DISTRIBUZIONI UNIDIMENSIONALI

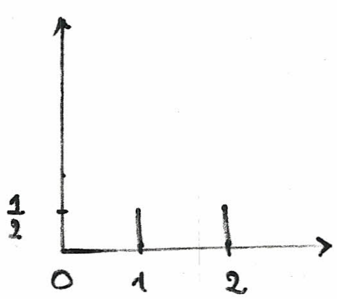
Introduciamo ora alcune famiglie parametriche di funzioni di densità unidimensionali.

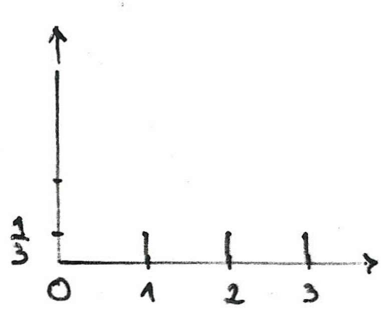
DISTRIBUZIONI DISCRETE

a) X v.a. discreta e finita

1) UNIFORME

$$f(x) = f(x;n) = \begin{cases} \frac{1}{n} & x = 1, \dots, n \\ 0 & \text{altrove} \end{cases}$$





$$x_j = j \qquad j = 1, \dots, n$$

$$\underline{E[X]} = \sum_{j=1}^{n} j \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$\underline{var[X]} = \sum_{j=1}^{n} j^2 \cdot \frac{1}{n} - \left(\frac{n+1}{2}\right)^2$$

$$= \frac{n(n+1)(2n+1)}{6} \cdot \frac{1}{n} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

$$\underline{m(t)} = E\left[e^{tX}\right] = \sum_{J=1}^{n} e^{tx_J} \cdot \frac{1}{n} = \sum_{J=1}^{n} e^{tJ} \cdot \frac{1}{n} = \frac{1}{n} \sum_{J=1}^{n} (e^t)^J$$

$$= \frac{1}{n}\left(e^t + e^{2t} + \dots + e^{nt}\right)$$

$$= \frac{1}{n} e^t \left(\frac{1 + e^t + \dots + e^{(n-1)t}}{(1-e^t)}\right) \cdot (1-e^t)$$

$$= \frac{e^t}{n} \frac{(1-e^{nt})}{1-e^t}$$

<u>Esempio</u> : Lotteria

90 biglietti numerati da 1 a 90.

l'organizzatore tiene per se 6 biglietti il cui numero x è tale che $9 < x < 16$. (cioè 10, 11, 12, 13, 14, 15)

1) Calcolare la probabilità che l'organizzatore vinca il 1° premio che corrisponde al 1° numero estratto su una certa ruota in un dato giorno.

2) Qual è la v.a. idonea?

1) • Prob. classica $p = \frac{f}{n} = \frac{6}{90} = \frac{1}{15}$

$A = \{$ l'organizzatore vince $\}$

• $P[A] = P\left[\bigcup_{i=10}^{15} A_i\right] \underset{\uparrow}{=} \sum_{i=10}^{15} P[A_i] = 6 \cdot \frac{1}{90} = \frac{1}{15}$

eventi indip. e incomp.

2) La v.a. X : $A_x \longleftrightarrow x \qquad x = 1, \ldots, 90$

Tutti i valori di X sono equiprobabili

$\Rightarrow$ X è v.a. discreta uniforme di parametro $n = 90$

$$P[X = x] = \frac{1}{90} \qquad x = 1, 2, \ldots, 90$$

$$f(x) = \begin{cases} \frac{1}{90} & x = 1, \ldots, 90 \\ 0 & \text{altrove} \end{cases}$$

$$P[9 < X < 16] = P[9 < X \leq 15] = F(15) - F(9)$$

$$= \sum_{x_j \leq 15}' f(x_j) - \sum_{x_j \leq 9}' f(x_j)$$

$$= 15 \cdot \frac{1}{90} - 9 \cdot \frac{1}{90} = \frac{1}{15}$$

se avessi chiesto

$$P[9 \leq X < 16] = P[9 \leq X \leq 15] = F(15) - F(9) + P(X = 9)$$

$$= 15 \cdot \frac{1}{90} - 9 \cdot \frac{1}{90} + \frac{1}{90} = \frac{1}{18}$$

2) BERNOULLI

$$f(x) = f(x,p) = \begin{cases} p^{x}(1-p)^{1-x} & x=0 \text{ o } x=1 \\ 0 & \text{altrove} \end{cases}$$

$q = 1-p \quad , \quad 0 \leq p \leq 1.$

Utilizzo: negli esperimenti il cui esito è dicotomico. "successo" o "insuccesso".

$X = 1$ successo

$X = 0$ insuccesso

$$f(X=0) = P[X=0] = 1-p$$

$$f(X=1) = P[X=1] = p$$

se p è la probabilità di successo.

$$E[X] = 1 \cdot p + 0 \cdot q = p$$

$$\text{var}[X] = 1^2 \cdot p + 0^2 \cdot q - p^2 = p(1-p) = pq$$

$$m(t) = E[e^{tX}] = p e^{tx}\Big|_{x=1} + q e^{tx}\Big|_{x=0}$$

$$= pe^{t} + q$$

In altre parole una v.a. è bernoulliana se può assumere solo i valori 0 e 1.

3) BINOMIALE

$$f(x) = f(x; n, p) = B(x; n, p) = \begin{cases} \binom{n}{x} p^x q^{n-x} & x=0,\dots,n \\ 0 & \text{altrove} \end{cases}$$

$0 \le p \le 1$, $q = 1-p$, $n > 0$.

Ricordando la formula del BINOMIO DI NEWTON

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

calcoliamo

$$\underline{m(t)} = E[e^{tX}] = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x}$$

$$= (pe^t + q)^n$$

$$m'(t) = n (pe^t + q)^{n-1} pe^t$$

$$m'(0) = np$$

$$m''(t) = n(n-1)(pe^t+q)^{n-2}(pe^t)^2 + n(pe^t+q)^{n-1} pe^t$$

$$= npe^t (pe^t+q)^{n-2} [(n-1)pe^t + pe^t + q]$$

$$= npe^t (pe^t+q)^{n-2} (npe^t + q)$$

$$m''(0) = np(np+q)$$

$E[X] = m'(0) = np$

$var[X] = E[X^2] - (E[X])^2 = m''(0) - [m'(0)]^2$

$= np(np+q) - n^2p^2 = npq$

PROPRIETÀ

$P[X < x] = P[X \leq x-1]$

$P[X > x] = 1 - P[X \leq x]$

$P[X \geq x] = 1 - P[X < x] = 1 - P[X \leq x-1]$

$P[X = x] = P[X \leq x] - P[X < x]$
$= P[X \leq x] - P[X \leq x-1]$

N.B. Nel discreto $P[X=x] \neq P[X \leq x]$.

OSS: La f. di densità di Bernoulli è la binomiale con $n = 1$.

SCHEMA DI BERNOULLI

Una situazione molto frequente è quella in cui si presenta una successione di esperimenti casuali tra loro indipendenti ognuno dei quali dà luogo a due possibili risultati: "successo", "insuccesso".

- Se esistono solo due possibili risultati mutuamente esclusivi,
- Se la probabilità di successo p è la stessa in ogni prova,
- se le prove sono indipendenti

$\Rightarrow$ X v.a. che conta il n° di successi in n prove è detta BINOMIALE.

<u>Esempio</u>: Lancio di una moneta
Esce testa con probabilità p, $0 \leq p \leq 1$, esce croce con probabilità $q = 1-p$.
La moneta viene lanciata n volte.

Qual è la probabilità di ottenere k volte testa ed $n-k$ volte croce?

$$\Omega = \{\omega : \omega = (\omega_1, \dots, \omega_n) \text{ dove } \omega_i = 1 \text{ oppure } \omega_i = 0 \quad i = 1, \dots, n\}$$

se $1 =$ testa
$0 =$ croce

$$\omega = (\underbrace{1, \dots, 1}_{k \text{ volte}}, \underbrace{0, \dots, 0}_{n-k \text{ volte}})$$

Sia

$A_i = \{ \omega : \omega_i = 1 \} = \{$ il risultato dell'i-esimo lancio è T $\}$

$P[A_i] = p$

$\omega = A_1 \cap \dots \cap A_k \cap \bar{A}_{k+1} \cap \dots \cap \bar{A}_n$

L'indipendenza implica.

$$P[\{\omega\}] = P[A_1] \cdots P[A_k] \cdot P[\bar{A}_{k+1}] \cdots P[\bar{A}_n]$$
$$= p^k q^{n-k}$$

Quante sono le configurazioni possibili?

Tante quante le combinazioni di n oggetti di cui "k" ed "$n-k$" uguali fra loro.

Dal calcolo combinatorio questo numero è

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

$$f(x) = P[X = x \text{ successi}] = \underline{\binom{n}{x} p^x (1-p)^{n-x}}$$

$$F(x) = P[X \leq x] = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$$

è la funzione di ripartizione.

Esempio

Il 5% dei chip di memoria prodotti da una macchina sono difettosi. Determinare la probabilità che su 4 chip scelti a caso

- uno sia dif.
- nessuno sia dif.
- meno di due siano dif.

Se X è v.a. che indica il n° dei chip difettosi su un totale di 400 chip determinare μ_x e σ_x.

$n = 4 \qquad p = 0.05$

- $P(X=1) = \binom{4}{1}(0.05)^1(0.95)^3 = 0.171$
- $P(X=0) = \binom{4}{0}(0.05)^0(0.95)^4 = 0.816$
- $P(X<2) = P(X=0) + P(X=1) = 0.985$

$n = 400 \qquad p = 0.05$

$\mu_x = np = 20$

$\sigma_x^2 = npq = 400 \cdot 0.05 \cdot 0.95 = 19$

$\sigma_x = \sqrt{19} \simeq 4.36.$

RELAZIONE DI RICORRENZA

$$P[X = x+1] = \frac{n-x}{x+1} \frac{p}{1-p} P[X = x]$$

4) IPERGEOMETRICA

$$f(x) = f(x;\, N, D, n) = \begin{cases} \dfrac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} & x = 0, 1, \dots, n \\ 0 & \text{altrove} \end{cases}$$

$N \in \mathbb{Z}^+$

$D \leq N \in \mathbb{Z}^+ \cup \{0\}$

$n \leq N \in \mathbb{Z}^+$

Si può provare:

$$E[X] = n\,\frac{D}{N}$$

$$\text{var}[X] = \frac{n D (N-D)(N-n)}{N^2 (N-1)}$$

<u>Utilizzo</u>: nel campionamento o estrazioni senza reimmissione.

<u>Esempio</u>: Una scatola di N pezzi contiene D pezzi difettosi e N − D accettabili.
Si estraggono senza reimmissione e in modo casuale n pezzi dalla scatola di cui x saranno difettosi ed $n-x$ accettabili.

X v.a. ipergeometrica conta il n° di pezzi difettosi contenuti nel campione estratto.

Esempio

Un rivenditore acquista chip a lotti di 10.

In ogni lotto controlla a caso 3 chip.

Decide di accettare il lotto solo se nessuno dei 3 pezzi è difettoso.

Sapendo che il 30% dei lotti ha 4 pezzi difettosi ed il 70% ha 1 pezzo difettoso, calcolare la percentuale di rifiuto del rivenditore.

$A = \{ \text{accetta il lotto} \}$

$R = \bar{A} = \{ \text{rifiuto il lotto} \}$

$D_i = \{ \text{il lotto ha "i" pezzi difettosi} \}$

$$P[A] = P[A \mid D_4] \cdot \underbrace{P[D_4]}_{=\frac{30}{100}} + P[A \mid D_1] \underbrace{P[D_1]}_{=\frac{70}{100}} \quad \text{TH. P. TOTALI}$$

$$P[A \mid D_4] = \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} = \frac{1 \cdot \frac{6!}{3!\,3!}}{\frac{10!}{3!\,7!}} = \frac{1}{6}$$

$$P[A \mid D_1] = \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} = \frac{1 \cdot \frac{9!}{3!\,6!}}{\frac{10!}{3!\,7!}} = \frac{7}{10}$$

$$\Rightarrow P[A] = \frac{1}{6} \cdot \frac{3}{10} + \frac{7}{10} \cdot \frac{7}{10} = \frac{54}{100}$$

$$P[R] = 1 - P[A] = 46\%$$

5) POISSON b) X. v.a. discreta, ma non finita

$$f(x) = f(x;\lambda) = P(x,\lambda) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & x=0,1,2,\dots \\ 0 & \text{altrove} \end{cases}$$

$\lambda > 0$.

$$m(t) = E[e^{tX}] = \sum_{x=0}^{\infty} \frac{e^{tx} e^{-\lambda} \lambda^{x}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^{t})^{x}}{x!} \overset{(\cdot)}{=}$$

ANALISI $\Rightarrow$ $e^{x} = \sum_{n=0}^{\infty} \dfrac{x^{n}}{n!}$

$$\overset{(\cdot)}{=} e^{-\lambda}(e^{\lambda e^{t}}) = e^{\lambda(e^{t}-1)}$$

$$m'(t) = e^{\lambda(e^{t}-1)} \lambda e^{t}$$

$$m'(0) = \lambda$$

$$m''(t) = \lambda e^{t} e^{\lambda(e^{t}-1)}(\lambda e^{t}+1)$$

$$m''(0) = \lambda(1+\lambda)$$

$\Rightarrow$ $\underline{E[X] = \lambda}$

$\underline{var[X]} = m''(0) - m'(0)^{2} = \lambda(1+\lambda) - \lambda^{2} = \underline{\lambda}$

<u>Utilizzo</u>: negli esperimenti in cui è possibile fare un qualche tipo di conteggio, in un certo intervallo di tempo o di spazio.

