 LEZIONE 6 06/04/09

Esempio di applicazioni alle v.c. congiunte

Data la d.d.p. congiunta f_{XY}, calcolare la f.d.d. e la d.d.p. di

$$Z = X + Y$$

Determiniamo prima $F_Z(z)$ e, per derivazione, $f_Z(z)$

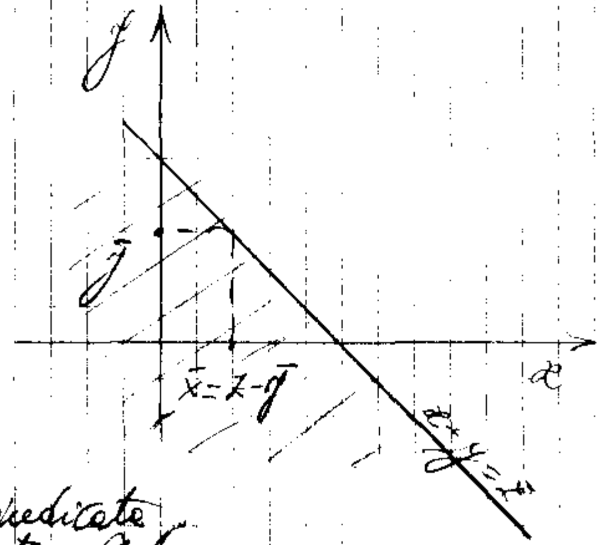
Ricordiamo che

$D_Z = \{(x,y) : g(x,y) \le z\}$

essendo $g(x,y) = x + y$

Perciò

$D_Z = \{(x,y) : x + y \le z\}$



Dato uno z generico la funzione $x + y = z$ è rappresentata dalla retta indicata mentre la regione tratteggiata rappresenta la condizione

$$x + y \le z$$

Ricordiamo (pagg 94 e 95) che

$$\cancel{F_Z(z) = \int_{-\infty}^{+\infty}}$$

$$F_Z(z) = \iint_{D_Z} f_{XY}(x,y)\,dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{z-y} f_{XY}(x,y)\,dx \cdot dy$$

L'integrale più interno è calcolato rispetto ad x perciò le strisce in cui è scomposta la regione hanno l'asse longitudinale in direzione dell'asse x.

L'integrale più esterno calcolato rispetto ad y indica che le strisce vengono poi affiancate in modo tale da coprire tutta la regione.

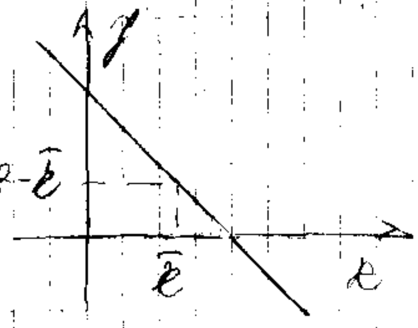
Una striscia orizzontale è definita per un determinato y e si estende da $-\infty$ fino a

$$\bar{x} = z - \bar{y}$$

Osserviamo che nell'estensione dell'integrale interno la y è bloccata ad un determinato valore.

Le strisce possono essere costruite parallelamente all'asse y e si estenderebbero da $-\infty$ fino a

$$\bar{y} = z - \bar{x}$$



Perciò si può anche scrivere

$$F_Z(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x} f_{XY}(x,y)\, dy\, dx$$

Indichiamo e definiamo $\tilde{y}$. $\quad \tilde{y} := y + x$

con la conseguenza che $\quad d\tilde{y} = dy$

Quando $y = z - x$, ne consegue: $\quad \tilde{y} = z$

$$y = \tilde{y} - x$$

Usando la variabile y l'integrale è esteso da $-\infty$ fino a $z-x$.

Usando la variabile $\tilde{y}$, l'integrale è esteso da $-\infty$ fino a z.

