots, y_n\}$$



È molto complicato calcolare la funzione di ripartizione congiunta $F_{x,y}(x, y)$, pertanto utilizzeremo direttamente la funzione di probabilità $p(X = x_i, Y = y_j) = p_{x,y}(x_i, y_j)$. Al variare di i e j , ho l'intera distribuzione della coppia (x, y) . La funzione di probabilità, deve rispettare le seguenti proprietà:

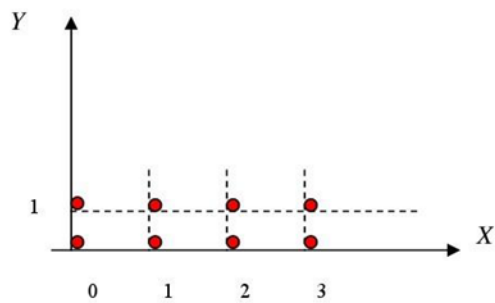
- 1) $p_{x,y}(x_i, y_j) \geq 0$
- 2) $\sum_i \sum_j p_{x,y}(x_i, y_j) = 1$

Esempio

Consideriamo 3 lanci di una moneta regolare e consideriamo la variabile casuale X , ovvero il numero totale di teste in 3 lanci, e la variabile casuale Y , cioè il numero di teste al primo lancio. Si avrà quindi che:

$$X = \{0, 1, 2, 3\}$$

$$Y = \{0, 1\}$$



Si ha che:

se $x = 0$, $y = 0$ si ha $1/8$

se $x = 1$, $y = 0$ si ha $2/8$

se $x = 2$, $y = 0$ si ha $1/8$

se $x = 3$, $y = 0$ si ha 0

se $x = 0$, $y = 1$ si ha 0

se $x = 1$, $y = 1$ si ha $1/8$

se $x = 2$, $y = 1$ si ha $2/8$

se $x = 3$, $y = 1$ si ha $1/8$

4.1 – ESERCIZIO SULLA DISTRIBUZIONE NORMALE DI GAUSS

È data la variabile casuale X , ovvero il tempo di vita di dispositivi elettronici distribuiti da un'azienda. Sappiamo che $X \sim N(\mu; \sigma^2)$, e che $\mu = 327, \sigma^2 = 231$. L'unità di misura è l'ora.

- 1) Calcolare la probabilità che un dispositivo abbia durata compresa tra 295 e 343 ore, cioè:
 $p(295 < X < 343)$ Riconduciamo la distribuzione normale a quella standard:

$$Z = \frac{X - 327}{\sqrt{231}} \sim N(0,1)$$

Quindi in questo caso abbiamo:

$$Z = \frac{X - 327}{\sqrt{231}}$$

Pertanto avremo:

$$\begin{aligned} p(295 \leq X \leq 343) &= p\left(\frac{295 - 327}{\sqrt{231}} \leq Z \leq \frac{343 - 327}{\sqrt{231}}\right) = p(-2,11 \leq Z \leq 1,05) = \Phi(2,11) - \Phi(-1,05) = \\ &= 0,8531 - 1 + 0,9826 = 0,8531 - 0,0174 = 0,8357 \rightarrow 83,57\% \end{aligned}$$

$$N.B. \Phi(-z) = 1 - \Phi(z)$$

- 2) Calcolare il tempo di vita superato dal 23% dei dispositivi.

$$0,23 = p(X > t) = p\left(Z > \frac{t - 327}{\sqrt{231}}\right) = 1 - p\left(Z \leq \frac{t - 327}{\sqrt{231}}\right) = 1 - \Phi\left(\frac{t - 327}{\sqrt{231}}\right)$$

$$0,77 = \Phi\left(\frac{t - 327}{\sqrt{231}}\right)$$

$$\Phi^{-1}(0,77) = \frac{t - 327}{\sqrt{231}}$$

$$0,74 = \frac{t - 327}{\sqrt{231}}$$

$$t = 0,74 \cdot \sqrt{231} + 327 = 338,25$$

- 3) Calcolare la probabilità che un dispositivo duri più di 350 ore dato che funziona da 300 ore.

$$p(X > 350 | X > 300) = \frac{p((X > 350) \cap (X > 300))}{p(X > 300)} = \frac{p(X > 350)}{p(X > 300)} = \frac{0,0655}{0,9625} = 0,0681$$

$$p(X > 350) = p\left(Z > \frac{350 - 327}{\sqrt{231}}\right) = 1 - \Phi\left(\frac{350 - 327}{\sqrt{231}}\right) = 1 - \Phi(1,51) = 1 - 0,9345 = 0,0655$$

$$p(X > 300) = p\left(Z > \frac{300 - 327}{\sqrt{231}}\right) = 1 - \Phi\left(\frac{300 - 327}{\sqrt{231}}\right) = 1 - \Phi(-1,78) = 1 - 1 + 0,9625 = 0,9625$$

N.B. La distribuzione normale non ha assenza di memoria

- 4) Calcolare $p(Z > -18,23)$

$$p(Z > -18,23) = 1 - p(Z \leq -18,23) = 1 - 1 + p(Z \leq 18,23) = 1 - 1 + (\approx 1) \approx 1$$

Questo valore non è possibile trovarlo sulle tavole, quindi sarà praticamente 1.

- 5) Supponiamo ora di non conoscere più i valori della media e della varianza, ma sappiamo che $X \sim N(\mu; \sigma^2)$. Sappiamo inoltre che $p(X < 200) = p(X > 300) = 0,025$. Dobbiamo calcolare la media e la varianza.

$$\begin{aligned} \begin{cases} 0,025 = p(X < 200) \\ 0,025 = p(X > 300) \end{cases} &= \begin{cases} 0,025 = p\left(Z < \frac{200 - \mu}{\sigma}\right) \\ 0,025 = 1 - p\left(Z < \frac{300 - \mu}{\sigma}\right) \end{cases} = \begin{cases} 0,025 = \Phi\left(\frac{200 - \mu}{\sigma}\right) \\ 0,025 = 1 - \Phi\left(\frac{300 - \mu}{\sigma}\right) \end{cases} = \\ &= \begin{cases} \Phi^{-1}(0,025) = \frac{200 - \mu}{\sigma} \\ \Phi^{-1}(0,975) = \frac{300 - \mu}{\sigma} \end{cases} = \begin{cases} -1,96 = \frac{200 - \mu}{\sigma} \\ 1,96 = \frac{300 - \mu}{\sigma} \end{cases} = \begin{cases} \mu = 1,96 \cdot \sigma + 200 \\ 1,96 \cdot \sigma = 300 - 1,96 \cdot \sigma - 200 \end{cases} = \\ &= \begin{cases} \mu = 1,96 \cdot \sigma + 200 \\ 2 \cdot 1,96 \cdot \sigma = 100 \end{cases} = \begin{cases} \mu = 1,96 \cdot \sigma + 200 \\ \sigma = \frac{100}{2 \cdot 1,96} = 25,51 \end{cases} = \begin{cases} \mu = 1,96 \cdot 25,51 + 200 = 250 \\ \sigma = 25,51 \end{cases} = \begin{cases} \mu = 250 \\ \sigma = 25,51 \end{cases} \end{aligned}$$

4.2 – ESERCIZIO SULLE TRASFORMAZIONI VARIABILI CASUALI

Sia X una variabile casuale uniforme e continua in $(0,1)$, cioè $X \sim U(0;1)$. La variabile casuale Y è una trasformazione della variabile X , tale che: $Y = g(x) = e^x$.

- 1) Si determinino la funzione di ripartizione $F_Y(y)$ e la fz. di densità $f_Y(y)$ della trasformata:

$$F_Y(y) = p(Y \leq y) = p(e^x \leq y) = p(X \leq \ln y) = F_X(\ln y) = \ln y$$

(abbiamo potuto effettuare la trasformazione dall'esponenziale al logaritmo senza cambiare di segno, in quanto il logaritmo è funzione monotona crescente)

Quindi $F_Y(y) = \ln y$ per $0 \leq \ln y \leq 1$, cioè $1 \leq y \leq e$.