P. es:
- n° chiamate telefoniche ricevute ogni ora da un centralino
- n° di clienti che entrano in un ufficio in un intervallo di tempo.

- n° incidenti mortali stradali in una regione in una settimana
- il n° di refusi in una pagina di un libro.

$\Rightarrow$ X v.a. che conta il n° di volte in cui si verifica un EVENTO RARO in un certo intervallo.

$\lambda > 0$ è il n° medio di realizzazioni dell'evento nell'intervallo considerato.

La funzione di ripartizione è:

$$F(x) = P[X \leq x] = \sum_{k=0}^{x} \frac{e^{-\lambda} \lambda^k}{k!}$$

PROPRIETÀ

le stesse definite per la binomiale

RELAZIONE DI RICORRENZA

$$P[X = x+1] = \frac{\lambda}{x+1} P[X = x]$$

Esempio

In un'azienda ogni giorno si assentano in media 1.8 operai. Calcolare la probabilità che un certo giorno ci siano 3 operai assenti contemporaneamente.

$\lambda = 1.8$ (piccolo)

$$P[X = 3] = e^{-1.8} (1.8)^3 / 3! = 0.16$$

Esempio

Il 3% delle lampadine uscite da una fabbrica è difettoso.
In un campione di 100 lampadine, calcolare la probabilità che 2 siano difettose.

1) con la binomiale

$n = 100 \;, \; x = 2 \;, \; p = 0.03$

$$P[X=2] = \binom{100}{2} (0.03)^2 (0.97)^{98} \approx 0.225$$

2) con la poissoniana

$n = 100 \;, \; x = 2 \;, \; p = 0.03 \qquad \lambda = np = 3$

$$P[X=2] = \frac{e^{-3}\, 3^2}{2!} \approx 0.224$$

APPROSSIMAZIONE

$$B(x; n, p) \xrightarrow[p \to 0]{n \to \infty} P(x; \lambda = np) \qquad \underset{np}{\underset{\parallel}{E[X_B]}} \xrightarrow[n \to \infty]{} \underset{\lambda}{\underset{\parallel}{E[X_P]}}$$

Empiricamente basta che

$n \geq 50$ e $p \leq 0.1$

X. v.a. di Poisson può essere utilizzato come approssimazione di una binomiale di parametri (n, p) quando n è molto grande e p molto piccolo.

cioè $\quad B(x; n, p) \xrightarrow[\substack{n\to\infty \\ p\to 0}]{} P(x; \lambda)$

$$B(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

Poniamo $np = \lambda$

$$= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)\cdots(n-x+1)}{x!} \frac{\lambda^x}{n^x} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^x}$$

Ricordando che

$$\left(1+\frac{1}{n}\right)^n \xrightarrow[n\to\infty]{} e \quad \left(1-\frac{1}{n}\right)^n \xrightarrow[n\to\infty]{} e^{-1} \quad \Rightarrow \left(1-\frac{\lambda}{n}\right)^n \xrightarrow[n\to\infty]{} e^{-\lambda}$$

mentre

- $\left(1-\frac{\lambda}{n}\right)^x \xrightarrow[n\to\infty]{} 1$
- $\frac{n(n-1)\cdots(n-x+1)}{n^x} \xrightarrow[n\to\infty]{} 1 \quad \Leftarrow \begin{cases} x=1 & \frac{n}{n} \\ x=2 & \frac{n(n-1)}{n^2} \\ x=3 & n(n-1)(n-2)/n^3 \\ \vdots & \end{cases}$

$$\Rightarrow \xrightarrow[n\to\infty]{} \frac{\lambda^x e^{-\lambda}}{x!} = P(x; \lambda)$$

6) GEOMETRICA

$$f(x) = f(x,p) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \dots \\ 0 & \text{altrove} \end{cases}$$

$0 < p \leq 1$, $q = 1-p$

Si può provare:

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

$\mathrm{var}[X] = \frac{1-p}{p^2} = q/p^2$

X.v.a. è detta TEMPO DI ATTESA, cioè quanto, in termini di n° di insuccessi, si deve attendere prima di ottenere un successo (es: il ritardo di un numero nel gioco del lotto)

$E[X] = \frac{1}{p}$ è detta TEMPO DI RITORNO, cioè la media del n° di ripetizioni da tentare per avere un successo.

OSSERVAZIONE

Per studiare l'istante di un primo successo bisogna considerare un numero arbitrariamente grande di prove e quindi non può più essere utilizzato lo schema di Bernoulli che descrive un numero prefissato di prove.

PROPRIETÀ

La mancanza di memoria.

In uno schema successo - insuccesso supponiamo di non aver ottenuto alcun successo nelle prime k prove.

Qual è la probabilità di dover attendere ancora n prove per avere il primo successo?

È la stessa che si avrebbe se le prime k prove non avessero avuto luogo.

Ciò è ovvio poiché in uno schema a prove ripetute indipendenti i risultati delle prime non influenzano le successive.

SCHEMA

Ω sp. campione

$\{E_n\}$ $n = 1, 2, \ldots$ successione di eventi indipendenti

$P[E_n] = p$, $q = 1 - p$

X v.a. che rappresenta il tempo di attesa del 1° successo cioè

$X = n$ se $E_1, \ldots, E_{n-1}$ non si verificano

E_n si verifica.

$$P[X=n] = P[\bar{E}_1 \cap \ldots \cap \bar{E}_{n-1} \cap E_n] =$$

$$= P[\bar{E}_1] \cdots P[\bar{E}_{n-1}] \cdot P[E_n]$$

$$= q^{(n-1)} p$$

$$= p(1-p)^{n-1}$$

Esempio

Ritardo di un numero nel gioco del lotto.

E_n = { nella n-esima settimana, a partire da una settimana "0", è estratto un certo numero su una data ruota }

Calcolare la probabilità che il numero abbia accumulato un RITARDO di n settimane.

$$P[X>n] = 1-P[X\leq n] = 1-F(n)$$

$$= 1 - \sum_{x=1}^{n} p(1-p)^{x-1} = 1 - \sum_{x=1}^{n} (1-q)q^{x-1}$$

$$= 1-(1-q+q-q^2+q^2-q^3+ \quad +q^{n-1}-q^n)$$

$$= 1-(1-q^n) = q^n = (1-p)^n$$

Calcolare la distribuzione di probabilità dell'attesa residua

Cioè calcolare la probabilità che un certo n° venga estratto solo tra K settimane sapendo che ha un ritardo di m settimane.

$$P[X=k+m \mid X>m] = \frac{P[X=m+k, X>m]}{P[X>m]}$$

$$= \frac{P[X=m+k]}{P[X>m]}$$

$$= \frac{p\,q^{(m+k)-1}}{q^m}$$

$$= p(1-p)^{k-1} = P[X=k]$$

⇒ mancanza di memoria

OSSERVAZIONE

Se l'evento è il superamento di una certa soglia "a" che avviene con probabilità $q = P[X > a]$, il tempo di ritorno di a abbiamo visto che è:

$$T = E[X = a] = \frac{1}{p} = \frac{1}{1-q}$$

misurato in n° di osservazioni.

Esempio

Consideriamo il massimo mensile di portata al colmo di un corso d'acqua supponendo che la distribuzione sia identica per ogni mese.

Sapendo che la probabilità di non superamento è 0.98 calcolare il tempo di ritorno.

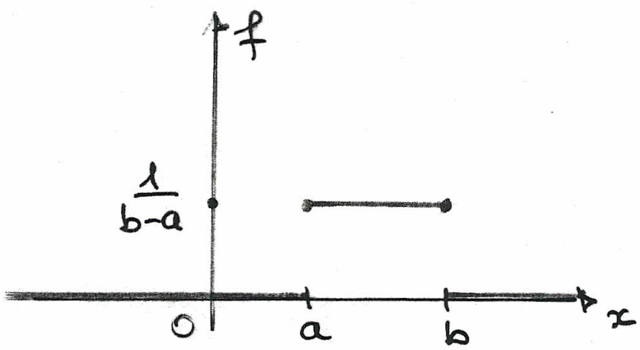
$$q = P[X > a] = 0.98 \Rightarrow T = \frac{1}{1-0.98} = \frac{1}{0.02} = 50 \text{ mesi}$$

Calcolato in anni $\frac{T}{12} = 4,17$ anni.

DISTRIBUZIONI CONTINUE

1) RETTANGOLARE

$$f(x) = f(x; a, b) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{altrove} \end{cases}$$



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

$$\mathrm{var}[X] = \int_a^b x^2 \cdot \frac{1}{b-a}\, dx - \left[\frac{a+b}{2}\right]^2$$

$$= \frac{(b^3-a^3)}{3(b-a)} - \frac{(a+b)^2}{4} = \boxed{\frac{(b-a)^2}{12}}$$

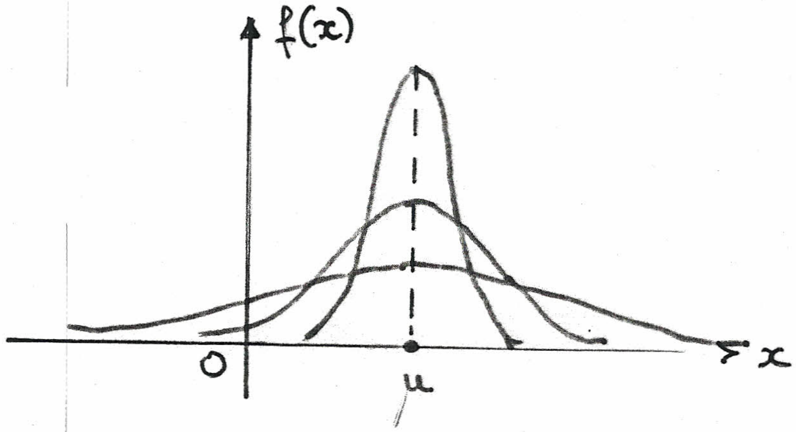
$$m(t) = E[e^{tx}] = \int_a^b e^{tx} \cdot \frac{1}{b-a}\, dx$$

$$= \frac{1}{b-a} \cdot \frac{1}{t}\left(e^{tb} - e^{ta}\right)$$

2) NORMALE

$$f(x) = N(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

$\mu \in \mathbb{R}$, $\sigma > 0$.



$x = \mu$ pto di MAX

$x = \mu \pm \sigma$ flessi

- N simmetrica rispetto a μ. (media ≡ mediana)
- N ha un max per $x = \mu$. (media ≡ moda)
- cambiare il valore di μ equivale a traslare su Ox il grafico di N senza deformarlo
- cambiare il valore di σ equivale a modificare la forma del grafico di N senza traslarlo.

$$\underline{m(t)} = E\left[e^{tX}\right] = \int_{-\infty}^{+\infty} e^{tx} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

lunghi calcoli

$$\overset{(*)}{=} \underline{e^{\mu t + \frac{\sigma^2 t^2}{2}}}$$

$$m'(t) = (\mu + \sigma^2 t)\, e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$$m'(0) = \mu$$

$$m''(t) = \sigma^2 e^{\mu t + \frac{\sigma^2 t^2}{2}} + (\mu + \sigma^2 t)^2 e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$$m''(0) = \sigma^2 + \mu^2$$

$$\underline{E[X]} = m'(0) = \boxed{\mu}$$

$$\underline{var[X]} = E[X^2] - (E[X])^2 = m''(0) - (m'(0))^2 = \boxed{\sigma^2}$$

La funzione di ripartizione si deve lasciare indicata come:

$$\underset{\substack{\| \\ P[X \le x]}}{F(x;\mu,\sigma)} = \int_{-\infty}^{x} N(u;\mu,\sigma)\, du$$

- X v.a. normale è detta normale standard o ridotta se $\mu = 0$, $\sigma = 1$.

- X v.a. normale $\Rightarrow$ $Z = \dfrac{X-\mu}{\sigma}$ è v.a. ridotta.

Infatti:

$$\underline{E[Z]} = \frac{1}{\sigma} E[X-\mu] = \boxed{0}$$

$$\underline{var[Z]} = \frac{1}{\sigma^2} var[X-\mu] = \frac{1}{\sigma^2}\Big\{E[(X-\mu)^2] - (\underset{\substack{||| \\ 0}}{E[X-\mu]})^2\Big\}$$

$$= \frac{1}{\sigma^2}\Big[\underset{\substack{\| \\ \sigma^2+\mu^2}}{E[X^2]} - 2\underset{\substack{\| \\ \mu}}{E[X]}\mu + \mu^2\Big] = \boxed{1}$$

La funzione di densità di v.a. standard:

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

indipendente da μ e da σ.

La funzione di ripartizione per v.a. normale standard è:

$$\Phi(z) = \int_{-\infty}^{z} \varphi(u)\, du$$

PROPRIETÀ

Sia X v.a. normale (μ, σ^2).

1. $P[a \leq X \leq b] = P\left[\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right]$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

2. $P[X \geq a] = P[a \leq X \leq \infty] = P\left[\frac{a-\mu}{\sigma} \leq Z \leq \infty\right]$

$$= \Phi(\infty) - \Phi\left(\frac{a-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

si pone $\Phi(-\infty) = 0$ e $\Phi(+\infty) = 1$

N.B.: La mediana di una v.a. continua con funzione di ripartizione F è il valore m tale che $F(m) = \frac{1}{2}$.