L'integrale diventa

$$F_Z(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z} f_{XY}(x, \tilde{y} - x)\, d\tilde{y}\, dx$$

Per il calcolo della d.d.p. $f_Z(z)$ ricordiamo che

$$\frac{d}{dz} \int_{a}^{z} g(y)\, dy = g(z)$$

Si ottiene

$$f_Z(z) = \frac{dF_Z}{dz} = \left(\text{portando } d/dz \text{ dentro il primo integrale}\right)$$

$$= \int_{-\infty}^{+\infty} \frac{d}{dz} \int_{-\infty}^{z} f_{XY}(x, \tilde{y} - x)\, d\tilde{y}\, dx =$$

$$= \int_{-\infty}^{+\infty} f_{XY}(x, z-x)\, dx$$

Per simmetria si ottiene

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{XY}(z-y, y)\, dy$$

• **Teorema**

Se X ed Y sono v.c. indipendenti la d.d.p. di

$$Z = X + Y$$

è data da

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\, dx = \text{(per simmetria)}$$

$$= \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\, dy$$

Occorre sostituire le relazioni di pag. 93 (per v.c. indipendenti)

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

con $y = z - x$ e $x = z - y$

I due ultimi integrali sono detti di convoluzione.

Si dice anche che f_Z è la convoluzione di f_X e f_Y, usando la simbologia

$$f_Z = f_X * f_Y$$

Esempio. Determinare la d.d.p. di $Z = X + Y$ essendo X e Y v.c. indipendenti ed entrambe

a) uniformi su $[0, 1]$

b) esponenziali con $E[X] = E[Y] = \frac{1}{\lambda}$

a) $f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{altrove} \end{cases}$

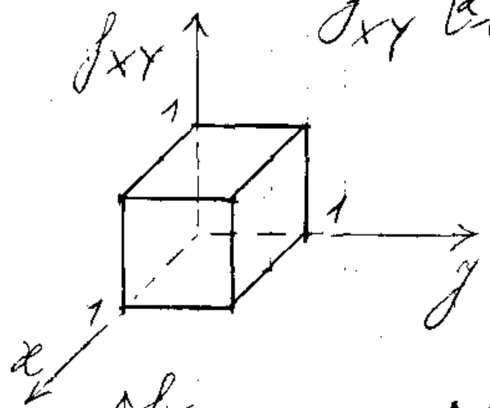
b) $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrove} \end{cases}$

a) Si può procedere in due modi

— a1) Calcoliamo la $F_Z(z)$

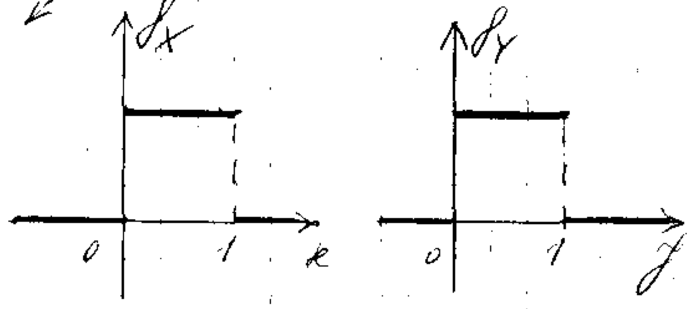
Serve la d.d.f. congiunta

$$f_{XY}(x,y) = f_X(x)\, f_Y(y) \quad \text{(per l'indipendenza)}$$

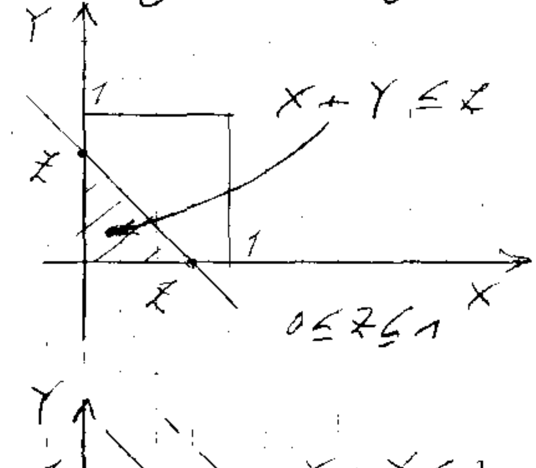


La $f_{XY}(x,y)$ è diversa da zero solo se x e y sono compresi tra 0 e 1.

$$f_{XY}(x,y) = \begin{cases} 1 & 0\le x\le 1,\ 0\le y\le 1 \\ 0 & \text{altrove} \end{cases}$$