Sappiamo inoltre che:

$$f_X(x) = 1 \text{ per } 0 \leq x \leq 1$$

$$F_X(x) = \int_0^x f_X(t) \cdot dt = \int_0^x 1 \cdot dt = [t]_0^x = x$$

$$\text{Calcoliamo ora } f_Y(y) = dF_Y(y) = F_Y'(y) = \frac{1}{y} \text{ per } 1 \leq y \leq e$$

oppure in maniera analoga:

$$f_Y(y) = F_X'(\ln y) \cdot \frac{1}{y} = f_X(\ln y) \cdot \frac{1}{y} = 1 \cdot \frac{1}{y} = \frac{1}{y}$$

cioè:

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{dg^{-1}(y)}{dy} \right|$$

- 2) Calcolare la mediana e la media di Y

$$F_Y(y) = 0,5$$

$$\ln y = 0,5$$

$$y = e^{0,5} = \sqrt{e}$$

$$E[Y] = \int_{-\infty}^{+\infty} y \cdot f_Y(y) \cdot dy = \int_1^e y \cdot \frac{1}{y} \cdot dy = \int_1^e 1 \cdot dy = [y]_1^e = e - 1$$

oppure:

$$E[Y] = E[e^x] = \int_{-\infty}^{+\infty} e^x \cdot f_X(x) \cdot dx = \int_0^1 e^x \cdot f_X(x) \cdot dx = [e^x]_0^1 = e - 1$$

3) Calcolare la varianza di Y

$$V[Y] = E[Y^2] - (E[Y])^2 = \left(\int_1^e y \cdot dy \right) - (e - 1)^2 = \left[\frac{y^2}{2} \right]_1^e - (e - 1)^2 = \frac{e^2}{2} - \frac{1}{2} - e^2 - 1 + 2e$$

oppure:

$$V[Y] = V[e^x] = E[e^{2x}] - (E[e^x])^2 = \left[\frac{e^{2x}}{2} \right]_0^1 - (e - 1)^2 = \frac{e^2}{2} - \frac{1}{2} - e^2 - 1 + 2e$$

5.1 – VARIABILI CASUALI DISCRETE BIDIMENSIONALI: RIEPILOGO

Consideriamo due variabili casuali discrete X e Y tali che:

$$X = \{x_1, x_2, x_3, \dots, x_n\}$$

$$Y = \{y_1, y_2, y_3, \dots, y_n\}$$

È molto complicato calcolare la funzione di ripartizione congiunta $F_{x,y}(x, y) = p(X \leq x, Y \leq y)$, pertanto utilizzeremo direttamente la funzione di probabilità $p(X = x_i, Y = y_j) = p_{x,y}(x_i, y_j)$.

5.2 – VARIABILI CASUALI CONTINUE BIDIMENSIONALI

Consideriamo due variabili casuali continue X e Y , che hanno, in quanto continue, i seguenti possibili valori:

$$X \in (a, b)$$

$$Y \in (c, d)$$

Vogliamo calcolare la funzione di ripartizione congiunta. Essa risulta pari all'integrale doppio della funzione di densità congiunta delle variabili X ed Y :

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

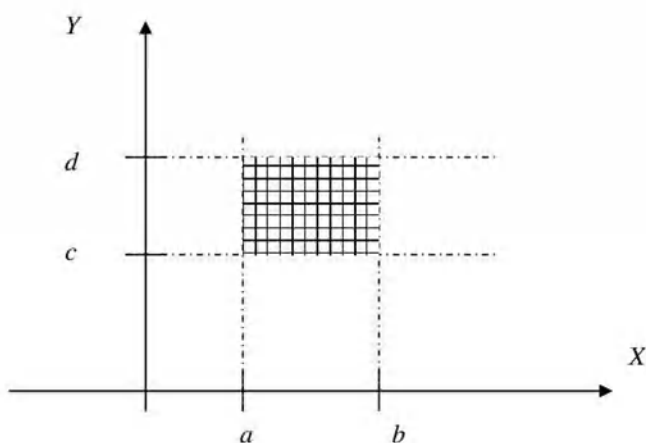
Se la funzione di ripartizione così definita esiste $\forall x, y$ allora dirò che $f_{x,y}$ è la funzione di densità della variabile casuale doppia continua (X, Y) .

La funzione di densità congiunta $f_{x,y}$, per essere tale, deve rispettare le seguenti proprietà:

$$1) f_{x,y}(x, y) \geq 0$$

$$2) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{x,y}(x, y) \cdot dx \cdot dy = 1$$

Vogliamo calcolare congiuntamente $p(a \leq X \leq b)$ e $p(c \leq Y \leq d)$, cioè vogliamo calcolare $p(a \leq X \leq b, c \leq Y \leq d)$, che è rappresentata in figura. (infatti sappiamo che la probabilità è il **volume** della figura nel caso di una variabile casuale continua), cioè:



che è pari a:

$$\int_a^b \int_c^d f_{x,y}(x, y) \cdot dx \cdot dy$$

Nel caso delle variabili casuali continue, inoltre, si ha che:

$$f_{x,y} = \frac{\partial^2 F_{x,y}(x, y)}{\partial x \cdot \partial y}$$

Esempio di variabili casuale continua bidimensionale

Consideriamo due variabili casuali continue X e Y , che hanno una funzione di ripartizione congiunta così definita:

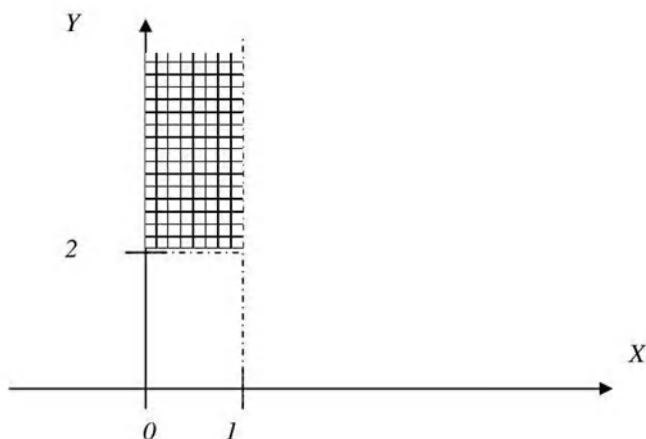
$$F_{x,y}(x, y) = \begin{cases} (1 - e^{-x}) \cdot (1 - e^{-y}) & \text{per } x > 0, y > 0 \\ 0 & \text{altrove} \end{cases}$$

Ne deduciamo quindi che i possibili valori delle due variabili casuali X e Y sono i seguenti:

$$X \in (0, +\infty)$$

$$Y \in (0, +\infty)$$

Vogliamo calcolare $p(0 \leq x \leq 1, y > 2)$, cioè:



$$\begin{aligned} p(0 \leq x \leq 1, y > 2) &= \int_0^1 \int_2^{+\infty} f_{x,y}(x, y) \cdot dx \cdot dy = \int_0^1 \int_2^{+\infty} e^{-x} \cdot e^{-y} \cdot dx \cdot dy = \int_0^1 e^{-x} \cdot \left[\int_2^{+\infty} e^{-y} \cdot dy \right] \cdot dx = \\ &= \int_0^1 e^{-x} \cdot \left[-e^{-y} \right]_2^{+\infty} \cdot dx = \int_0^1 e^{-x} \cdot e^{-2} \cdot dx = e^{-2} \cdot \int_0^1 e^{-x} \cdot dx = e^{-2} \cdot \left[-e^{-x} \right]_0^1 = e^{-2} \cdot [1 - e^{-1}] = e^{-2} - e^{-3} = 0,085 \end{aligned}$$

infatti si ha che:

$$f_{x,y} = \frac{\partial^2 F_{x,y}(x, y)}{\partial x \cdot \partial y} = \frac{\partial^2 (1 - e^{-x}) \cdot (1 - e^{-y})}{\partial x \cdot \partial y} = \frac{\partial (1 - e^{-x}) \cdot e^{-y}}{\partial y} = e^{-x} \cdot e^{-y} \quad \text{per } x, y > 0$$

5.3 – DALLA CONGIUNTA ALLE MARGINALI

Abbiamo visto precedentemente che è sempre possibile passare dalla funzione di ripartizione congiunta alle funzioni marginali:

$$F_{x,y}(x, y) \Rightarrow \begin{pmatrix} F_x(x) \\ F_y(y) \end{pmatrix} \quad \text{SI'}$$

mentre non è possibile il contrario.