1 2 3
<u>media</u>, <u>mediana</u>, <u>moda</u>

1) $x_1, \dots, x_n$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

2) $x_1, \dots, x_n$ ordine crescente

n dispari x_k posizionato in $\frac{(n+1)}{2}$

n pari x_J ottenuto facendo la media aritmetica tra il valore posizionato in $\frac{n}{2}$ e quello in $\left(\frac{n}{2}+1\right)$

3) se esiste, il valore che ha frequenza max.

La mediana di v.a. continua con F funz. di rip. è il valore m: $F(m) = \frac{1}{2}$.

CALCOLI LUNGHI (*) ①

$$e^{tx}\cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2} = \underbrace{e^{t\mu}\, e^{-t\mu}}\cdot e^{tx}\cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$= e^{\mu t}\left[e^{t(x-\mu)}\cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\right]$$

$$= e^{\mu t}\, e^{\left[t(x-\mu)-\frac{1}{2\sigma^2}(x-\mu)^2\right]} \overset{(\circ)}{=}$$

$$t(x-\mu)-\frac{1}{2\sigma^2}(x-\mu)^2 = -\frac{(x-\mu)^2-2\sigma^2 t(x-\mu)}{2\sigma^2} \quad \underline{\pm\,\sigma^4 t^2} \text{ al numeratore}$$

$$= -\frac{\left[(x-\mu)-\sigma^2 t\right]^2-\sigma^4 t^2}{2\sigma^2}$$

$$= -\frac{\left[(x-\mu)-\sigma^2 t\right]^2}{2\sigma^2}+\frac{\sigma^2 t^2}{2}$$

$$\overset{(\circ)}{=} e^{\mu t+\frac{\sigma^2 t^2}{2}}\quad e^{-\frac{(x-\mu-\sigma^2 t)^2}{2\sigma^2}}$$

$$m(t) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{\mu t+\frac{\sigma^2 t^2}{2}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu-\sigma^2 t)^2}{2\sigma^2}}\,dx$$

devo dimostrare che

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu-\sigma^2 t)^2}{2\sigma^2}}\,dx = 1.$$

<u>Cambio variabili</u>

$$\frac{x-\mu-\sigma^2 t}{\sqrt{2}\,\sigma} = y \qquad dx = dy\cdot\sqrt{2}\,\sigma \qquad \begin{matrix} x=\infty & y=\infty \\ x=-\infty & y=-\infty\end{matrix}$$

$$\frac{1}{\cancel{\sigma}\cancel{\sqrt{2}}\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-y^2}\,\cancel{\sqrt{2}}\,\cancel{\sigma}\,dy = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-y^2}\,dy.$$

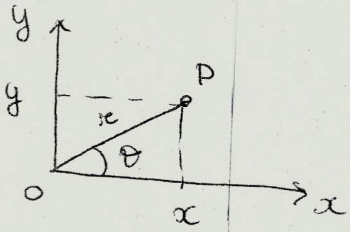
pertanto devo dimostrare che

$$\int_{-\infty}^{+\infty} e^{-y^2}\,dy = \sqrt{\pi}$$

Considero

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = I$$

$$\underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{+\infty}}_{\int_{\mathbb{R}^2}} e^{-x^2-y^2}\,dx\,dy = \int_{-\infty}^{\infty} e^{-x^2}\,dx \cdot \int_{-\infty}^{\infty} e^{-y^2}\,dy = I^2$$



(x,y) <u>coord. cart</u> $\longrightarrow$ <u>coord. polari</u> (r, θ)

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \qquad \begin{matrix} 0 < r < \infty \\ 0 < \theta < 2\pi \end{matrix}$$

Jacobiano della trasformazione

$$J = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

$$|J| = r\left(\cos^2\theta + \sin^2\theta\right) = r$$

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy = \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,r\,dr\,d\theta$$

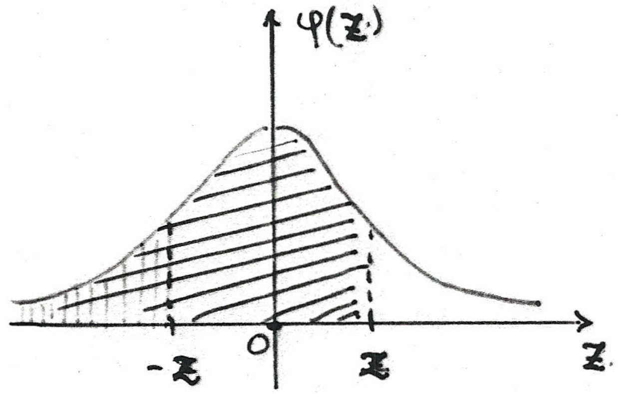
$$= \underbrace{\int_0^{2\pi} d\theta}_{\substack{= \cancel{2\pi} \\ [\theta]_0^{2\pi} \\ = 2\pi}} \int_0^{\infty} r\,e^{-r^2}\,dr$$

$$r\,e^{-r^2} = -\frac{1}{2}\cdot\underbrace{\left(-2r\,e^{-r^2}\right)}_{f' \,\to\, f = e^{-r^2}}$$

$$= 2\pi \cdot \left(-\frac{1}{2}\right) \underbrace{\left[e^{-x^2} \right]_0^\infty}_{= -1} = \pi .$$

$$I^2 = \pi \quad \Rightarrow \quad \boxed{I = \sqrt{\pi}} \quad \text{c.v.d.}$$

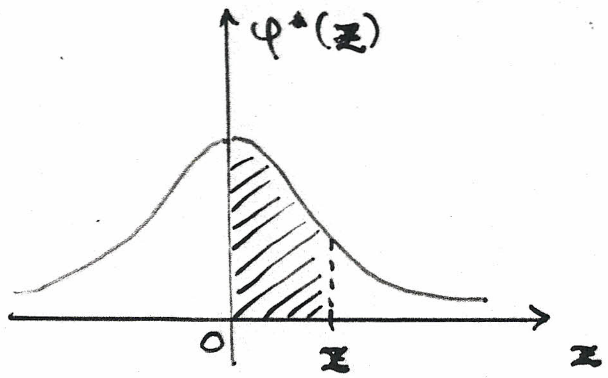
OSS.: $\Phi(z) = 1 - \Phi(-z)$ $\qquad \boxed{\Phi(-z) = 1 - \Phi(z)}$



$$\Phi^*(z) = \int_{\boxed{0}}^{z} \varphi(u)\,du$$

$$\Phi(z) = \frac{1}{2} + \Phi^*(z)$$

$$\Phi^*(-z) = -\Phi^*(z)$$



UTILIZZO: ha un ruolo centrale in statistica (TLC) e nella teoria degli errori.
Risulta essere la forma limite di alcune distribuzioni di probabilità.

Esempio

Il peso di 2 confezioni è una X v.a. normale con $\mu = 250$ g., $\sigma = 3$g. Calcolare la probabilità che il peso sia minore di 245 g.

$$Z = \frac{X-\mu}{\sigma} = \frac{X-250}{3}$$

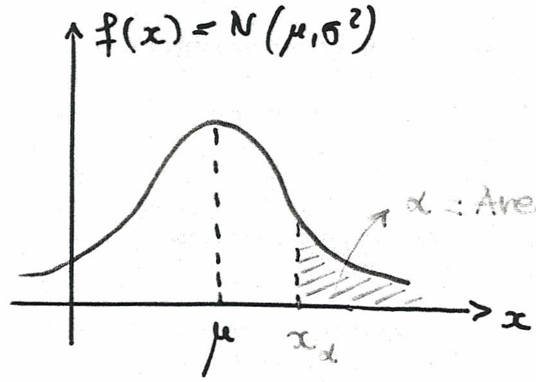
se $x = 245 \Rightarrow Z = -1.67$

$$P(X<245) = P(Z<-1.67) = \Phi(-1.67) = 1 - \Phi(\underline{1.67})$$

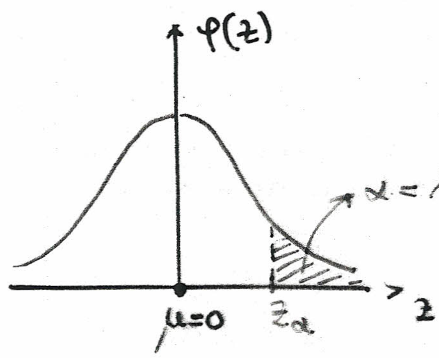
$$= 1 - \underline{0.9525} = 0.0475$$

PROBLEMA INVERSO

Dato $\alpha \in (0,1)$, determinare x_α : $P(X > x_\alpha) = \alpha$



ovvero per la normale standard, z_α: $P(Z > z_\alpha) = \alpha$



Esempio

Dato Z, determinare z_α:

1. $P(Z < z_\alpha) = 0.9953$
2. $P(Z > z_\alpha) = 0.2743$
3. $P(0 \leq Z \leq z_\alpha) = 0.377$
4. $P(|Z| \leq z_\alpha) = 0.5762$
5. $P(z_\alpha < Z < 1.6) = 0.7865$

1. dalle tavole $z_\alpha = 2.6$

2. $P(Z < z_\alpha) = 1 - P(Z > z_\alpha) = 1 - 0.2743 = 0.7257$

dalle tavole $z_\alpha = 0.6$

3. $\underbrace{P(0 \leq Z \leq z_\alpha)}_{= 0.377} = P(Z < z_\alpha) - 0.5$

$\Rightarrow P(Z < z_\alpha) = 0.877$ dalle tavole $z_\alpha = 1.16$

4. $\underbrace{P(|Z| < z_\alpha)}_{= 0.5762} = P(-z_\alpha < Z < z_\alpha) \overset{\text{simmetria } Z}{=} 2P(0 < Z < z_\alpha)$

$= 2[P(Z < z_\alpha) - 0.5]$

$= 2P(Z < z_\alpha) - 1$

$\Rightarrow P(Z < z_\alpha) = \dfrac{1 + 0.5762}{2} = 0.7881$

dalle tavole $z_\alpha = 0.8$.

5. $\underbrace{P(z_\alpha < Z < 1.6)}_{= 0.7865} = \underbrace{P(Z < 1.6)}_{= 0.9452 \text{ (dalle tavole)}} - P(Z < z_\alpha)$

$\Rightarrow P(Z < z_\alpha) = \underline{0.1587} < 0.5$!

$\Rightarrow$ z_α si trova alla sinistra dell' origine.

Cerco z_α^* simmetrico di z_α. (a destra dell' origine)

$$P(Z < z_\alpha) = P(Z > z_\alpha^*) = 0.1587$$

ma

$$P(Z > z_\alpha^*) = 1 - P(Z < z_\alpha^*)$$

$\Rightarrow P(Z < z_\alpha^*) = 0.8413$ e dalle tavole $z_\alpha^* = 1$

perciò $z_\alpha = -1$.

Esempio

A) $P(Z \geq z_\alpha) = 1\% = 0.01 \quad \Rightarrow z_\alpha \cong 2.326$

$P(Z \geq z_\alpha) = 5\% = 0.05 \quad \Rightarrow z_\alpha \cong 1.645$

$P(Z \geq z_\alpha) = 2.5\% = 0.025 \quad \Rightarrow z_\alpha \cong 1.96$

B) X v.a. normale $\mu = 19$, $\sigma^2 = 49$ ($\Rightarrow \sigma = 7$)

x_α: $P(X > x_\alpha) = 20\% = 0.20$

$$0.20 = P(X > x_\alpha) = P\left(\frac{X-\mu}{\sigma} > \frac{x_\alpha - \mu}{\sigma}\right)$$

$$= P\Big(Z > \underbrace{\frac{x_\alpha - 19}{7}}_{= z_\alpha}\Big)$$

dalla tavola dei percentili $z_\alpha \cong 0.842$ "5"

$\Rightarrow x_\alpha \cong 19 + 7 \cdot 0.842 \cong 24.89$

Esempio

La potenza W dissipata da una resistenza è proporzionale al quadrato della differenza di potenziale V ai suoi capi:

$$W = r V^2 \qquad r \text{ costante}$$

DATI:

$r = 3$

$V \sim N(\mu = 6, \sigma = 1)$

Calcolare $E[W]$, $P[W > 120]$.

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

ora

$$E[W] = E[r V^2] = r E[V^2] = r \{ var[V] + (E[V])^2 \}$$

$$= 3[1 + 36] = 111$$

$$P[W > 120] = P[3V^2 > 120] = P[V^2 > 40] = P[V > \sqrt{40}]$$

$$Z = \frac{V - \mu_V}{\sigma_V} = V - 6$$

$$P[V > \sqrt{40}] = P[V - 6 > \sqrt{40} - 6] \cong P[Z > 0.3246]$$

$$= 1 - P[Z < \underbrace{0.3246}_{= z_\alpha}]$$

$0.32 \to 0.62552$

$0.33 \to 0.62930$

facendo la media

$0.3246 \longrightarrow 0.62741$

$$\simeq 1 - 0.62741 \simeq 0.37259$$

APPROSSIMAZIONI

$$B(x; n, p) \xrightarrow[n \to \infty]{} N(x; \mu = np, \sigma^2 = npq) \rightarrow N(0,1)$$

$\downarrow$ discreta $\qquad\qquad$ $\downarrow$ continua $\qquad\qquad$ $Z = \dfrac{X - np}{\sqrt{npq}}$

basta che ~~$np \geq 5$~~ e ~~$nq \geq 5$~~ $npq \geq 10$

$$P(x; \lambda) \xrightarrow[\lambda \to \infty]{} N(x; \mu = \lambda, \sigma = \sqrt{\lambda}) \rightarrow N(0,1)$$

$\sigma^2 = \lambda$ $\qquad$ $Z = \dfrac{X - \lambda}{\sqrt{\lambda}}$

$\downarrow$ discreta $\qquad\qquad$ $\downarrow$ continua

basta che $\lambda \geq 10$.

Ipergeometrica $\longrightarrow$ Binomiale

$$\frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} \xrightarrow[N \to \infty]{} \binom{n}{x} p^x (1-p)^{n-x}$$

$$E[X_{iper}] = E[X_{bin}]$$

$$n \frac{D}{N} = np$$

$$\Downarrow$$

$$\boxed{\frac{D}{N} = p}$$

3) GAMMA

$$f(x) = f(x; r, \lambda) = \begin{cases} \dfrac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrove} \end{cases}$$

$r, \lambda > 0$ e

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt \quad , \forall x > 0$$

detta FUNZIONE GAMMA DI EULERO.

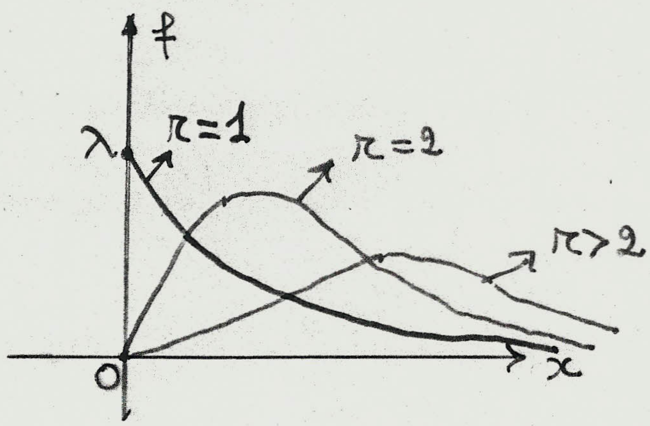
Si prova:

$E[X] = \boxed{\dfrac{r}{\lambda}}$

$var[X] = \boxed{\dfrac{r}{\lambda^2}}$

$m(t) = \left(\dfrac{\lambda}{\lambda - t}\right)^r \quad , \lambda > t$

Nel caso $r = 1$, la densità gamma è la densità esponenziale.



N.B.: $\Gamma(1) = \int_0^\infty e^{-t}\, dt = 1$

4) ESPONENZIALE

$$f(x) = f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \quad , \lambda > 0 \\ 0 & \text{altrove} \end{cases}$$

Abbiamo già ricavato:

$E[X] = \boxed{\frac{1}{\lambda}}$; $var[X] = \boxed{\frac{1}{\lambda^2}}$; $m(t) = \boxed{\frac{\lambda}{\lambda - t}}$ $\quad t < \lambda$

UTILIZZO: come modello per la durata di vita di diversi fenomeni.

- La v. aleatoria GAMMA è il corrispondente continuo della v.a. geometrica (discreta).

ASSENZA DI MEMORIA O MANCANZA DI USURA

Finchè un oggetto funziona si comporta come se fosse nuovo.

X v.a. esponenziale che rappresenta il tempo di vita di un oggetto.

=> la distribuzione del tempo di vita residua è la medesima sia che l'oggetto stia funzionando da un tempo t, sia che esso sia nuovo.

(è anche una X v.a. esponenziale)

Esempio

X v. a. esponenziale che rappresenta il n° di km percorsi da un' automobile prima che la batteria si fermi.

$\mu_X = 10'000$ km.

Calcolare la probabilità che si possa fare un viaggio di 5'000 km senza sostituire la batteria.

Per la mancanza di usura, il tempo di vita residua è v.a. esponenziale (in migliaia di km.):

$$\mu_{v.r.} = \frac{1}{\lambda} \quad \Rightarrow \quad \underline{\lambda = \frac{1}{10}}$$

$$P[v.r. > 5] = 1 - P[v.r. \leq 5] = 1 - F(5)$$

ma $\quad F(x) = 1 - e^{-\lambda x}$

perciò $\quad F(5) = 1 - e^{-5\lambda}$

$$P[v.r. > 5] = e^{-5\lambda} = e^{-\frac{1}{2}} \cong 0.6$$

5. CHI-QUADRO: χ^2

E' una distribuzione Γ (gamma) con $\lambda = \frac{1}{2}$ ed $\alpha = \frac{k}{2}$ dove k sono i gradi di libertà.

$$f(x) = \chi^2(x; k) = \begin{cases} \dfrac{1}{2^{\frac{k}{2}}\,\Gamma\left(\frac{k}{2}\right)} \, x^{\frac{k}{2}-1} e^{-\frac{x}{2}} & , x \geq 0 \\ 0 & \text{altrove} \end{cases}$$

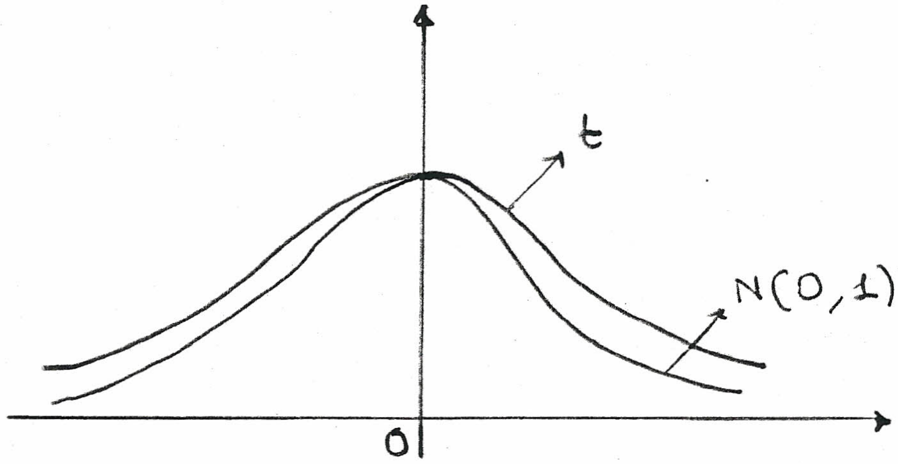
Si prova

$E[X] = \boxed{k}$; $\quad var[X] = \boxed{2k}$.

6 STUDENT t

$$t(x, n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\sqrt{n\pi}} \cdot \left(1 + \frac{x^2}{n}\right)^{-\frac{(n+1)}{2}}$$

n = gradi di libertà



Ha lo stesso tipo di grafico della normale standard, ma con code più pesanti.

DISTRIBUZIONI CONGIUNTE

Nello studio di molti esperimenti casuali può essere coinvolta più di una variabile casuale; perciò estenderemo le definizioni di funzione di ripartizione e di funzione di densità di una v.a. a più variabili casuali.

<u>DEF.</u>: Dato $(\Omega, \mathcal{A}, P)$ s. di p., chiamiamo VARIABILE ALEATORIA (O CASUALE) n-DIMENSIONALE una funzione $X = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ tale che $\forall (x_1, \dots, x_n) \in \mathbb{R}^n$ si ha:

$$\{\omega \in \Omega : X_1(\omega) \leq x_1, \dots, X_n(\omega) \leq x_n\} \in \mathcal{A}$$

Perciò una v.a. n-dim. è un complesso di n v.a. che associano un numero ad un risultato.

<u>DEF</u>: Data $X = (X_1, \dots, X_n)$ definita su $(\Omega, \mathcal{A}, P)$ chiamiamo FUNZIONE DI RIPARTIZIONE CONGIUNTA di $X_1, \dots, X_n$ la funzione $F_{X_1, \dots, X_n} : \mathbb{R}^n \to [0,1]$:
$\forall x = (x_1, \dots, x_n)$

$$F_{X_1, \dots, X_n}(x) = P[X_1 \leq x_1, \dots, X_n \leq x_n]$$

CASO DISCRETO

DEF: La v. a. $X=(X_1,\dots,X_n)$ è detta v.a. discreta n-dim se può assumere valori solo in un insieme numerabile $(x_1,\dots,x_n)$ di punti di $\mathbb{R}^n$.

DEF: Dato $(\Omega, \mathcal{A}, P)$ s. di p. e $X=(X_1,\dots,X_n)$ v.a. discreta n-dim, si definisce FUNZIONE DI DENSITÀ DISCRETA CONGIUNTA di $X_1,\dots,X_n$ la funzione

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n)=P[X_1=x_1,\dots X_n=x_n]$$

PROPRIETÀ

$$\sum' f_{X_1,\dots,X_n}(x_1,\dots,x_n)=1$$

Limitandoci al caso $n=2$ elenchiamo le proprietà della funzione di ripartizione congiunta $F_{X,Y}$ relativa alle v.a. congiunte X, Y.

PROPRIETÀ DI $F_{X,Y}$

1) $\lim\limits_{x\to-\infty} F(x,y)=0 \quad \forall y$

$\lim\limits_{y\to-\infty} F(x,y)=0 \quad \forall x$

2) $\lim\limits_{\substack{x\to+\infty\\ y\to+\infty}} F(x,y)=1$

3) se $x_1 < x_2$ e $y_1 < y_2$ allora

$$P[x_1 < X \leq x_2 ; y_1 < Y \leq y_2] =$$

$$F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$$

4) $F(x,y)$ è continua da destra in ciascuna variabile:

$$\lim_{h \to 0^+} F(x+h, y) = \lim_{h \to 0^+} F(x, y+h) = F(x,y).$$

DEF: Date X, Y v.a. congiunte discrete si ha:

$$F_{X,Y}(\bar{x}, \bar{y}) = \sum_{\substack{x_i \leq \bar{x} \\ y_i \leq \bar{y}}} f_{X,Y}(x_i, y_i)$$

DEF: Data $F_{X,Y}$ funzione di ripartizione congiunta di X, Y definiamo FUNZIONI DI RIPARTIZIONE MARGINALI le seguenti funzioni:

$$F_X(x) = F_{X,Y}(x, \infty) \qquad F_Y(y) = F_{X,Y}(\infty, y)$$

N.B. Tale def. vale anche nel caso continuo.

DEF: Date X, Y v.a. discrete congiunte, le funzioni $f_X(\cdot)$ e $f_Y(\cdot)$ sono dette FUNZIONI DI DENSITÀ MARGINALI

Si ricavano dalla f. di densità congiunta $f_{X,Y}(x,y)$ così:

$$f_X(x_k) = \sum_J f_{X,Y}(x_k, y_J)$$

$$f_Y(y_k) = \sum_J f_{X,Y}(x_J, y_k)$$

N.B. Dalla f. di densità congiunta è sempre possibile ricavare le f. di densità marginali ma non vale il viceversa.

<u>DEF</u>: Date X, Y v.a. congiunte con f. di densità congiunta $f_{X,Y}$, la funzione

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

se $f_Y(y)>0$ è detto FUNZIONE DI DENSITÀ CONDIZIONATA di X dato $Y=y$ e viene indicato con $f_{X,Y}(\cdot|y)$.

Analogamente si definisce

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

se $f_X(x)>0$.

N.B. Questa def. vale anche nel caso continuo.

<u>DEF</u>: Date X, Y v.a. congiunte si definisce FUNZIONE DI RIPARTIZIONE CONDIZIONATA di Y dato $X=x$ la funzione

$$F_{Y|X}(y|x) = P[Y \leq y \mid X=x] \quad \text{per } f_X(x)>0$$

$$= \sum_{\{j:\, y_j \leq y\}} f_{Y|X}(y_j|x)$$

che viene indicata con $F_{Y|X}(\cdot|x)$.

Esempio Lancio di 2 tetraedri (poliedri regolari a 4 facce) aventi facce numerate da 1 a 4.

$X =$ n° sulla faccia rivolta verso il basso del 1° t.

$Y =$ n° più grande tra quelli indicati sulle facce rivolte verso il basso.