Caso discreto

Nel caso discreto abbiamo visto che è complicato calcolare la funzione di ripartizione congiunta $F_{x,y}(x,y) = p(X \leq x, Y \leq y)$. Pertanto si utilizza la funzione di probabilità:

$$p(X = x_i, Y = y_j) = p_{x,y}(x_i, y_j)$$

Quindi voglio trovare, a partire da $p_{x,y}(x_i, y_j)$, la funzione di probabilità della variabile casuale X , ovvero $p_x(x_i)$. Essa si trova nel seguente modo:

$$p_x(x_i) = p(X = x_i) = \sum_j p_{x,y}(x_i, y_j)$$

Ad esempio, se considero l'esercizio prima considerato nel caso discreto, dove si ha che:

se $x = 0, y = 0$ si ha $1/8$

se $x = 1, y = 0$ si ha $2/8$

se $x = 2, y = 0$ si ha $1/8$

se $x = 3, y = 0$ si ha 0

se $x = 0, y = 1$ si ha 0

se $x = 1, y = 1$ si ha $1/8$

se $x = 2, y = 1$ si ha $2/8$

se $x = 3, y = 1$ si ha $1/8$

se voglio calcolare ad esempio $p_x(0) = p(X = 0)$, avrò che:

$$p_x(0) = p(X = 0) = \sum_j p_{x,y}(x_i, y_j) = p(Y = 0) + p(Y = 1) = \frac{1}{8} + 0 = \frac{1}{8}$$

Caso continuo

Nel caso continuo, la funzione di densità di X , risulta essere pari a:

$$f_x(x) = \int_{-\infty}^{+\infty} f_{x,y}(x,y) \cdot dy$$

Analogamente si avrà che la funzione di densità marginale di Y sarà pari a:

$$f_y(y) = \int_{-\infty}^{+\infty} f_{x,y}(x,y) \cdot dx$$

Ad esempio, se considero l'esercizio prima considerato nel caso continuo, dove si ha che:

$$f_{x,y}(x,y) = \begin{cases} e^{-x} \cdot e^{-y} & \text{per } x > 0, y > 0 \\ 0 & \text{altrove} \end{cases}$$

se voglio trovare la funzione di densità di X , avrò che:

$$f_x(x) = \int_{-\infty}^0 f_{x,y}(x,y) \cdot dy + \int_0^{+\infty} f_{x,y}(x,y) \cdot dy = 0 + \int_0^{+\infty} e^{-x-y} \cdot dy = 0 + e^{-x} \cdot \int_0^{+\infty} e^{-y} \cdot dy = 0 + e^{-x} \cdot [e^{-y}]_0^{+\infty} = e^{-x}$$

Quindi la funzione di densità di X , in questo caso, risulta pari alla funzione di densità di un'esponenziale di parametro 1 (cioè con $\lambda = 1$). Si ha quindi che: $X \sim \text{Exp}(1)$. Analogamente, se si effettuano i calcoli, si ha che la funzione di densità di Y , in questo caso, risulta pari alla funzione di densità di un'esponenziale di parametro 1. Si ha quindi che: $Y \sim \text{Exp}(1)$.

5.4 – MOMENTI DI UNA VARIABILI CASUALE BIDIMENSIONALE

Consideriamo due variabili casuali continue X e Y . Partendo dalla funzione di densità congiunta posso calcolare la funzione di densità marginale. Quindi, ad esempio, posso calcolare $f_x(x)$ ed utilizzarla per calcolare il valore atteso di X (e anche la varianza):

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f_x(x) \cdot dx$$

Voglio però calcolare il valore atteso della variabile casuale bidimensionale, cioè $E[XY]$. Esso risulta pari a:

$$E[X \cdot Y] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f_{x,y}(x, y) \cdot dx \cdot dy$$

Se voglio invece calcolare come variano congiuntamente le due variabili, calcolo la covarianza:

$$COV[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

se $COV[X, Y] = 0$, significa che X e Y sono incorrelate.

Ad esempio, se considero l'esercizio prima considerato nel caso continuo, si ha che:

$$\begin{aligned} E[X \cdot Y] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f_{x,y}(x, y) \cdot dx \cdot dy = \int_0^{+\infty} \int_0^{+\infty} x \cdot y \cdot e^{-x-y} \cdot dx \cdot dy = \int_0^{+\infty} x \cdot e^{-x} \cdot \left[\int_0^{+\infty} y \cdot e^{-y} \cdot dy \right] \cdot dx = \\ &= \int_0^{+\infty} x \cdot e^{-x} \cdot 1 \cdot dx = 1 \end{aligned}$$

infatti:

$$E[Y] = \int_0^{+\infty} y \cdot e^{-y} \cdot dy = 1$$

$$E[X] = \int_0^{+\infty} x \cdot e^{-x} \cdot dx = 1$$

Quindi si conclude dicendo che:

$$COV[X, Y] = 1 - 1 \cdot 1 = 0$$

cioè X e Y sono incorrelate.

5.5 – DISTRIBUZIONE CASUALE BIDIMENSIONALE NORMALE BIVARIATA

La distribuzione bidimensionale normale bivariata ha la seguente funzione di densità congiunta:

$$f_{x,y}(x, y) = \frac{1}{2\pi \cdot \delta_x \cdot \delta_y \cdot \sqrt{1-\rho^2}} \cdot \exp\left(-\frac{1}{2 \cdot (1-\rho^2)} \cdot \left(\frac{(x-\mu_x)^2}{\delta_x^2} + \frac{(y-\mu_y)^2}{\delta_y^2} - \frac{2\rho \cdot (x-\mu_x)(y-\mu_y)}{\delta_x \cdot \delta_y}\right)\right)$$

dove $-1 \leq \rho \leq 1$ ed è il coefficiente di correlazione lineare.

Le due variabili casuali X e Y dipendono quindi, nel caso della distribuzione normale bivariata, dai parametri μ_x (ovvero la media di X), μ_y (ovvero la media di Y), δ_x (ovvero lo scarto quadratico medio di X), δ_y (ovvero lo scarto quadratico medio di Y), ρ (ovvero il coefficiente di correlazione lineare).

Come sono le distribuzioni marginali di X e Y ?

Si ha che:

$$f_x(x) = \int_{-\infty}^{+\infty} f_{x,y}(x, y) \cdot dy = \frac{1}{\sqrt{2\pi} \cdot \delta_x} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(x-\mu_x)^2}{\delta_x^2}\right)$$

dove

$$\exp\left(-\frac{1}{2} \cdot \frac{(x-\mu_x)^2}{\delta_x^2}\right)$$

equivale alla densità di normale unidimensionale $X \sim N(\mu_x, \delta_x^2)$.

Analogamente si ricava che $Y \sim N(\mu_y, \delta_y^2)$.

Nel caso di una distribuzione normale bivariata, si ha che la covarianza è pari a:

$$COV[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y] = \left(\iint xy \cdot f_{x,y}(x, y) \cdot dx \cdot dy \right) - \mu_x \cdot \mu_y = \rho \cdot \delta_x \cdot \delta_y$$

Quindi si ricava che il coefficiente di correlazione lineare è pari a:

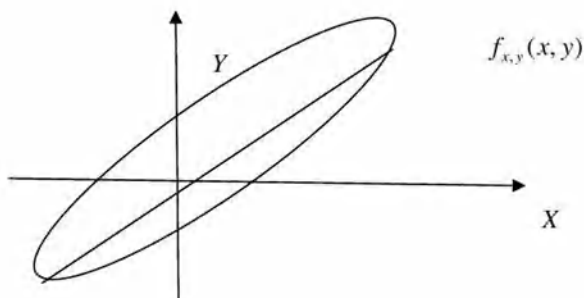
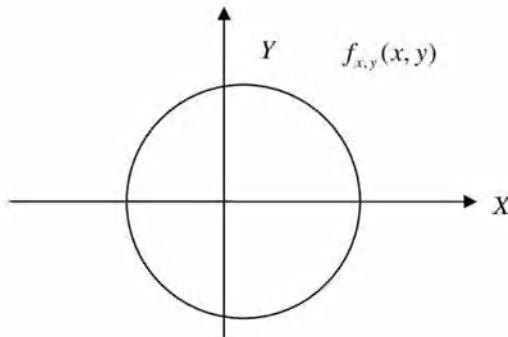
$$\rho = \frac{COV[X, Y]}{\delta_x \cdot \delta_y}$$

Se $\rho = 0$, le variabili sono incorrelate.

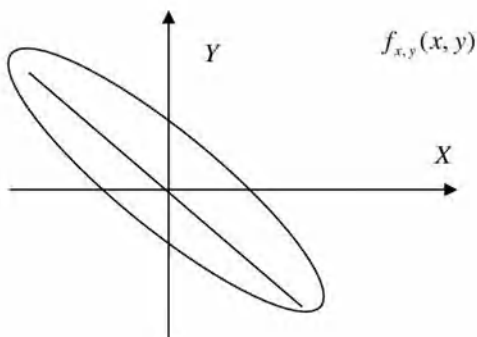
Ma come è fatta graficamente la funzione di densità della distribuzione normale bivariata?

Consideriamo il caso in cui $\mu_x = \mu_y = 0$ (se anziché 0 c'è un qualsiasi altro valore o i valori delle due medie sono diversi tra di loro, viene semplicemente traslato il grafico in un nuovo centro), $\delta_x^2 = \delta_y^2 = 1$ (se anziché 1 c'è un qualsiasi altro valore stirò o meno gli assi). Si ha che:

se $\rho = 0$, la funzione di densità è una circonferenza



$\rho > 0$, la funzione di densità è un'ellisse, con asse $y = x$ (la bisettrice). Tanto più la curva si avvicina a 1 (correlazione positiva perfetta), tanto più le curve si schiacciano rispetto alla bisettrice. Tanto più si avvicina a 0, tanto meno sono concentrate l'ellissi.