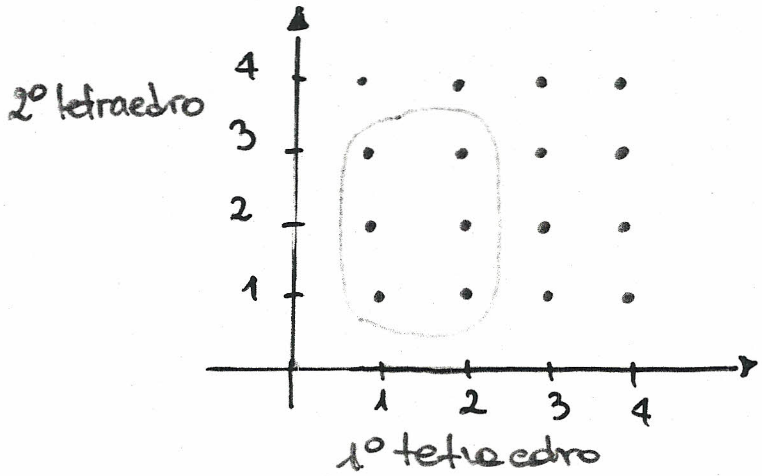
valori di X e Y congiunti

$(1,1), (1,2), (1,3), (1,4)$

$(2,2), (2,3), (2,4)$

$(3,3), (3,4)$

$(4,4)$



$N(\Omega) = 16$

vale l'equiprobabilità

P. es. $(x,y)=(2,3)$ voglio calcolare $F_{X,Y}(2,3)$:

$$F_{X,Y}(2,3) = P[X \leq 2 ; Y \leq 3] = \frac{6}{16}$$

Tabella delle $f_{X,Y}(x,y)$:

(x,y)	$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(2,2)$	$(2,3)$	$(2,4)$	$(3,3)$	$(3,4)$	$(4,4)$
$f_{X,Y}(x,y)$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{4}{16}$

"COSTRUZIONE"

1° T , 2° T

- $1 \quad 2 \quad \Rightarrow (1,2)$
- $\begin{cases} 2 & 1 \\ 2 & 2 \end{cases} \Rightarrow (2,2)$
- $2 \quad 3 \quad \Rightarrow (2,3)$
- $\begin{cases} 3 & 1 \\ 3 & 2 \\ 3 & 3 \end{cases} \Rightarrow (3,3)$

In quanti modi ottengo $(2,2)$? $f_{X,Y} = \frac{2}{16}$

In quanti modi ottengo $(3,3)$? $f_{X,Y} = \frac{3}{16}$

Ricordando che $F_{X,Y}(\bar{x},\bar{y}) = \sum_{\substack{x_i \le \bar{x} \\ y_i \le \bar{y}}} f_{X,Y}(x_i, y_i)$ posso calcolare p. es:

$$F(2,3) = \sum_{\substack{x_i \le 2 \\ y_i \le 3}} f_{X,Y}(x_i, y_i) = f(1,1) + f(1,2) + f(1,3) + f(2,2) + f(2,3) = \frac{6}{16}$$

$$F(4,3) = \sum_{\substack{x_i \le 4 \\ y_i \le 3}} f_{X,Y}(x_i, y_i) = f(1,1) + f(1,2) + f(1,3) + f(2,2) + f(2,3) + f(3,3) = \frac{9}{16}$$

In tal modo si ottiene, $\forall$ pto $\in \Omega$, la $F_{X,Y}(x,y)$, cioè:

$$\begin{cases} F(1,1) = f(1,1) = 1/16 \\ F(1,2) = f(1,1) + f(1,2) = 2/16 \\ F(1,3) = f(1,1) + f(1,2) + f(1,3) = 3/16 \\ F(1,4) = 4/16 \end{cases}$$

$$\begin{cases} F(2,1) = f(1,1) = 1/16 \\ F(2,2) = f(1,1) + f(1,2) + f(2,2) = 4/16 \\ F(2,3) = 6/16 \\ F(2,4) = 8/16 \end{cases}$$

$$\begin{cases} F(3,1) = f(1,1) = 1/16 \\ F(3,2) = f(1,1) + f(1,2) + f(2,2) = 4/16 \\ F(3,3) = 9/16 \\ F(3,4) = 12/16 \end{cases}$$

$$\begin{cases} F(4,1) = 1/16 \\ F(4,2) = 4/16 \\ F(4,3) = 9/16 \\ F(4,4) = 1 \end{cases}$$

Per il calcolo delle densità marginali:

$$
\begin{cases}
f_Y(1) = f_{X,Y}(1,\underline{1}) = \frac{1}{16} \\
f_Y(2) = f_{X,Y}(1,\underline{2}) + f_{X,Y}(2,\underline{2}) = \frac{3}{16} \\
f_Y(3) = f_{X,Y}(1,\underline{3}) + f_{X,Y}(2,\underline{3}) + f_{X,Y}(3,\underline{3}) = \frac{5}{16} \\
f_Y(4) = f_{X,Y}(1,\underline{4}) + f_{X,Y}(2,\underline{4}) + f_{X,Y}(3,\underline{4}) + f_{X,Y}(4,\underline{4}) = \frac{7}{16}
\end{cases}
$$

$$
\begin{cases}
f_X(1) = f_{X,Y}(\underline{1},1) + f_{X,Y}(\underline{1},2) + f_{X,Y}(\underline{1},3) + f_{X,Y}(\underline{1},4) = \frac{4}{16} \\
f_X(2) = f_{X,Y}(\underline{2},2) + f_{X,Y}(\underline{2},3) + f_{X,Y}(\underline{2},4) = \frac{4}{16} \\
f_X(3) = f_{X,Y}(\underline{3},3) + f_{X,Y}(\underline{3},4) = \frac{4}{16} \\
f_X(4) = f_{X,Y}(\underline{4},4) = \frac{4}{16}
\end{cases}
$$

Osservazione: le densità marginali si calcolano più semplicemente sommando per righe o per colonne i valori della tabella (di $f_{X,Y}$) a doppia entrata (vedi lucido successivo).

Tabella dei valori di $F_{X,Y}$

$4 \le y$	0	4/16	8/16	12/16	1
$3 \le y < 4$	0	3/16	6/16	9/16	9/16
$2 \le y < 3$	0	2/16	4/16	4/16	4/16
$1 \le y < 2$	0	1/16	1/16	1/16	1/16
$y < 1$	0	0	0	0	0
	$x < 1$	$1 \le x < 2$	$2 \le x < 3$	$3 \le x < 4$	$4 \le x$

Tabella dei valori di $f_{X,Y}$

4	1/16	1/16	1/16	4/16	$\rightarrow f_Y(4)$
3	1/16	1/16	3/16		$\rightarrow f_Y(3)$
2	1/16	2/16			$\rightarrow f_Y(2)$
1	1/16				$\rightarrow f_Y(1)$
y/x	1	2	3	4	
	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	
	$f_X(1)$	$f_X(2)$	$f_X(3)$	$f_X(4)$	

INDIPENDENZA

DEF: Dato $X=(X_1,..,X_m)$ v.a. m-dim (discreta o continua con funzione di densità congiunta $f_{X_1,..,X_m}$, diciamo che $X_1,...,X_m$ sono v.a. INDIPENDENTI sse

$$f_{X_1,..,X_m}(x_1,..,x_m)=f_{X_1}(x_1)\cdots f_{X_m}(x_m)$$

per tutti gli $x_1,..,x_m$.

Cioè la f. di densità congiunta si scrive come PRODOTTO delle f. di densità marginali.

Nel caso 2-D si ha:

X, Y indipendenti sse $f_{X,Y}(x,y)=f_X(x)\cdot f_Y(y)$.

Dalla definizione di densità condizionata si ha:

$f_{X,Y}(x,y)=f_{Y|X}(y|x)\,f_X(x)$

perciò $f_{Y|X}(y|x)=f_Y(y)$

cioè la densità condizionata di Y dato $X=x$ è la densità non condizionata di Y.

Per mostrare che due v.a. non sono indipendenti basta ~~di~~mostrare che $f_{Y|X}(y|x)$ dipende da x.

OSSERVAZIONE

Si può provare che se $X_1,..,X_m$ sono v.a. indipendenti e $g_1,..,g_m$ sono m funzioni tali che $Y_k=g_k(X_k)$ $\forall\ k=1,..,m$ siano v.a. allora $Y_1,..,Y_m$ sono indipendenti.

- Qual è la densità di Y dato X = 2?

$$f_{Y|X}(2|2) = \frac{f_{X,Y}(2,2)}{f_X(2)} = \frac{2/16}{4/16} = \frac{1}{2}$$

$$f_{Y|X}(3|2) = \frac{f_{X,Y}(2,3)}{f_X(2)} = \frac{1/16}{4/16} = \frac{1}{4}$$

$$f_{Y|X}(4|2) = \frac{f_{X,Y}(2,4)}{f_X(2)} = \frac{1/16}{4/16} = \frac{1}{4}.$$

- Qual è la densità di Y dato X = 3?

$$f_{Y|X}(y|3) = \begin{cases} \frac{3}{4} & \text{per } y = 3 \\ \frac{1}{4} & \text{per } y = 4. \end{cases}$$

- Le v.a. X, Y sono indipendenti? OVVIAMENTE NO!

$$f_{Y|X}(2|3) = P[Y=2 \mid X=3] \equiv 0 \quad (Y > X)$$

$$f_Y(2) = P[Y=2] = \frac{3}{16}$$

$$f_{Y|X}(y|x) \neq f_Y(y)$$

Esempio

Da un gruppo di 12 batterie (3 nuove, 4 usate, 5 difettose) ne vengono scelte 3 a caso. Indicato con

X = n° batterie nuove

Y = n° batterie usate

tra quelle scelte, determinare la $f_{X,Y}(x,y)$.

Possibili risultati:

NNN $\rightarrow f(3,0)$

NNU $\rightarrow f(2,1)$

NND $\rightarrow f(2,0)$

NUU $\rightarrow f(1,2)$

NDD $\rightarrow f(1,0)$

NUD $\rightarrow f(1,1)$

UUU $\rightarrow f(0,3)$

UUD $\rightarrow f(0,2)$

UDD $\rightarrow f(0,1)$

DDD $\rightarrow f(0,0)$

(non conta l'ordine).

$$f(3,0) = \frac{\binom{3}{3}}{\binom{12}{3}} = \dots = \frac{1}{220} \qquad f(2,1) = \frac{\binom{3}{2}\binom{4}{1}}{\binom{12}{3}} = \dots = \frac{12}{220}$$

$$f(2,0) = \frac{\binom{3}{2}\binom{5}{1}}{\binom{12}{3}} = \dots = \frac{15}{220} \qquad f(1,2) = \frac{\binom{3}{1}\binom{4}{2}}{\binom{12}{3}} = \dots = \frac{18}{220}$$

$$f(1,0) = \frac{\binom{3}{1}\binom{5}{2}}{\binom{12}{3}} \dots = \frac{30}{220} \qquad f(1,1) = \frac{\binom{3}{1}\binom{4}{1}\binom{5}{1}}{\binom{12}{3}} = \dots = \frac{60}{220}$$

$$f(0,3) = \frac{\binom{4}{3}}{\binom{12}{3}} = \frac{4}{220} \qquad f(0,2) = \frac{\binom{4}{2}\binom{5}{1}}{\binom{12}{3}} = \frac{30}{220}$$

$$f(0,1) = \frac{\binom{4}{1}\binom{5}{2}}{\binom{12}{3}} = \frac{40}{220} \qquad f(0,0) = \frac{\binom{5}{3}}{\binom{12}{3}} = \frac{10}{220}$$

Tabella di $f_{X,Y}$.

	$Y=0$	$Y=1$	$Y=2$	$Y=3$	$f_X(x)$
$X=0$	$1/22$	$2/11$	$3/22$	$1/55$	$41/110$
$X=1$	$3/22$	$3/11$	$9/110$	0	$27/55$
$X=2$	$1/15$	$3/55$	0	0	$27/220$
$X=3$	$1/220$	0	0	0	$1/220$
$f_Y(y)$	$14/55$	$28/55$	$12/55$	$1/55$	1

MODELLO DI V.A. N-DIM (caso discreto)

DISTRIBUZIONE MULTINOMIALE

Tale distribuzione è associata a prove ripetute e indipendenti che generalizzano il caso delle prove di Bernoulli a 2 esiti a quello con più di due esiti.

Supponiamo che esistano $k+1$ esiti possibili distinti di un tentativo. Siano $s_1, \dots s_{k+1}$ tali esiti.

Sia $p_i = P[s_i]$ $\quad i = 1, \dots, k+1$ con

$$\sum_{i=1}^{k+1} p_i = 1 \quad \Rightarrow \quad p_{k+1} = 1 - (p_1 + \dots + p_k)$$

Ripetiamo la prova n volte.

Sia X_i il n° di volte che si ottiene s_i sugli n tentativi, per $i = 1, \dots, k+1$.

Se le prove sono ripetute e indipendenti si ha:

$$f_{X_1, \dots, X_k}(x_1, \dots, x_k) = \frac{n!}{x_1! \cdots x_{k+1}!} \, p_1^{x_1} \cdots p_{k+1}^{x_{k+1}}$$

dove $\displaystyle\sum_{i=1}^{k+1} x_i = n$ e $\displaystyle X_{k+1} = n - \sum_{i=1}^{k} X_i$.

Pertanto cerco la probabilità che da n prove si abbiano esattamente x_1 esiti del tipo s_1, x_2 esiti del tipo s_2, ... x_{k+1} esiti del tipo s_{k+1}.

Un particolare ordinamento ha probabilità

$$p_1^{x_1} \cdots p_{k+1}^{x_{k+1}}$$

e di ordinamenti ne esistono:

$$\frac{n!}{x_1! \cdots x_{k+1}!}$$

<u>Esempio</u>

Su un quantitativo di merce, il 10% viene pagato in ritardo, il 30% viene restituito.

Vengono effettuati 20 ordini.

Calcolare la probabilità che 3 ordini vengano pagati in ritardo e 5 vengano restituiti (su un totale di 20).

<u>IPOTESI</u>: indipendenza

$n = 20$

$$x_1 = 3, \quad x_2 = 5 \quad x_3 = n - x_1 - x_2 = 20 - 3 - 5 = 12$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$p_1 = 0.1 \quad p_2 = 0.3 \quad p_3 = 0.6 = 1 - p_1 - p_2$$

il valore cercato è:

$$P = \frac{20!}{3!\,5!\,12!}(0.1)^3 \cdot (0.3)^5 \cdot (0.6)^{12} \simeq 0.037.$$

VALORE ATTESO

Estendiamo ora il concetto di valore atteso di una v.a. al caso di più variabili.

CASO DISCRETO

DEF: Data $X=(X_1, \dots X_n)$ v.a. n-dim con densità $f_{X_1,\dots,X_n}$ e data una funzione $g: \mathbb{R}^n \to \mathbb{R}$, definiamo VALORE ATTESO di g:

$$E[g(X_1,\dots,X_n)] = \sum g(x_1,\dots,x_n) f_{X_1,\dots,X_n}(x_1,\dots,x_n)$$

CASI PARTICOLARI

- $g(X_1,\dots,X_n) = X_i$

  $E[g] = E[X_i] = \mu_{X_i}$

- $g(X_1,\dots,X_n) = (X_i - \mu_{X_i})^2$

  $E[g] = E[(X_i - \mu_{X_i})^2] = var[X_i] = \sigma^2_{X_i}$

CASO 2-D

DEF: Date X, Y v.a. definite su $(\Omega, \mathcal{A}, P)$, definiamo COVARIANZA di X e Y la quantità:

$$\sigma_{X,Y} = cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)]$$

definiamo COEFFICIENTE DI CORRELAZIONE la quantità:

$$\rho_{X,Y} = \rho[X,Y] = \frac{cov[X,Y]}{\sigma_X \cdot \sigma_Y} \quad , \sigma_X > 0 \; , \sigma_Y > 0$$

La covarianza e il coefficiente di correlazione sono misure di una relazione lineare tra X e Y.

PROPRIETÀ : $|\rho_{X,Y}| \leq 1$

Nel caso $Y = mX + q$ (eq. della retta $\Rightarrow$ dipend. lineare)

$E[Y] = E[mX + q] = m\,E[X] + q$ cioè:

$\mu_Y = m\,\mu_X + q$

$var[Y] = var[mX + q] = var[mX] = m^2 var[X]$ cioè:

$\sigma_Y^2 = m^2 \sigma_X^2 \quad \Rightarrow \quad \sigma_Y = m\,\sigma_X > 0$

allora

$$cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = E[(mX+q-m\mu_X-q)(X-\mu_X)]$$
$$= E[m(X-\mu_X)^2] = m\,var[X] = m\,\sigma_X^2 \gtrless 0$$

dipende da $m \gtrless 0$

Pertanto

$$|\rho_{X,Y}| = \frac{|m|\sigma_X^2}{m\,\sigma_X \cdot \sigma_X} = 1$$

N.B.: le def. di $\sigma_{X,Y}$ e $\rho_{X,Y}$ valgono anche nel caso continuo.

PROPRIETÀ DELLA COVARIANZA

1) $cov[X,Y] = E[XY] - \mu_X \mu_Y$

Dalla definizione

$$cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = E[XY - \mu_Y X - \mu_X Y + \mu_X \mu_Y]$$

$$= E[XY] - \mu_Y \underbrace{E[X]}_{=\mu_X} - \cancel{\mu_X \underbrace{E[Y]}_{=\mu_Y}} + \cancel{\mu_X \mu_Y}$$

$$= E[XY] - \mu_X \mu_Y$$

2) $cov[aX, bY] = ab\, cov[X,Y]$, $a, b \in \mathbb{R}$

banale

3) $cov[X,Y] = cov[Y,X]$

banale

4) $cov[X,X] = var[X]$

$$cov[X,X] = E[(X-\mu_X)^2] = var[X]$$

5) $cov[X+Y, Z] = cov[X,Z] + cov[Y,Z]$

$$cov[X+Y,Z] \underset{\text{per } 1)}{=} E[(X+Y)Z] - E[(X+Y)]\,E[Z]$$

$$= E[XZ + YZ] - (E[X] + E[Y])\,E[Z]$$

$$= E[XZ] + E[YZ] - E[X]E[Z] - E[Y]E[Z]$$

$$= cov[X,Z] + cov[Y,Z]$$

Esempio

Riprendiamo l'esempio del lancio dei due tetraedri

- $g(X,Y) = XY$

$E[XY] = \sum' xy\, f_{X,Y}(x,y)$

$= 1\cdot 1\cdot \frac{1}{16} + 1\cdot 2\cdot \frac{1}{16} + 1\cdot 3\cdot \frac{1}{16} + 1\cdot 4\cdot \frac{1}{16} +$

$2\cdot 2\cdot \frac{2}{16} + 2\cdot 3\cdot \frac{1}{16} + 2\cdot 4\cdot \frac{1}{16} +$

$3\cdot 3\cdot \frac{3}{16} + 3\cdot 4\cdot \frac{1}{16} + 4\cdot 4\cdot \frac{4}{16} = \frac{135}{16}$

- $g(X,Y) = X$

$E[X] = \sum' x\, f_{X,Y}(x,y)$

- $g(X,Y) = Y$

$E[Y] = \sum' y\, f_{X,Y}(x,y)$

ma $\sum' x\, f_{X,Y}(x,y) = 1\cdot \sum_y f_{X,Y}(1,y) + 2 \sum'_y f_{X,Y}(2,y) +$

$3\cdot \sum'_y f_{X,Y}(3,y) + 4 \sum'_y f_{X,Y}(4,y)$

$= 1\cdot \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}\right) + 2\cdot\left(\frac{2}{16} + \frac{1}{16} + \frac{1}{16}\right) + 3\left(\frac{3}{16} + \frac{1}{16}\right)\cdot 4\cdot \frac{4}{16}$

$= \frac{40}{16} = \frac{5}{2}$

analogamente si ricava

$E[Y] = 1\cdot \sum_x f_X(x,1) + 2\cdot \sum_x f_X(x,2) + 3 \sum_x f_X(x,3) + 4 \sum_x f_X(x,4)$

$= 1\cdot \frac{1}{16} + 2\cdot\left(\frac{1}{16} + \frac{2}{16}\right) + 3\cdot\left(\frac{1}{16} + \frac{1}{16} + \frac{3}{16}\right) +$

$+ 4\left(\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{4}{16}\right) = \frac{50}{16} = \frac{25}{8}$

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

$E[X+Y] = E[X] + E[Y] = \frac{90}{16} = \frac{45}{8}$

OSSERVAZIONE

I valori $E[X]$ e $E[Y]$ si possono calcolare più velocemente attraverso le funzioni di densità marginali.

$$E[X] = \sum_i x f_X(x) = 1 \cdot f_X(1) + 2 \cdot f_X(2) + 3 \cdot f_X(3) + 4 \cdot f_X(4)$$

$$= (1+2+3+4) \cdot \frac{4}{16} = \frac{40}{16} = \frac{5}{2}$$

$$E[Y] = \sum_i y f_Y(y) = 1 \cdot f_Y(1) + 2 \cdot f_Y(2) + 3 \cdot f_Y(3) + 4 \cdot f_Y(4)$$

$$= 1 \cdot \frac{1}{16} + 2 \cdot \frac{3}{16} + 3 \cdot \frac{5}{16} + 4 \cdot \frac{7}{16} = \frac{50}{16} = \frac{25}{8}$$

Possiamo determinare

$$\text{cov}[X, Y] = E[XY] - E[X] \cdot E[Y]$$

$$= \frac{135}{16} - \frac{5}{2} \cdot \frac{25}{8} = \frac{10}{16} = \frac{5}{8}$$

Per determinare $\rho_{X,Y}$ dobbiamo prima calcolare σ_X, σ_Y.

$$E[X^2] = \sum_i x^2 f_X(x) = 1 \cdot f_X(1) + 2^2 \cdot f_X(2) + 3^2 \cdot f_X(3) + 4^2 \cdot f_X(4)$$

$$= \frac{4}{16} \cdot (1+4+9+16) = \frac{120}{16} = \frac{15}{2}$$

$$\text{var}[X] = E[X^2] - (E[X])^2 = \frac{15}{2} - \left(\frac{5}{2}\right)^2 = \frac{5}{4}$$

$$\Rightarrow \sigma_X = \frac{\sqrt{5}}{2}$$

$$E[Y^2] = \sum y^2 f_Y(y) = 1 \cdot f_Y(1) + 4 \cdot f_Y(2) + 9 \cdot f_Y(3) + 16 f_Y(4)$$

$$= 1 \cdot \frac{1}{16} + 4 \cdot \frac{3}{16} + 9 \cdot \frac{5}{16} + 16 \cdot \frac{7}{16} = \frac{170}{16} = \frac{85}{8}$$

$$\text{var}[Y] = E[Y^2] - (E[Y])^2 = \frac{85}{8} - \left(\frac{25}{8}\right)^2 = \frac{55}{64}$$

$$\Rightarrow \sigma_Y = \frac{\sqrt{55}}{8}$$

$$\Rightarrow \rho_{X,Y} = \frac{\frac{5}{8}}{\frac{\sqrt{5}}{2} \cdot \frac{\sqrt{55}}{8}} = \frac{2}{\sqrt{11}}$$

DEF: Date X, Y v.a. e data $g = g(X,Y)$. Definiamo VALORE ATTESO CONDIZIONATO di g dato $X = x$ la quantità:

$$E[g(X,Y) \mid X = x] = \sum_i g(x, y_i) f_{Y|X}(y_i \mid x)$$

In particolare se $g(X,Y) = Y$

$$E[g(X,Y) \mid X = x] = E[Y \mid X = x] = E[Y \mid x]$$

Nell'esempio del lancio dei due tetraedri ricadiamo che

$$f_{Y|X}(y \mid x=2) = \begin{cases} \frac{1}{2} & y = 2 \\ \frac{1}{4} & y = 3 \\ \frac{1}{4} & y = 4 \end{cases}$$

voglio calcolare $E[Y \mid X=2]$.

$$E[Y \mid X=2] = \sum_i y\, f_{Y|X}(y \mid x=2)$$

$$= 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = \frac{11}{4}$$

Esempio

Prodotti classificati a seconda dei loro difetti e della fabbrica che li ha prodotti.

X_2 = v.a. che indica la fabbrica ($X_2 = 1$, $X_2 = 2$)

X_1 = v.a. che indica il n° di difetti per pezzo
($X_1 = 0, 1, 2, 3$)

Data la seguente tabella della $f_{X,Y}$.

	$X_1=0$	$X_1=1$	$X_1=2$	$X_1=3$	f_{X_2}
$X_2=1$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{2}$
$X_2=2$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$
f_{X_1}	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{5}{16}$	$\frac{3}{8}$	1

determinare:

1) $f_{X_1}(x_1)$; $f_{X_2}(x_2)$

2) media e varianza di X_i $i=1,2$

3) $\operatorname{cov}[X_1, X_2] \Rightarrow \rho_{X_1,X_2}$

$$E[X_1] = \sum x_1 f_{X_1}(x_1) = 0\cdot\frac{3}{16} + 1\cdot\frac{1}{8} + 2\cdot\frac{5}{16} + 3\cdot\frac{3}{8} = \frac{15}{8}$$

$$E[X_2] = \sum x_2 f_{X_2}(x_2) = 1\cdot\frac{1}{2} + 2\cdot\frac{1}{2} = \frac{3}{2}$$

$$E[X_1^2] = \sum x_1^2 f_{X_1}(x_1) = 0\cdot\frac{3}{16} + 1\cdot\frac{1}{8} + 4\cdot\frac{5}{16} + 9\cdot\frac{3}{8} = \frac{19}{4}$$

$$\Rightarrow \operatorname{var}[X_1] = \frac{19}{4} - \frac{225}{64} = \frac{79}{64}$$

$$E[X_2^2] = \sum x_2^2 f_{X_2}(x_2) = 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{2} = \frac{5}{2}$$

$$\Rightarrow var[X_2] = \frac{5}{2} - \frac{9}{4} = \frac{1}{4}$$

$$cov[X_1, X_2] = E[X_1 X_2] - E[X_1] \cdot E[X_2]$$

$$E[X_1 X_2] = \sum x_1 x_2 f_{X_1, X_2}(x_1, x_2)$$

$$= 0 \cdot 1 \cdot \frac{1}{8} + 1 \cdot 1 \cdot \frac{1}{16} + 2 \cdot 1 \cdot \frac{3}{16} + 3 \cdot 1 \cdot \frac{1}{8} +$$

$$0 \cdot 2 \cdot \frac{1}{16} + 1 \cdot 2 \cdot \frac{1}{16} + 2 \cdot 2 \cdot \frac{1}{8} + 3 \cdot 2 \cdot \frac{1}{4}$$

$$= \frac{47}{16}$$

$$\Rightarrow cov[X_1, X_2] = \frac{47}{16} - \frac{15}{8} \cdot \frac{3}{2} = \frac{1}{8}.$$

$$\rho = \frac{cov[X_1, X_2]}{\sigma_{X_1} \sigma_{X_2}} = \frac{\frac{1}{8}}{\frac{\sqrt{79}}{8} \cdot \frac{1}{2}} = 2/\sqrt{79}$$

<u>Esempio</u> : scritto del 17.07.07

URNA ← 12 palline numerate $\begin{cases} 2 \text{ hanno } ① \\ 4 \quad " \quad ② \\ 2 \quad " \quad ③ \\ 4 \quad " \quad ④ \end{cases}$

Estratta una pallina

X = n° inciso sulla pallina estratta

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

Determinare

1) $f_{X,Y}$; f_X; f_Y

2) indipendenza

3) $\operatorname{cov}[X,Y]$

4) $P[X>2 \mid Y=\frac{1}{2}]$

$P(1) = \frac{2}{12} = \frac{1}{6}$; $P(2) = \frac{4}{12} = \frac{1}{3}$; $P(3) = \frac{1}{6}$; $P(4) = \frac{1}{3}$

$X=1 \Rightarrow Y = \frac{1}{2}$

$X=2 \Rightarrow Y=0$ $\qquad \Rightarrow X = \{1, 2, 3, 4\}$

$X=3 \Rightarrow Y = \frac{1}{2}$ $\qquad Y = \{0, \frac{1}{2}, 2\}$

$X=4 \Rightarrow Y=2$

Y \ X	X=1	X=2	X=3	X=4	f_Y
Y=0	0	$\frac{1}{3}$	0	0	$\frac{1}{3}$
$Y=\frac{1}{2}$	$\frac{1}{6}$	0	$\frac{1}{6}$	0	$\frac{1}{3}$
Y=2	0	0	0	$\frac{1}{3}$	$\frac{1}{3}$
f_X	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	1

- $f_{X,Y} = f_X \cdot f_Y$?