$\rho < 0$, la funzione di densità è un'ellisse, con asse $y = -x$ (la bisettrice). Tanto più la curva si avvicina a -1 (correlazione negativa perfetta), tanto più le curve si schiacciano rispetto alla bisettrice, tanto più tende a 0, tanto più tende a diventare un cerchio.

5.6 – VARIABILI CASUALI BIDIMENSIONALI: INDIPENDENZA

Sono date due variabili casuali X e Y . Quando X è indipendente da Y ? Dal punto di vista intuitivo, significa che se conosco X , questo non mi dà alcuna informazione su Y . Se considero due intervalli qualsiasi $A = (a, b)$, $C = (c, d)$, dire che X è indipendente da Y significa che:

$$p((X \in A) \cap (Y \in C)) = p(X \in A) \cdot p(Y \in C) \quad (1)$$

Se la relazione vale per qualunque intervallo A , C , gli eventi che dipendono da X sono indipendenti quindi dagli eventi che dipendono da Y .

In generale, dati n eventi $X_1, X_2, \dots, X_n$ e considerati n intervalli qualsiasi $A_1, A_2, \dots, A_n$, vale che:

$$p((X_1 \in A_1) \cap (X_2 \in A_2) \cap \dots \cap (X_n \in A_n)) = p(X_1 \in A_1) \cdot p(X_2 \in A_2) \cdot \dots \cdot p(X_n \in A_n)$$

se gli eventi sono indipendenti.

Come deve essere la funzione di ripartizione congiunta $F_{x,y}(x, y) = p(X \leq x, Y \leq y)$ perché X e Y siano indipendenti? Se considero due intervalli A e C tali che $A = (-\infty, x]$ e $C = (-\infty, y]$, dire che:

$$F_{x,y}(x, y) = p(X \leq x, Y \leq y)$$

equivale a dire:

$$F_{x,y}(x, y) = p(X \in A, Y \in C)$$

cioè equivale alla (1):

$$F_{x,y}(x, y) = p(X \in A, Y \in C) = p(X \in A) \cdot p(Y \in C)$$

Se X ed Y sono indipendenti si ha quindi:

$$F_{x,y}(x, y) = p(X \leq x, Y \leq y) = p(X \leq x) \cdot p(Y \leq y)$$

ma $p(X \leq x) = F_x(x)$ e $p(Y \leq y) = F_y(y)$ e quindi si conclude che:

$$F_{x,y}(x, y) = F_x(x) \cdot F_y(y)$$

ovvero *se due variabili casuali qualsiasi sono indipendenti la funzione di ripartizione congiunta è il prodotto delle funzioni di ripartizione marginali*. Analogamente vale il viceversa.

Ma normalmente, se ho una coppia di variabili casuali non utilizzo la funzione di ripartizione, ma la funzione di probabilità (se le variabili casuali sono discrete) o la funzione di densità (se le variabili casuali sono continue). Che relazione hanno tra di loro queste funzioni?

Caso discreto

Considero due variabili casuali discrete X ed Y . Anziché utilizzare la funzione di ripartizione congiunta (difficile da ottenere), utilizzo la funzione di probabilità congiunta:

$$p_{x,y}(x_i, y_j) = p(X = x_i, Y = y_j)$$

Se le due variabili sono indipendenti, secondo la definizione si avrà che:

$$p_{x,y}(x_i, y_j) = p(X = x_i) \cdot p(Y = y_j)$$

ma $p(X = x_i) = p_x(x)$ e $p(Y = y_j) = p_y(y)$. Quindi si conclude che:

$$p_{x,y}(x_i, y_j) = p_x(x) \cdot p_y(y) \quad (2)$$

ovvero *se due variabili casuali discrete sono indipendenti la funzione di probabilità congiunta è il prodotto delle funzioni di probabilità marginali*. Analogamente vale il viceversa.

Ad esempio, se considero l'esercizio prima considerato nel caso discreto, dove si ha che:

se $x = 0$, $y = 0$ si ha $1/8$

se $x = 1$, $y = 0$ si ha $2/8$

se $x = 2$, $y = 0$ si ha $1/8$

se $x = 3$, $y = 0$ si ha 0

se $x = 0$, $y = 1$ si ha 0

se $x = 1$, $y = 1$ si ha $1/8$

se $x = 2$, $y = 1$ si ha $2/8$

se $x = 3, y = 1$ si ha $1/8$

ovvero si hanno le funzioni di probabilità congiunte, abbiamo visto che è possibile calcolare le funzioni di probabilità marginali. Ad esempio:

$$p_x(0) = \sum_j p_{x,y}(x_i, y_j) = p(Y=0) + p(Y=1) = \frac{1}{8} + 0 = \frac{1}{8}$$

$$p_y(0) = \sum_i p_{x,y}(x_i, y_j) = p(X=0) + p(X=1) + p(X=2) + p(X=3) = \frac{1}{8} + \frac{2}{8} + \frac{1}{8} + 0 = \frac{1}{2}$$

In generale, per vedere se X ed Y sono indipendenti, dovrò verificare che $\forall x_i, y_j$ vale la (2).

Ad esempio, vediamo se funziona per $x = 0, y = 0$:

$$p_{x,y}(0,0) \stackrel{?}{=} p_x(0) \cdot p_y(0)$$

$$\frac{1}{8} \stackrel{?}{=} \frac{1}{8} \cdot \frac{1}{2}$$

$$\frac{1}{8} \neq \frac{1}{16}$$

Si conclude che X ed Y non sono indipendenti.

Caso continuo

Considero due variabili casuali continue X ed Y . Utilizzo la funzione di densità congiunta, che sappiamo essere legata alla funzione di ripartizione congiunta dalla seguente relazione:

$$f_{x,y}(x,y) = \frac{\partial^2 F_{x,y}(x,y)}{\partial x \cdot \partial y}$$

ma se le due variabili sono indipendenti si ha che vale sempre che:

$$F_{x,y}(x,y) = F_x(x) \cdot F_y(y)$$

Quindi si ha che:

$$f_{x,y}(x,y) = \frac{\partial^2 F_x(x) \cdot F_y(y)}{\partial x \cdot \partial y} = \frac{\partial F_x(x)}{\partial x} \cdot f_y(y) = f_x(x) \cdot f_y(y)$$

cioè:

$$f_{x,y}(x,y) = f_x(x) \cdot f_y(y) \quad (3)$$

ovvero *se due variabili casuali continue sono indipendenti la funzione di densità congiunta è il prodotto delle funzioni di densità marginali*. Analogamente vale il viceversa.

5.7 – PROPRIETÀ DELLE VARIABILI CASUALI BIDIMENSIONALI INDIPENDENTI

Se X ed Y sono due variabili casuali indipendenti, allora si ha che:

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

infatti si ha (considerando ad esempio il caso continuo) che:

$$\begin{aligned} E[X \cdot Y] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f_x(x) \cdot f_y(y) \cdot dx \cdot dy = \int_{-\infty}^{+\infty} x \cdot f_x(x) \cdot \left[\int_{-\infty}^{+\infty} y \cdot f_y(y) \cdot dy \right] \cdot dx = \int_{-\infty}^{+\infty} x \cdot f_x(x) \cdot E[Y] \cdot dx = \\ &= E[Y] \cdot \int_{-\infty}^{+\infty} x \cdot f_x(x) \cdot dx = E[Y] \cdot E[X] \end{aligned}$$

da cui si ricava che:

$$COV[X, Y] = 0$$

infatti si ha che $COV[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y] = E[X] \cdot E[Y] - E[X] \cdot E[Y] = 0$.

Pertanto se due variabili casuali sono indipendenti, allora esse sono anche incorrelate, ma non vale il viceversa (infatti se due variabili casuali sono incorrelate, non è detto che esse siano indipendenti).

5.8 – INDIPENDENZA DI DUE VARIABILI CASUALI NORMALI BIVARIATE

Considero due variabili casuali continue X e Y normali bivariate. Voglio trovare la condizione per cui esse sono indipendenti. Per quanto dimostrato precedentemente, deve valere la (3).

Quando abbiamo trattato la distribuzione normale bivariata abbiamo ricavato le funzioni di densità marginali a partire dalla congiunta ed abbiamo scoperto che $X \sim N(\mu_x, \sigma_x^2)$ e che $Y \sim N(\mu_y, \sigma_y^2)$.

Sappiamo inoltre che la funzione di densità marginale è data da:

$$f_{x,y}(x, y) = \frac{1}{2\pi \cdot \sigma_x \cdot \sigma_y \cdot \sqrt{1-\rho^2}} \cdot \exp\left(-\frac{1}{2 \cdot (1-\rho^2)} \cdot \left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho \cdot (x-\mu_x)(y-\mu_y)}{\sigma_x \cdot \sigma_y}\right)\right)$$

e che le funzioni di densità marginali sono:

$$f_x(x) = \frac{1}{\sqrt{2\pi \cdot \sigma_x^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(x-\mu_x)^2}{\sigma_x^2}\right)$$

$$f_y(y) = \frac{1}{\sqrt{2\pi \cdot \sigma_y^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(y-\mu_y)^2}{\sigma_y^2}\right)$$

Quando vale $f_{x,y}(x, y) = f_x(x) \cdot f_y(y)$?