  $f_{X,Y}(1,0) = 0$ ma $f_X(1) = \frac{1}{6}$ e $f_Y(0) = \frac{1}{3}$ $\Rightarrow$ NO.

- $\mathrm{cov}[X,Y] = E[XY] - \mu_X \mu_Y$

  $\mu_X = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{3} = \frac{8}{3}$

  $\mu_Y = 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} = \frac{5}{6}$

  $E[XY] = \cancel{2 \cdot 0 \cdot \frac{1}{3}} + 1 \cdot \frac{1}{2} \cdot \frac{1}{6} + 3 \cdot \frac{1}{2} \cdot \frac{1}{6} + 4 \cdot 2 \cdot \frac{1}{3} = 3$

  $\Rightarrow \mathrm{cov}[X,Y] = \frac{7}{9}$

- $P[X > 2 \mid Y = \frac{1}{2}] = P[X = 3 \mid Y = \frac{1}{2}] + P[X = 4 \mid Y = \frac{1}{2}]$

  $$= \frac{f_{X,Y}(3, \frac{1}{2})}{f_Y(\frac{1}{2})} + \frac{f_{X,Y}(4, \frac{1}{2})}{f_Y(\frac{1}{2})}$$

  $$= \frac{\frac{1}{6}}{\frac{1}{3}} + \frac{0}{\frac{1}{3}} = \frac{1}{2}$$

TEOREMA

Siano X, Y v.a. indipendenti e g_1, g_2 funzioni tali che $g_1 = g_1(X)$, $g_2 = g_2(Y)$ siano v.a.

$$\Rightarrow \quad E[g_1 \cdot g_2] = E[g_1] \cdot E[g_2]$$

Dim:

$$E[g_1(X) \cdot g_2(Y)] = \sum_i \sum_J g_1(x_i)\, g_2(y_J)\, f_{X,Y}(x_i, y_J)$$

$$= \sum_i \sum_J g_1(x_i)\, g_2(y_J)\, f_X(x_i) \cdot f_Y(y_J)$$

$$= \sum_i g_1(x_i) f_X(x_i) \cdot \sum_J g_2(y_J) f_Y(y_J)$$

$$= E[g_1] \cdot E[g_2]$$

COROLLARIO

Se X, Y v.a. indipendenti $\Rightarrow \mathrm{cov}[X, Y] \equiv 0$.

Dim:

Scelto $g_1(X) = X - \mu_X$

$g_2(Y) = Y - \mu_Y$

$$\mathrm{cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = E[g_1 \cdot g_2] \overset{\text{TEOREMA}}{=} E[g_1] \cdot E[g_2]$$

$$= E[(X-\mu_X)] \cdot E[(Y-\mu_Y)]$$

$$= \{E[X] - \mu_X\} \cdot \{E[Y] - \mu_Y\}$$

$$= (\mu_X - \mu_X) \cdot (\mu_Y - \mu_Y) \equiv 0.$$

OSSERVAZIONE

Date X, Y v.a. indipendenti, se $g_1(X) = X$ e $g_2(Y) = Y$, dal teorema si ha:

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

$\Rightarrow$

$$\operatorname{cov}[X, Y] = E[XY] - E[X] \cdot E[Y] \equiv 0.$$

N.B.: Il viceversa non è vero

cioè:

$$\operatorname{cov}[X, Y] = 0 \not\Rightarrow X, Y \text{ v.a. indipendenti}$$

cioè se $\operatorname{cov}[X, Y] = 0$ non vale che $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$

DEF: X, Y. v.a. si dicono NON CORRELATE SE E SOLO SE

$$\operatorname{cov}[X, Y] = 0.$$

SOMMA DI VARIABILI CASUALI

Date n v.a. $X_1, \dots, X_n$, si consideri la funzione:

$$g(X_1, \dots, X_n) = X_1 + \dots + X_n$$

si ha che

1) $$E[g] = \sum_{i=1}^{n} E[X_i]$$

2) $$\mathrm{var}[g] = \sum_{i=1}^{n} \mathrm{var}[X_i] + 2\sum_i \sum_{j,\, i \neq j} \mathrm{cov}[X_i, X_j]$$

La 1) è ovvia. Dimostriamo la 2):

$$\mathrm{var}[g] = \mathrm{var}\Big[\sum_{i=1}^{n} X_i\Big] = E\Big[\Big(\sum_{i=1}^{n} X_i - E\Big[\sum_{i=1}^{n} X_i\Big]\Big)^2\Big]$$

$$\overset{\text{per 1)}}{=} E\Big[\Big(\sum_{i=1}^{n} (X_i - E[X_i])\Big)^2\Big]$$

$$= E\Big[\sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - E[X_i])(X_j - E[X_j])\Big]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} E\Big[(X_i - E[X_i])(X_j - E[X_j])\Big]$$

si spezza nella somma in cui $i = j$ perciò

$$\sum_{i=1}^{n} E\Big[(X_i - E[X_i])^2\Big] = \sum_{i=1}^{n} \mathrm{var}[X_i]$$

e nella somma in cui $i \neq j$ perciò

$$2\sum_i \sum_{j,\, i \neq j} E\Big[(X_i - E[X_i])(X_j - E[X_j])\Big) =$$

$$= 2\sum_i \sum_{j,\, i \neq j} \mathrm{cov}[X_i, X_j]$$

- $g = X + Y$

$$var[g] = E[(X+Y)^2] - (E[X+Y])^2$$

$$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$$

$$= E[X^2] + 2E[XY] + E[Y^2] - \{(E[X])^2 + 2E[X]E[Y] + (E[Y])^2\}$$

$$= var[X] + var[Y] + 2[E[XY] - E[X]E[Y]]$$

$$= var[X] + var[Y] + 2\,cov(X,Y)$$

- $g = X + Y + Z$

$var[g]$

$$\cancel{E[g]} = \cancel{E[X+Y+Z]} = E[(X+Y+Z)^2] - (E[X+Y+Z])^2$$

$$= var[X] + var[Y] + var[Z] + 2\,cov(X,Y) +$$

$$+ 2\,cov(X,Z) + 2\,cov(Y,Z)$$

<u>Quanti sono i termini del tipo</u>

$$\sum_i \sum_{\substack{j \\ i \neq j}} cov(x_i, x_j) \quad ?$$

se

$n=2$ $\quad cov(x_1, x_2) \rightarrow 1$

$n=3$ $\quad cov(x_1, x_2),\ cov(x_1, x_3),\ cov(x_2, x_3) \rightarrow 3$

$n=4$ $\quad cov(x_1, x_2),\ cov(x_1, x_3),\ cov(x_1, x_4)$

$\quad cov(x_2, x_3),\ cov(x_2, x_4),\ cov(x_3, x_4) \rightarrow 6$

$\vdots$

$n=k$ $\quad \dots\dots \Rightarrow \dfrac{k(k-1)}{2}$

$\Rightarrow$ se $g = x_1 + \dots + x_n$ i termini sono $\dfrac{n(n-1)}{2}$

in generale

3) $E[a_1 X_1 + \dots + a_n X_n] = a_1 E[X_1] + \dots + a_n E[X_n]$

cioè

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

4) $\operatorname{var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \operatorname{var}[X_i] + 2 \sum\sum_{i \neq j} a_i a_j \operatorname{cov}[X_i, X_j]$

per $n = 2$

$\operatorname{var}[X \pm Y] = \operatorname{var}[X] + \operatorname{var}[Y] \pm 2 \operatorname{cov}[X, Y]$

- Se $X_1, \dots, X_n$ sono v.a. non correlate $\Rightarrow$

5) $\operatorname{var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \operatorname{var}[X_i]$ (ipotesi più forte indipendenza)

MEDIA CAMPIONARIA

Date n v.a. $X_1, \dots, X_n$ a 2 a 2 non correlate (ipotesi più forte: indipendenti) e identicamente distribuite con media μ e varianza σ^2, si consideri la funzione:

- $g(X_1, \dots, X_n) = \bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$

chiamata MEDIA CAMPIONARIA.

si ha:

1) $E[\bar{X}_n] = \mu$

2) $\operatorname{var}[\bar{X}_n] = \frac{\sigma^2}{n}$

Infatti

$$E[\bar{X}_n] = E\left[\frac{1}{n}(X_1 + \dots + X_n)\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] =$$

$$= \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \cdot (n\mu) = \mu$$

Posto $a_i = \frac{1}{n}$ $\quad i = 1, \dots, n$

$$\mathrm{var}[\bar{X}_n] = \mathrm{var}\left[\frac{1}{n}(X_1 + \dots + X_n)\right] = \mathrm{var}\left[\frac{1}{n} X_1 + \dots + \frac{1}{n} X_n\right]$$

$$\overset{\text{per (4)}}{=} \frac{1}{n^2} \mathrm{var}[X_1] + \dots + \frac{1}{n^2} \mathrm{var}[X_n] + 2 \cdot \frac{1}{n^2} \sum_i \sum_{j,\, i \neq j} \underbrace{\mathrm{cov}[X_i, X_j]}_{= 0 \text{ non correlate}}$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{var}[X_i] = \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n}$$

Se considero invece la funzione:

- $g(X_1, \dots, X_n) = Y = X_1 + \dots + X_n$ $\quad$ i. i. d. oppure i. d. e non correlate.

si ha

1) $E[Y] = n\mu$

2) $\mathrm{var}[Y] = n\sigma^2$

Esempio

Determinare la varianza del n° di teste su 10 lanci indipendenti di una moneta non truccata.

$$I_J = \begin{cases} 1 & \text{se il } J\text{-esimo lancio è testa} \\ 0 & \text{se il } J\text{-esimo lancio è croce} \end{cases}$$

$$\text{n° totale delle teste} = \sum_{J=1}^{10} I_J$$

I lanci sono indipendenti $\Rightarrow \mathrm{var}[\sum \cdots] = \sum \mathrm{var}[\cdots]$

$$\mathrm{var}\left[\sum_{J=1}^{10} I_J\right] = \sum_{J=1}^{10} \mathrm{var}[I_J] = \mathrm{var}[I_1] + \cdots + \mathrm{var}[I_m]$$

$$= 10\, \mathrm{var}[I]$$

$$\mathrm{var}[I] = E[I^2] - (E[I])^2$$

Ma $\underline{I^2 = I}$

$$\Rightarrow \mathrm{var}[I] = E[I] - (E[I])^2 = E[I](1 - E[I])$$

Ora

$$E[I] = 1 \cdot \underset{\frac{1}{2}}{\underset{\parallel}{P[I=1]}} + 0 \cdot \underset{\frac{1}{2}}{\underset{\parallel}{P[I=0]}} = \frac{1}{2}$$

$$\Rightarrow \mathrm{var}[I] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\Rightarrow \mathrm{var}\left[\sum_{J=1}^{10} I_J\right] = \frac{5}{2}.$$

Date n variabili casuali $X_1, \dots, X_n$ tali che

$$\forall i, j \quad i \neq j \quad \text{cov}[X_i, X_j] = \alpha,$$

verificare che

$$\lim_{n \to \infty} \text{var}\left[\frac{X_1 + \dots + X_n}{n}\right] = \alpha$$

$$\text{var}\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{1}{n^2} \text{var}[X_1 + \dots + X_n]$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} \text{var}[X_i] + 2 \sum_{i,j}' \text{cov}[X_i, X_j]\right]$$

questi termini sono $\frac{n(n-1)}{2}$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} \text{var}[X_i] + \cancel{2}\, \frac{n(n-1)}{\cancel{2}} \cdot \alpha\right]$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \text{var}[X_i] + \frac{(n-1)n}{n^2}\, \alpha$$

$$\lim_{n \to \infty} \text{var}\left[\frac{X_1 + \dots + X_n}{n}\right] = \alpha$$

DISTRIBUZIONI DI FUNZIONI DI VARIABILI CASUALI

Date n v.a. $X_1, \dots, X_n$ e k funzioni $g_1, \dots, g_k$ delle n v.a. Vogliamo trovare la distribuzione congiunta di $Y_1, \dots, Y_k$ con $Y_i = g_i(X_1, \dots, X_n)$.

Se è nota la distribuzione congiunta di $X_1, \dots, X_n$ la funzione di ripartizione congiunta di $Y_1, \dots, Y_k$ soddisfa:

$$F_{Y_1, \dots, Y_k}(y_1, \dots, y_k) = P[Y_1 \le y_1, \dots, Y_k \le y_k]$$

$$= P[g_1(X_1, \dots, X_n) \le y_1, \dots, g_k(X_1, \dots, X_n) \le y_k]$$

per $y_1, \dots, y_k$ fissati.

In generale non è facile da valutare.

Analizzeremo un metodo per calcolare la distribuzione di funzioni di variabili casuali detto metodo DIRETTO.