La relazione vale nel caso in cui $\rho = 0$.

Nella distribuzione normale bivariata $\rho = 0$, ovvero il coefficiente di correlazione nullo, significa che c'è indipendenza e non solo incorrelazione.

Questa relazione particolare vale solo nel caso di una distribuzione normale bivariata e non sempre.

5.9 – VARIABILI CASUALI BIDIMENSIONALI: DIPENDENZA

Come misuro la dipendenza tra due variabili casuali X ed Y ? Dal punto di vista intuitivo, se io conosco una delle due variabili, ad esempio Y , mi chiedo qual è la nuova distribuzione di X conoscendo il valore di Y . Voglio cioè calcolare:

$$X | Y = y$$

Caso discreto

Considero la funzione di probabilità marginale di X , ovvero:

$$p_x(x_i) = p(X = x_i)$$

voglio calcolare la dipendenza tra X e Y , quindi utilizzerò la funzione di probabilità condizionata:

$$p(X = x_i | Y = y_j) = \frac{p((X = x_i) \cap (Y = y_j))}{p(Y = y_j)}$$

ma $p((X = x_i) \cap (Y = y_j)) = p_{x,y}(x_i, y_j)$ (cioè la fz. di probabilità congiunta) e $p(Y = y_j) = p_y(y_j)$ (cioè la fz. di probabilità marginale di y). Si ha quindi che:

$$p(X = x_i | Y = y_j) = \frac{p((X = x_i) \cap (Y = y_j))}{p(Y = y_j)} = \frac{p_{x,y}(x_i, y_j)}{p_y(y_j)} \quad \forall x_i$$

che è una funzione di probabilità a tutti gli effetti (e quindi rispetta le due proprietà di una funzione di probabilità) che esprime la dipendenza tra X e Y .

Esempio

Considero 3 lanci di una moneta. La variabile casuale discreta X è il numero di teste (T) nei primi due lanci ed Y è il numero di teste negli ultimi due lanci. Pertanto i possibili valori delle variabili sono $X \in \{0,1,2\}$ e $Y \in \{0,1,2\}$. Voglio calcolare la nuova distribuzione di Y sapendo che $X = 2$, cioè:

$$p(Y = y_j | X = 2) = \frac{p_{y,x}(y_j, 2)}{p_x(2)} = \frac{p(Y = y_j, X = 2)}{p(X = 2)} \quad \text{con } j = 0, 1, 2$$

	$y = 0$	$y = 1$	$y = 2$	$p_x(x_i)$
$x = 0$	1/8	1/8	0	1/4
$x = 1$	1/8	1/4	1/8	1/2
$x = 2$	0	1/8	1/8	1/4
$p_y(y_j)$	1/4	1/2	1/4	

Si ha che :

$$p_x(0) = p(X = 2) = 0 + \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

Quindi:

	$p(Y = y_j, X = 2)$	$p(Y = y_j X = 2)$
$y = 0$	0	0
$y = 1$	1/8	1/2
$y = 2$	1/8	1/2

Caso continuo

Con ragionamento analogo a quello visto nel caso discreto, si ottiene che:

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

$\forall x$

5.1 – ESERCIZIO SULLE VARIABILI CASUALI BIDIMENSIONALI DISCRETE

Si consideri la funzione di probabilità di una variabile casuale bidimensionale discreta (X, Y) , cioè

$p_{x,y}(x, y) = p(X = x, Y = y)$, definita dalla tabella seguente:

	$y = 0$	$y = 1$	$y = 2$
$x = 0$	$\frac{1-\theta}{2}$	$\frac{1-\theta}{3}$	$\frac{1-\theta}{6}$
$x = 1$	$\frac{\theta}{2}$	$\frac{\theta}{3}$	$\frac{\theta}{6}$

1. Per quali valori di θ questa tabella fornisce una funzione che sia una funzione di probabilità?

$$\begin{aligned} \begin{cases} p_{x,y}(x, y) \geq 0 \\ \sum_i \sum_j p_{x,y}(x, y) = 1 \end{cases} &= \begin{cases} 1-\theta \geq 0 \wedge \theta \geq 0 \\ \frac{1-\theta}{2} + \frac{1-\theta}{3} + \frac{1-\theta}{6} + \frac{\theta}{2} + \frac{\theta}{3} + \frac{\theta}{6} = 1 \end{cases} = \begin{cases} \theta \leq 1 \wedge \theta \geq 0 \\ \frac{1-\theta + \theta + 2 - 2\theta + 2\theta + 3 - 3\theta + 3\theta}{6} = 1 \end{cases} \\ &= \begin{cases} 0 \leq \theta \leq 1 \\ \frac{1+3+2}{6} = 1 \end{cases} = \begin{cases} 0 \leq \theta \leq 1 \\ 1 = 1 \end{cases} = \begin{cases} 0 \leq \theta \leq 1 \\ \forall \theta \in R \end{cases} = 0 \leq \theta \leq 1 \end{aligned}$$

2. Si determinino le funzioni di probabilità delle variabili casuali marginali

$$p_x(x) = \sum_y p_{x,y}(x, y) = \sum_{y=0}^2 p_{x,y}(x, y) = \begin{cases} \frac{1-\theta}{2} + \frac{1-\theta}{3} + \frac{1-\theta}{6} & \text{se } x = 0 \\ \frac{\theta}{2} + \frac{\theta}{3} + \frac{\theta}{6} & \text{se } x = 1 \end{cases} = \begin{cases} 1-\theta & \text{se } x = 0 \\ \theta & \text{se } x = 1 \end{cases}$$

Si conclude che $X \sim Be(\theta)$

$$p_y(y) = \sum_x p_{x,y}(x, y) = \sum_{x=0}^1 p_{x,y}(x, y) = \begin{cases} \frac{1-\theta}{2} + \frac{\theta}{2} & \text{se } y = 0 \\ \frac{1-\theta}{3} + \frac{\theta}{3} & \text{se } y = 1 \\ \frac{1-\theta}{6} + \frac{\theta}{6} & \text{se } y = 2 \end{cases} = \begin{cases} \frac{1}{2} & \text{se } y = 0 \\ \frac{1}{3} & \text{se } y = 1 \\ \frac{1}{6} & \text{se } y = 2 \end{cases}$$

3. Si determini il momento misto di ordine due $E[XY]$, le due medie marginali $E[X]$ e $E[Y]$ e la covarianza $COV[X, Y]$.

$$E[X] = \sum_x x \cdot p_x(x) = 0 \cdot (1-\theta) + 1 \cdot \theta = \theta$$

$$E[Y] = \sum_y y \cdot p_y(y) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = \frac{2}{3}$$

$$E[XY] = \sum_x \sum_y xy \cdot p_{x,y}(x, y) = 0 \cdot 0 \cdot \frac{1-\theta}{2} + 0 \cdot 1 \cdot \frac{1-\theta}{3} + 0 \cdot 2 \cdot \frac{1-\theta}{6} + 1 \cdot 0 \cdot \frac{\theta}{2} + 1 \cdot 1 \cdot \frac{\theta}{3} + 1 \cdot 2 \cdot \frac{\theta}{6} = \frac{2}{3}\theta$$

$$COV[X, Y] = E[XY] - E[X] \cdot E[Y] = \frac{2}{3}\theta - \theta \cdot \frac{2}{3} = 0$$

Dato che la covarianza è nulla, implica che X ed Y sono incorrelate.

4. Calcolare $p(X=1|Y=1)$:

$$p(X=1|Y=1) = \frac{p((X=1) \cap (Y=1))}{p(Y=1)} = \frac{p_{x,y}(1,1)}{p_y(1)} = \frac{9}{3} \cdot 3 = 9$$

Dato che abbiamo scoperto che $X \sim Be(9)$, possiamo notare che $p(X=1|Y=1) = p(X=1) = 9$, il che significa che X e Y sono indipendenti per definizione.

5. Si determinino le funzioni di probabilità condizionate.

$$p_{X|Y=y}(x) = p(X=x|Y=y) = \frac{p((X=x) \cap (Y=y))}{p(Y=y)} = \frac{p_{x,y}(x,y)}{p_y(y)}$$

Si ha quindi:

$$p_{X|Y=0}(x) = p(X=x|Y=0) = \frac{p_{x,y}(x,0)}{p_y(0)} = \begin{cases} 1-9 & \text{se } x=0 \\ 9 & \text{se } x=1 \end{cases}$$

$$p_{X|Y=1}(x) = p(X=x|Y=1) = \frac{p_{x,y}(x,1)}{p_y(1)} = \begin{cases} 1-9 & \text{se } x=0 \\ 9 & \text{se } x=1 \end{cases}$$

$$p_{X|Y=2}(x) = p(X=x|Y=2) = \frac{p_{x,y}(x,2)}{p_y(2)} = \begin{cases} 1-9 & \text{se } x=0 \\ 9 & \text{se } x=1 \end{cases}$$

Dato che le distribuzioni condizionate sono uguali alla distribuzione marginale, se ne deduce che c'è indipendenza. Dato che c'è indipendenza c'è per forza incorrelazione (infatti la covarianza è nulla), ma non è detto che l'incorrelazione implichi l'indipendenza.

Analogamente si può calcolare:

$$p_{Y|X=x}(y) = p(Y=y|X=x) = \frac{p((Y=y) \cap (X=x))}{p(X=x)} = \frac{p_{x,y}(x,y)}{p_x(x)}$$

5.2 – ESERCIZIO SULLE VARIABILI CASUALI BIDIMENSIONALI CONTINUE

Si consideri la funzione di densità di una variabile casuale bidimensionale continua (X,Y) , cioè:

$$f_{x,y}(x,y) = k(x+y) \quad \text{con } 0 < x < 1, 0 < y < 2$$

1. Trovare il valore di k per cui la funzione sopra definita è la funzione di densità di una variabile casuale bidimensionale continua

$$\begin{aligned} \begin{cases} f_{x,y}(x,y) \geq 0 \\ 1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{x,y}(x,y) \cdot dx \cdot dy \end{cases} &= \begin{cases} k \geq 0 \\ 1 = \int_0^1 \int_0^2 k \cdot (x+y) \cdot dx \cdot dy \end{cases} = \begin{cases} k \geq 0 \\ 1 = k \cdot \int_0^1 \int_0^2 (x+y) \cdot dx \cdot dy \end{cases} = \\ &= \begin{cases} k \geq 0 \\ 1 = k \cdot \int_0^2 \left[\frac{x^2}{2} + xy \right]_0^1 \cdot dy \end{cases} = \begin{cases} k \geq 0 \\ 1 = k \cdot \int_0^2 \left[\frac{1}{2} + y \right] \cdot dy \end{cases} = \begin{cases} k \geq 0 \\ 1 = k \cdot \left[\frac{1}{2}y + \frac{y^2}{2} \right]_0^2 \end{cases} = \begin{cases} k \geq 0 \\ 1 = 3k \end{cases} = \begin{cases} k \geq 0 \\ k = \frac{1}{3} \end{cases} \Rightarrow k = \frac{1}{3} \end{aligned}$$

Si ha quindi che:

$$f_{x,y}(x,y) = \frac{x+y}{3} \quad \text{con } 0 < x < 1, 0 < y < 2$$

2. Si determinino le funzioni di densità delle variabili casuali marginali

$$f_x(x) = \int_{-\infty}^{+\infty} f_{x,y}(x,y) \cdot dy = \int_0^2 \left(\frac{x+y}{3} \right) \cdot dy = \frac{1}{3} \cdot \int_0^2 (x+y) \cdot dy = \frac{1}{3} \cdot \left[xy + \frac{y^2}{2} \right]_0^2 = \frac{1}{3} \cdot 2x + \frac{1}{3} \cdot 2 = \frac{2}{3} \cdot (x+1)$$

con $0 < x < 1$.

6.1 – RIEPILOGO: INDIPENDENZA E PROBABILITÀ CONDIZIONATE

Abbiamo precedentemente definito la probabilità condizionata come:

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x, y)}{p_Y(y)} \quad (\text{caso discreto})$$

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad (\text{caso continuo})$$

e l'indipendenza come:

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y) \quad (\text{caso discreto})$$

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) \quad (\text{caso continuo})$$

Ne risulta che, se vale:

$$p_{X|Y=y}(x) = p_X(x) \quad (\text{caso discreto})$$

$$f_{X|Y=y}(x) = f_X(x) \quad (\text{caso continuo})$$

allora c'è indipendenza.

Vogliamo ora calcolare la media o valore atteso di una variabile casuale condizionata:

$$E[X | Y] = \sum_{x_i} x_i \cdot p_{X|Y=y}(x) \quad (\text{caso discreto})$$

$$E[X | Y] = \int_{-\infty}^{+\infty} x \cdot f_{X|Y=y}(x) \cdot dx \quad (\text{caso continuo})$$

Ad esempio, se X ed Y sono due variabili casuali **normali bivariate**, sappiamo che $X | Y$ ha una forma del tipo:

$$X | Y \sim N\left(\mu_x + \frac{\delta_x}{\delta_y} \cdot \rho(y - \mu_y), \delta_x^2 \cdot (1 - \rho^2)\right)$$

mentre abbiamo studiato che X è così definito:

$$X \sim N(\mu_x, \delta_x^2)$$

Perché ci sia indipendenza è necessario che $\rho = 0$: infatti in questo modo si ha che $X | Y \sim X$.

I valori attesi di X ed $X | Y$ sono così definiti:

$$E[X] = \mu_x$$

$$E[X | Y] = \mu_x + \frac{\delta_x}{\delta_y} \cdot \rho(y - \mu_y)$$

6.2 – PROPRIETÀ DEL VALORE ATTESO ITERATO

Se voglio calcolare $E[X]$:

- Se è continuo utilizzo la definizione $\int_{-\infty}^{+\infty} x \cdot f_X(x) \cdot dx$, se è discreto utilizzo $\sum_i x_i \cdot p_X(x_i)$
- Se non conosco $f_X(x)$ o $p_X(x)$ e conosco $X | Y = y$ e Y , allora vale che

$$E[X | Y] = g(Y) = \int_{-\infty}^{+\infty} x \cdot f_{X|Y=y}(x) \cdot dx \quad \text{oppure} \quad E[X | Y] = \sum_{x_i} x_i \cdot p_{X|Y=y}(x_i)$$

Ne segue che:

$$E[X] = E_Y[E[X | Y]]$$

Ad esempio, nel caso della normale, se non sapessi quanto vale $E[X]$, passerei da $E[X | Y]$:

$$E[X | Y = y] = \mu_x + \frac{\delta_x}{\delta_y} \cdot \rho(y - \mu_y)$$

$$E[X] = E_Y \left[\mu_x + \frac{\delta_x}{\delta_y} \cdot \rho(y - \mu_y) \right] = \mu_x + \frac{\delta_x}{\delta_y} \cdot E_Y [\rho(y - \mu_y)] = \mu_x$$

infatti il valore atteso di y è μ_y e quindi il termine $E_Y [\rho(y - \mu_y)]$ si annulla.

6.3 – DISTRIBUZIONE DELLA SOMMA DI VARIABILI

Consideriamo $X_1, X_2, \dots, X_n$ variabili **indipendenti identicamente distribuite** (i.i.d.), tali cioè che $X_i = 1$ se è un successo, $X_i = 0$ se non è un successo (ognuna è una distribuzione di Bernoulli) e che $p(X_i = 1) = p$. Voglio valutare come è fatta la somma $X_1 + X_2 + X_3 + \dots + X_n$.

Si dice che $X_1 + X_2 + X_3 + \dots + X_n$ è la **somma** e $\frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$ **somma campionaria**. È importante calcolare la distribuzione della somma (automaticamente da essa posso poi ricavare la distribuzione della media) S_n e vogliamo calcolare una probabilità, ad esempio $p(a < S_n < b)$.

Distribuzione binomiale

Supponiamo di avere due variabili binomiali $X_1 \sim Bi(n_1, \theta)$ e $X_2 \sim Bi(n_2, \theta)$ indipendenti tra di loro. Da notare che $n_1 \neq n_2$ ma θ è lo stesso necessariamente. Vogliamo calcolare la distribuzione della somma, cioè qual è la distribuzione di $X_1 + X_2$. Ne deriva che:

$$X_1 + X_2 \sim Bi(n_1 + n_2, \theta)$$

Si dice che la distribuzione binomiale gode della **proprietà riproduttiva**. Questa proprietà vale analogamente per n variabili binomiali.

Distribuzione di Poisson

Supponiamo di avere due variabili di Poisson $X_1 \sim Po(\lambda_1)$ e $X_2 \sim Po(\lambda_2)$ indipendenti tra di loro. Vogliamo calcolare la distribuzione della somma, cioè qual è la distribuzione di $X_1 + X_2$.

Ne deriva che:

$$X_1 + X_2 \sim Po(\lambda_1 + \lambda_2)$$

Si dice che la distribuzione di Poisson gode della **proprietà riproduttiva**. Questa proprietà vale analogamente per n variabili di Poisson.

Distribuzione normale

Supponiamo di avere due variabili normali $X_1 \sim N(\mu_1, \delta_1^2)$ e $X_2 \sim N(\mu_2, \delta_2^2)$ indipendenti tra di loro. Vogliamo calcolare la distribuzione della somma, cioè qual è la distribuzione di $X_1 + X_2$.