APPLICAZIONI

Date $X_1, \dots, X_n$ v.a. indipendenti e (+ i.d.)

$$Y = X_1 + \dots + X_n$$

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

È possibile dimostrare:

1) X_i BERNOULLI di parametro p $\Rightarrow$ Y BINOMIALE di parametri (n, p)

2) X_i POISSON di parametro λ $\Rightarrow$ Y POISSON di parametro $n\lambda$

3) X_i ESPONENZIALI di param. λ $\Rightarrow$
   - Y GAMMA di parametri (n, λ)
   - $\bar{X}_n$ GAMMA di parametri $(n, n\lambda)$

4) X_i NORMALI di parametri (μ, σ^2) $\Rightarrow$
   - Y NORMALE di parametri $(n\mu, n\sigma^2)$
   - $\bar{X}_n$ NORMALE di parametri $\left(\mu, \frac{\sigma^2}{n}\right)$

2) Date n v.a. $X_1, \dots, X_n$ continue

$$Y_n = \max [X_1, \dots, X_n]$$

Per esempio, per un dato $\omega \in \Omega$, $Y_n(\omega)$ è il più grande dei numeri reali $X_1(\omega), \dots, X_n(\omega)$ $\forall \omega \in \Omega$.

$$F_{Y_n}(y) = P[Y_n \leq y] = P[X_1 \leq y, \dots, X_n \leq y]$$

- Ipotesi aggiuntiva $X_1, \dots, X_n$ sono indipendenti

$$P[X_1 \leq y, \dots, X_n \leq y] = P[X_1 \leq y] \cdot \dots \cdot P[X_n \leq y]$$

$$= \prod_{i=1}^{n} P[X_i \leq y] = \prod_{i=1}^{n} F_{X_i}(y)$$

- Ipotesi ulteriore $F_{X_1} = F_{X_2} = \dots = F_{X_n} := F_X$

$$\prod_{i=1}^{n} F_{X_i}(y) = [F_X(y)]^n$$

$$f_{Y_n}(y) = \frac{d}{dy} F_{Y_n}(y) = n [F_X(y)]^{n-1} f_X(y)$$

3) Date n v.a. $X_1, \dots, X_n$ continue

$$Y_1 = \min [X_1, \dots, X_n]$$

$$F_{Y_1}(y) = P[Y_1 \leq y] = 1 - P[Y_1 > y]$$

$$= 1 - P[X_1 > y, \dots, X_n > y]$$

Se X_i sono indipendenti, $i=1, \dots, n$

$$P[X_1>y, \dots, X_n>y] = P[X_1>y] \dots P[X_n>y]$$

$$= \prod_{i=1}^{n} P[X_i>y]$$

$$P[X_i>y] = 1-P[X_i \leq y] = 1 - F_{X_i}(y)$$

Pertanto

$$F_{Y_1}(y) = 1 - \prod_{i=1}^{n} [1-F_{X_i}(y)]$$

Se le F_{X_i} sono tutte uguali (ad F_X)

$$F_{Y_1}(y) = 1 - [1-F_X(y)]^n$$

$$f_{Y_1}(y) = \frac{d}{dy} F_{Y_1}(y) = n\,[1-F_X(y)]^{n-1} f_X(y)$$

Esempio

Una lampadina ha una distribuzione esponenziale con media pari a 100 ore.

10 lampadine vengono installate contemporaneamente.

- Qual è la distribuzione della durata della lampadina che si esaurisce prima?
- Qual è la sua durata attesa?

X_i = vita della i-esima lampadina

$Y_1 = \min[X_1, .., X_{10}]$ durata della lampadina che si esaurisce prima

X_i sono indipendenti, ident. distr. (i.i.d.)

$$f_{X_i}(x) = \frac{1}{100} e^{-\frac{1}{100}x} \qquad 0 < x < +\infty$$

perchè ogni X_i è esponenziale con $E[X] = 100$ e

$$E[X] = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{100}.$$

$$\Rightarrow F_{X_i}(x) = \left(1 - e^{-\frac{1}{100}x}\right) \qquad 0 < x < +\infty$$

Perciò

$$F_{Y_1}(y) = 1 - [1 - F_X(y)]^n$$

$$= 1 - \left(\cancel{1} - \cancel{1} + e^{-\frac{1}{100}y}\right)^{10} = 1 - \left(e^{-\frac{1}{100}y}\right)^{10}$$

$$f_{Y_1}(y) = n\,[1 - F_X(y)]^{n-1} f_X(y)$$

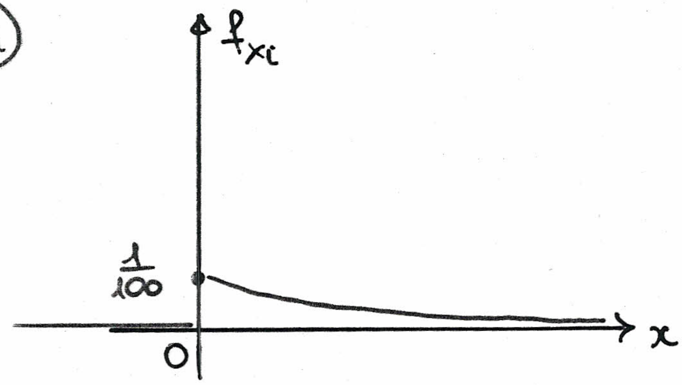
$$= 10\left(e^{-\frac{1}{100}y}\right)^9 \left(\frac{1}{100} e^{-\frac{1}{100}y}\right) \qquad 0 < y < \infty$$

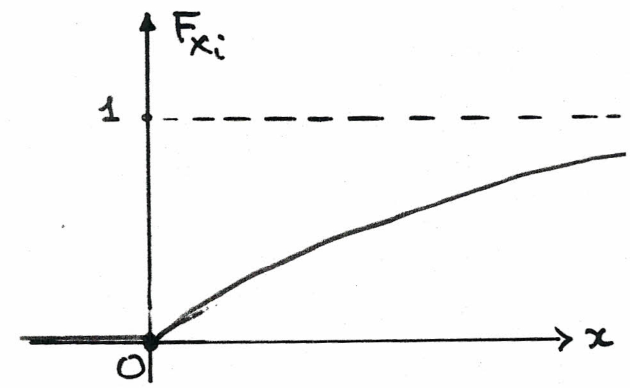
$$f_{Y_1}(y) = \frac{10}{100} \, e^{-\frac{10}{100} y} = \frac{1}{10} \, e^{-\frac{1}{10} y} \qquad 0 < y < \infty$$

cioè un'esponenziale con parametro $\lambda_1 = \frac{1}{10}$.

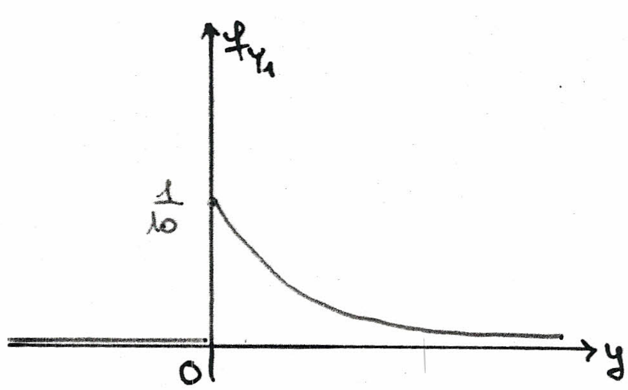
$$\Rightarrow E[Y_1] = \frac{1}{\lambda_1} = 10 \text{ ore}.$$

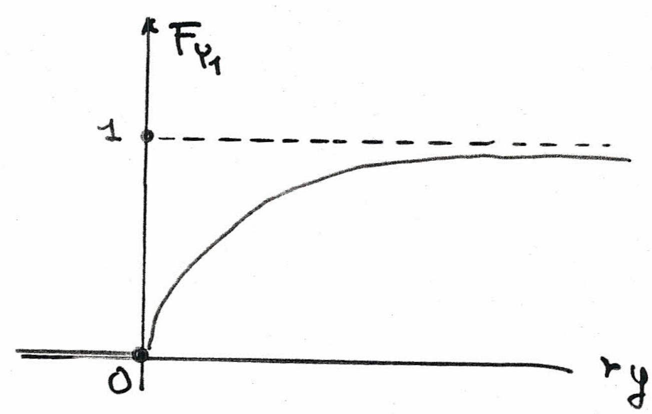
X_i





Y_1





Esempio

Le precipitazioni annuali a BS hanno distribuzione normale con media 12,08 mm e deviazione standard 3,1 mm. Le precipitazioni di anni successivi sono indipendenti.

1) Calcolare la probabilità che le precipitazioni dei prossimi 2 anni superino i 25 mm.

2) Calcolare la probabilità che le precipitazioni dell'anno prossimo superino quelle dell'anno successivo per più di 3 mm.

$X \sim N(\mu, \sigma^2)$ $\quad \mu = 12,08$

$\sigma = 3,1 \Rightarrow \sigma^2 = 9,61$

$$Z = \frac{X-\mu}{\sigma} = \frac{X - 12,08}{3.1} \sim N(0,1)$$

1

X_1 per 1° anno
X_2 per 2° anno
indipendenti

$$Y = X_1 + X_2 \sim N(\mu_Y, \sigma_Y^2)$$

dove

$$\mu_Y = E[Y] = E[X_1] + E[X_2] = 2E[X] = 24,16 \; (= n\mu) \; (n=2)$$

$$\sigma_Y^2 = var[Y] = var[X_1] + var[X_2] = 2\, var[X] = 19,22 \; (= n\sigma^2)$$

$$P[Y > 25] = P\left[\frac{Y-\mu_Y}{\sigma_Y} > \frac{25-\mu_Y}{\sigma_Y}\right] = P\left[Z > \frac{0.84}{\sqrt{19,22}}\right]$$

$$\simeq P[Z > 0.1916] \cong 1 - P[Z < 0.1916]$$

$$\cong 1 - 0.576 \simeq 0.424$$

② sfrutto il fatto che se

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

$Y = \alpha X + \beta \quad \alpha, \beta \in \mathbb{R}, \ \alpha \neq 0$

$Y \sim N(\mu_Y = \alpha \mu_X + \beta, \ \sigma_Y^2 = \alpha^2 \sigma_X^2)$

$P[X_1 > X_2 + 3] = P[X_1 - X_2 > 3] = ?$

$-X_2 \sim N(-\mu_X, \sigma_X^2)$ perchè $\alpha = -1, \beta = 0$

$T = X_1 - X_2$ è normale $\sim N(\mu_T, \sigma_T^2)$

$\mu_T = E[T] = E[X_1] - E[X_2] = \mu_X - \mu_X \equiv 0$

$\sigma_T^2 = var[T] = var[X_1] + var[X_2] = 2 var[X] = 19.22$

$P[T > 3] = P\left[\frac{T}{\sqrt{19.22}} > \frac{3}{\sqrt{19.22}}\right] \cong P[Z > 0.6843]$

$\cong 1 - P[Z < 0.6843] \cong 1 - 0.754 \cong 0.246$

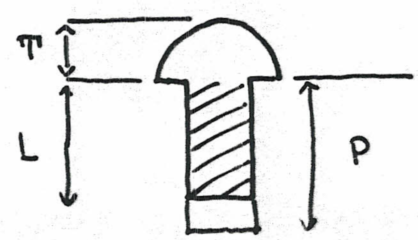
Esercizio

Una macchina lavora un pezzo producendo un foro con profondità P dove $P \sim N(\mu_P = 10\ mm, \sigma = 0{,}2\ mm)$

Nel foro va avvitata una vite con filetto L dove $L \sim N(\mu_L = 8{,}9\ mm, \sigma = 0.3\ mm)$ e testa T dove $T \sim N(\mu_T = 2\ mm, \sigma = 0.1\ mm)$.

P, L, T sono v.a. indipendenti

vincolo $T \leq 2.3\ mm$.



1. Calcolare la prob. che la sua testa sia troppo alta.
2. Calcolare la distribuzione della differenza P-L e la probabilità che il filetto non stia nel foro.
3. Calcolare la probabilità che la vite non sporga dal foro e che la testa sia alta meno di 2.3 mm.

1. $P[T > 2.3] = P\left[\frac{T-\mu_T}{\sigma_T} > \frac{2.3-2}{0.1}\right] = P\left[Z_T > \frac{0.3}{0.1}\right]$

$= P[Z_T > 3] = 1 - P[Z_T \leq 3]$ TAVOLE

$= 1 -$

$D = P - L$

$E[D] = E[P] - E[L] = \mu_P - \mu_L = 10 - 8.9 = 0.1$

$var[D] = var[P] + var[L] - 2\underbrace{cov(P, L)}_{=0}$ perché indip.

$= (0.2)^2 + (0.1)^2 = 0.04 + 0.01 = 0.05$

La differenza di normali è una normale.

$D \sim N(\mu_D = 0.1 \text{ mm}, \ \sigma^2 = 0.05 \text{ mm})$

$P[L>P] = P[P-L<0] = P[D<0] = P\left[\frac{D-\mu_D}{\sigma_D} < \frac{-0.1}{\sqrt{0.05}}\right]$

$= P\left[Z_D < -\sqrt{\frac{1}{5}}\right] \cong P[Z_D < -0.4472]$

$= P[Z_D > 0.4472] = 1 - P[Z_D \leq 0.4472]$ TAVOLE

$= 1 -$

$P[L<P] = 1 - P[L \geq P] = P[Z_D \leq 0.4472] =$

$P[T < 2.3] = 1 - P[T \geq 2.3] = P[Z_T < 3] =$