Ne deriva che:

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \delta_1^2 + \delta_2^2)$$

Si dice che la distribuzione normale gode della **proprietà riproduttiva**. Questa proprietà vale analogamente per n variabili normali.

Distribuzione gamma

Supponiamo di avere due variabili gamma $X_1 \sim Ga(\alpha_1, \lambda)$ e $X_2 \sim Ga(\alpha_2, \lambda)$ indipendenti tra di loro. Da notare che $\alpha_1 \neq \alpha_2$ ma λ è lo stesso necessariamente. Vogliamo calcolare la distribuzione della somma, cioè qual è la distribuzione di $X_1 + X_2$. Ne deriva che:

$$X_1 + X_2 \sim Ga(\alpha_1 + \alpha_2, \lambda)$$

Si dice che la distribuzione gamma gode della **proprietà riproduttiva** rispetto ad α , mentre rimane fisso λ . Questa proprietà vale analogamente per n variabili gamma.

Distribuzione esponenziale

Abbiamo visto che un caso particolare della distribuzione gamma è la distribuzione esponenziale. Supponiamo di avere due variabili esponenziali $X_1 \sim \text{Exp}(\lambda) \sim \text{Ga}(1, \lambda)$ e $X_2 \sim \text{Exp}(\lambda) \sim \text{Ga}(1, \lambda)$ indipendenti tra di loro. Da notare che $\alpha_1 = \alpha_2 = 1$, λ è lo stesso necessariamente. Vogliamo calcolare la distribuzione della somma, cioè qual è la distribuzione di $X_1 + X_2$. Ne deriva che:

$$X_1 + X_2 \sim \text{Ga}(2, \lambda)$$

In generale vale che:

$$X_1 + X_2 + X_3 + \dots + X_n \sim \text{Ga}(n, \lambda)$$

che è il tempo che devo aspettare affinché ci sia l' n -esimo arrivo.

Si dice che la distribuzione esponenziale **non** gode della **proprietà riproduttiva**. Questa proprietà vale analogamente per n variabili esponenziali.

Altre distribuzioni

Se considero altre distribuzioni cosa succede? È molto difficile ricavare la distribuzione esatta, possiamo più facilmente calcolare il valore atteso e la varianza della somma. Dato che le variabili utilizzate per ipotesi sono indipendenti identicamente distribuite si avrà che:

$$E[X_i] = \mu$$

$$V[X_i] = \sigma^2$$

Allora si avrà:

$$E[S_n] = n\mu$$

infatti vale sempre che $E[X_1 + X_2 + \dots + X_n] = E[X_1] + E[X_2] + \dots + E[X_n]$

$$V[S_n] = n\sigma^2$$

Infatti, dato che sono indipendenti per ipotesi, significa che c'è in correlazione e quindi la covarianza è pari a 0. Infatti in generale non vale sempre che $V[X_1 + X_2] = V[X_1] + V[X_2]$.

Facilmente si può calcolare il valore atteso e la varianza della somma campionaria:

$$E\left[\frac{S_n}{n}\right] = \frac{E[S_n]}{n} = n\mu \cdot \frac{1}{n} = \mu$$

$$V\left[\frac{S_n}{n}\right] = \frac{V[S_n]}{n^2} = n\sigma^2 \cdot \frac{1}{n^2} = \frac{\sigma^2}{n}$$

Abbiamo detto che è molto difficile calcolare precisamente la distribuzione della somma S_n . Cerchiamo allora un'approssimazione di $p(a < S_n < b)$. Dobbiamo ricorrere al concetto di **convergenza di variabili casuali**. Intuitivamente per $n \rightarrow \infty$, si considererà che $S_n \rightarrow S$.

Sappiamo che se una successione numerica a_n converge significa che $\lim_{n \rightarrow \infty} a_n = a$, cioè significa che i valori a_n per $n \rightarrow \infty$ si trovano sempre più vicini ad a , in un intorno di a .

Stiamo però ora considerando una successione di variabili casuali, quindi non numerica. Avremo che $X_n \rightarrow X$. Che cosa significa?

Convergenza in distribuzione

Se consideriamo la convergenza in distribuzione, essa intuitivamente ci dice non tanto che $X_n \rightarrow X$ ma che la distribuzione della variabile X_n si avvicina sempre più (tende) alla distribuzione di X .

La distribuzione di X_n , sia nel caso di variabile casuale discreta che continua, è così definita:

$$p(X_n \leq x) = F_{X_n}(x)$$

La distribuzione di X , sia nel caso di variabile casuale discreta che continua, è così definita:

$$p(X \leq x) = F_X(x)$$

Questo significa che X_n converge a X in distribuzione se la funzione di ripartizione di X_n converge alla funzione di ripartizione di X , cioè:

$$X_n \xrightarrow{d} X \quad \text{se} \quad F_{X_n}(x) \rightarrow F_X(x)$$

Ma per quali valori x deve verificarsi questa convergenza?

$$\forall x \in R \text{ dove } F(x) \text{ è continua}$$

cioè per qualunque x tranne i punti dove la funzione ha i salti, è discontinua.

Esempio

Consideriamo ad esempio una successione delle variabili casuali $Y_1, Y_2, \dots, Y_n$ indipendenti identicamente distribuite e uniforme sull'intervallo $[0,1]$. Poiché le Y sono uniformi sull'intervallo $[0,1]$, si ha che:

$$f_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{altrove} \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ y & 0 < y < 1 \\ 1 & y \geq 1 \end{cases}$$

Definiamo $X_n = \max\{Y_1, Y_2, Y_3, \dots, Y_n\}$. Vogliamo calcolare a quale X converge X_n . Abbiamo visto che dire che $X_n \xrightarrow{d} X$ equivale a $F_{X_n}(x) \rightarrow F_X(x)$.

$$F_{X_n}(x) = p(X_n \leq x) = p(Y_1 \leq x, \dots, Y_n \leq x) = p(Y_1 \leq x) \cdot \dots \cdot p(Y_n \leq x) = F_{Y_1}(x) \cdot \dots \cdot F_{Y_n}(x) = [F_{Y_1}(x)]^n$$

infatti, poiché sono tutte uniformi hanno la stessa distribuzione e quindi $F_{Y_1}(x) \cdot \dots \cdot F_{Y_n}(x) = [F_{Y_1}(x)]^n$.

Avremo quindi che $F_{X_n}(x)$ è così definita:

$$F_{X_n}(x) = \begin{cases} 0 & x \leq 0 \\ x^n & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

A quale $F_X(x)$ tende $F_{X_n}(x)$ quando $n \rightarrow +\infty$?

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 0 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases} = \begin{cases} 0 & x < 1 \\ 1 & x \geq 1 \end{cases}$$

$F_X(x)$ è cioè una variabile casuale discreta degenera in 1, cioè $F_X(x) \rightarrow 1$.

Avremo quindi in definitiva che:

$$F_{X_n}(x) \rightarrow 1$$

Convergenza in probabilità

Se consideriamo la convergenza in probabilità, essa intuitivamente va a vedere direttamente quando la variabile X_n si trova vicino ad X , cioè quando $X_n \rightarrow X$. A differenza della convergenza in

distribuzione che si occupa appunto delle distribuzioni, questa convergenza si occupa direttamente di variabili. In particolare, voglio trovare $p(|X_n - X| > \varepsilon)$, cioè la probabilità che X_n sia lontano da X per più di una quantità ε fissata piccola a piacere. Per $n \rightarrow +\infty$ si ha che la probabilità tende a:

$$p(|X_n - X| > \varepsilon) \rightarrow 0 \quad \forall \varepsilon > 0$$

che equivale a dire che:

$$p(|X_n - X| \leq \varepsilon) \rightarrow 1 \quad \forall \varepsilon > 0$$

Se una successione X_n converge in probabilità a X scriverò:

$$X_n \xrightarrow{p} X$$

Proprietà e relazioni tra le convergenze

- 1) se $X_n \xrightarrow{p} X$ allora $X_n \xrightarrow{d} X$
ma *non viceversa*, cioè la convergenza in probabilità è più forte della convergenza in distribuzione.
- 2) se $X_n \xrightarrow{d} c$ allora $X_n \xrightarrow{p} c$
dove c è una variabile casuale costante, cioè degenerare. Quindi quando parlo di convergenza di una successione ad un numero è analogo utilizzare la convergenza in probabilità o in distribuzione.
- 3) se vale che $E[X_n] \rightarrow c$, cioè la successione delle medie delle variabili casuali X_i converge ad una costante c quando $n \rightarrow +\infty$, e $V[X_n] \rightarrow 0$ per $n \rightarrow +\infty$, allora $X_n \xrightarrow{p} c$

Legge dei grandi numeri

Data una successione di variabili casuali $X_1, X_2, \dots, X_n$ indipendenti e identicamente distribuite (i.i.d.), calcolo la media campionaria:

$$\frac{S_n}{n} = \frac{\sum_{i=1}^n X_i}{n} = \overline{X_n}$$

Questa successione a che cosa converge per $n \rightarrow +\infty$? Per rispondere a questo quesito dobbiamo aggiungere l'ipotesi che esista la media $E[X_i] = \mu$. Si avrà allora che:

$$X_n \xrightarrow{p} \mu$$

Esempio

È data una successione di variabili casuali $X_1, X_2, \dots, X_n$ i.i.d., dove ciascuna variabile $X_i \sim Bi(1, \vartheta) \sim Be(\vartheta)$, con $\vartheta = 1$ se si ha un successo, $\vartheta = 0$ in caso contrario, cioè $p(X_i = 1) = E[X_i] = \vartheta$. Vogliamo ad esempio calcolare a che cosa tende la percentuale di volte che è uscito 56 sulla ruota di Venezia. La percentuale di volte che è uscito 56 sulla ruota di Venezia non è altro che la media campionaria sopra definita. Per la legge dei grandi numeri sopra definita si ha:

$$\frac{\sum_{i=1}^n X_i}{n} = \overline{X_n} \xrightarrow{p} \vartheta = \frac{1}{90}$$

Dimostrazione

Vogliamo dimostrare che $X_n \xrightarrow{p} \mu$, cioè la legge dei grandi numeri. Sappiamo però dalla proprietà

(3) sopra definita che per affermare che $X_n \xrightarrow{p} c$ basta dimostrare che $E[X_n] \rightarrow c$ e $V[X_n] \rightarrow 0$ per $n \rightarrow +\infty$. Ipotizziamo nella dimostrazione che $E[X_i] = \mu$, $V[X_i] = \delta^2$. Calcoliamo quindi:

$$E[X_n] = E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n} \cdot n\mu = \mu$$

$$V[X_n] = \frac{1}{n^2} \cdot n\delta^2 = \frac{\delta^2}{n}$$

quando $n \rightarrow +\infty$ si ha che:

$$E[X_n] \rightarrow \mu$$

$$V[X_n] \rightarrow 0$$

Dato che i due prerequisiti sono dimostrati, allora abbiamo che $X_n \xrightarrow{p} \mu$.

Esempio

Supponiamo che le nostre osservazioni siano i.i.d. con $X_i \sim \text{Exp}(\theta)$. Sappiamo che in una distribuzione esponenziale:

$$E[X_i] = \frac{1}{\theta}$$

Calcolo la media campionaria, cioè faccio la media degli n tempi di attesa che ho osservato:

$$\overline{X}_n = \frac{\sum_{i=1}^n x_i}{n} \rightarrow \frac{1}{\theta}$$

per la legge dei grandi numeri.

Teorema centrale del limite

Abbiamo dimostrato che $\overline{X}_n \xrightarrow{p} \mu$, con $n \rightarrow +\infty$ o in altre parole che $\overline{X}_n - \mu \xrightarrow{p} 0$. Ma questo risultato non mi dice niente sulla distribuzione della media campionaria. Fino ad ora non so calcolare probabilità come $p(\overline{X}_n < 3)$. Per rispondere a questa esigenza utilizzo il teorema centrale del limite.

Consideriamo una successione di variabili casuali $X_1, X_2, \dots, X_n$ i.i.d. e supponiamo che esistano $E[X_i] = \mu$, $V[X_i] = \delta^2$. Vogliamo vedere a che cosa converge $\overline{X}_n$. Vogliamo però una convergenza ad una vera e propria variabile casuale o ad una distribuzione, non ad una costante, come abbiamo visto in precedenza. Devo quindi standardizzare $\overline{X}_n$ facendo una semplice trasformazione lineare:

$$\overline{Z}_n = \frac{\overline{X}_n - E[\overline{X}_n]}{\sqrt{V[\overline{X}_n]}}$$

Una successione di questo tipo $\overline{Z}_n$ ha media e varianza $\forall n$ pari a:

$$E[\overline{Z}_n] = 0$$

$$V[\overline{Z}_n] = 1$$

Si avrà quindi:

$$\overline{Z}_n = \frac{\overline{X}_n - \mu}{\sqrt{\frac{\delta^2}{n}}} = \sqrt{n} \cdot \left(\frac{\overline{X}_n - \mu}{\delta} \right) \rightarrow N(0,1)$$

Il teorema centrale del limite ci dice che questa successione converge in distribuzione ad una normale $N(0, 1)$. Abbiamo quindi che, presa una **qualsiasi** variabile causale, la media campionaria standardizzata tende ad una normale. Si avrà che la funzione di ripartizione di $\overline{Z}_n$ tende alla funzione di ripartizione della normale $N(0, 1)$ per $n \rightarrow +\infty$:

$$p(\overline{Z}_n \leq z) \rightarrow \Phi(z)$$

$\forall z \in R$.

Abbiamo in definitiva che:

$$\overline{Z}_n = \sqrt{n} \cdot \left(\frac{\overline{X}_n - \mu}{\delta} \right) \approx N(0,1)$$

Quindi ad esempio $p(\overline{Z}_n \leq 1) \approx \Phi(1)$. Se voglio invece ad esempio calcolare $p(\overline{X}_n \leq 5)$, tenendo conto che conosco la media, la varianza ed n , avrò:

$$p\left(\sqrt{n} \cdot \left(\frac{\overline{X}_n - \mu}{\delta} \right) \leq \sqrt{n} \cdot \left(\frac{5 - \mu}{\delta} \right)\right) \approx \Phi\left(\sqrt{n} \cdot \left(\frac{5 - \mu}{\delta} \right)\right)$$

Quanto grande deve essere n perché valga questo, cioè perché si possa utilizzare il teorema centrale del limite? Dipende dal fatto che le variabili siano più o meno simmetriche. In generale se $n \geq 30$ l'approssimazione è già buona.

Esempio

Vogliamo valutare l'efficacia di una certa dieta alimentare per ridurre il tasso di colesterolo. Estraiamo a caso 100 persone ($n = 100$), somministriamo la dieta e poi verifichiamo i risultati. La dieta è efficace se almeno 65 persone su 100 hanno abbassato il livello di colesterolo. Consideriamo $X_1, X_2, \dots, X_{100}$ variabili, dove ciascuna variabile $X_i \sim Be(\theta)$, assume il valore 1 se si ha un successo, il valore 0 in caso contrario. Ci sarà una certa probabilità che la dieta sia efficace $p(X_i = 1) = \theta$. Avremo quindi che, se $S_{100} \geq 65$ allora la dieta è efficace, in caso contrario è inefficace. $\theta = 1/2$, essendo ogni variabile una Bernoulliana. Ma se assumiamo che $\theta = 1/2$, ipotizziamo che la dieta sia inefficace. Calcoliamo la probabilità che la dieta venga prescritta quando è inefficace, cioè:

$$p(S_{100} \geq 65, \theta = 1/2)$$

Dato che $X_i \sim Be(\theta) \sim Bi(1, \theta)$, si avrà che:

$$X = S_{100} = \sum_{i=1}^{100} X_i \sim Bi(100, 1/2)$$

Vogliamo in pratica calcolare $p(X \geq 65)$. Possiamo ad esempio utilizzare la funzione di probabilità della binomiale, ma è poco agevole. In alternativa utilizziamo il teorema centrale del limite.

Essendo $X_i \sim Be(\theta)$ e variabili i.i.d., abbiamo che $E[X_i] = \theta$, $V[X_i] = \theta \cdot (1 - \theta)$. Nel nostro caso, quindi, abbiamo che $E[X_i] = 1/2$, $V[X_i] = 1/4$. Si avrà:

$$\overline{Z}_n = \left(\frac{S_{100} - E[S_{100}]}{\sqrt{E[S_{100}]}} \right) \approx N(0,1)$$

$$\overline{Z}_n = \left(\frac{S_{100} - 100 \cdot \frac{1}{2}}{\sqrt{100 \cdot \frac{1}{4}}} \right) \approx N(0,1)$$

$$\overline{Z}_n = \left(\frac{S_{100} - 50}{5} \right) \approx N(0,1)$$

Voglio calcolare $p(\overline{Z}_n \geq 65)$:

$$p(S_{100} \geq 65) = p\left(\frac{S_{100} - 50}{5} \geq \frac{65 - 50}{5}\right) \approx p(\overline{Z}_n \geq 3) = \Phi(3) = 0,0013$$

che è la probabilità che la dieta venga prescritta quando non è efficace.

Tavole statistiche

Tavola I. Valori della f.r. $\Phi(z)$ della v.c. Normale standardizzata:

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx.$$

<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Tavole statistiche

segue Tavola I.

[illegible]