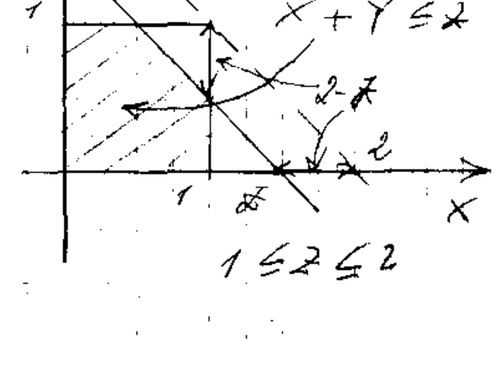


$$F_Z(z) = P(Z\le z) = P(X+Y\le z)$$



La d.d.f. indica la possibilità di scegliere un punto nel quadrato in modo equiprobabile.

I punti appartenenti alla zona tratteggiata hanno la proprietà di rendere $X + Y$ minore di z.



Se z è compreso tra 0 e 1 la zona è triangolare altrimenti è per completezza può essere considerata come un quadrato meno un triangolo.

Ricordando quanto detto alle pagg. 10 e ss. occorre fare il rapporto tra l'area tratteggiata e l'area del quadrato.

Se $0\le z\le 1$, l'area tratteggiata vale $\frac{z^2}{2}$

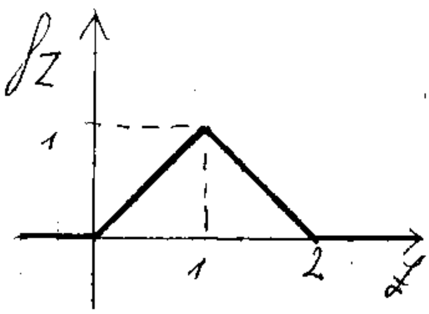
Se $1\le z\le 2$, l'area diventa $1-\frac{(2-z)^2}{2}$

Perciò

$$F_Z(z) = \begin{cases} 0 & z<0 \\ \frac{z^2}{2} & 0 \le z \le 1 \\ 1-\frac{(2-z)^2}{2} & 1 \le z \le 2 \\ 1 & z>2 \end{cases}$$

Si ottiene

$$f_Z(z) = \frac{dF_Z(z)}{dz} = \begin{cases} 0 & z<0 \\ z & 0 \le z \le 1 \\ 2-z & 1 \le z \le 2 \\ 0 & z>2 \end{cases}$$



N.B.: è opportuno controllare sempre che l'area sottesa da $f_Z(z)$ sia uguale a 1.

<u>Osservazione</u>: abbiamo sommato due v.c. indipendenti uniformi tra $[0,1]$.
La f_Z ottenuta non è uniforme.

-) a2 Calcolo diretto di $f_Z(z)$ mediante la convoluzione

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(k)\, f_X(z-k)\, dk$$

Ricordiamo che

$$f_X(k) = \begin{cases} 1 & 0 \le k \le 1 \\ 0 & \text{altrove} \end{cases}$$

Ne consegue che

$$f_X(z-k) = \begin{cases} 1 & 0 \le z-k \le 1 \iff z-1 \le k \le z \\ 0 & \text{altrove} \end{cases}$$

Si ricava che:

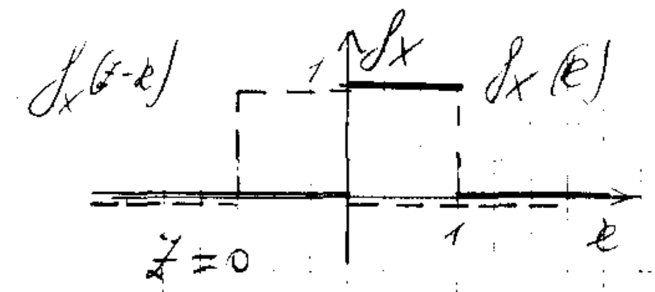
$$\cancel{f_Z(z)} = f_Z(z) = \int_{\max(0,z-1)}^{\min(1,z)} 1 \cdot 1 \cdot dk =$$

$$= \begin{cases} 0 & z<0 \\ \int_0^z dk = z & 0\le z\le 1 \\ \int_{z-1}^1 dk = 1-z+1 = 2-z & 1\le z\le 2 \\ 0 & z\ge 2 \end{cases}$$

L'integrale è calcolato nell'intervallo in cui la funzione $f_X(k)$ è diversa da zero

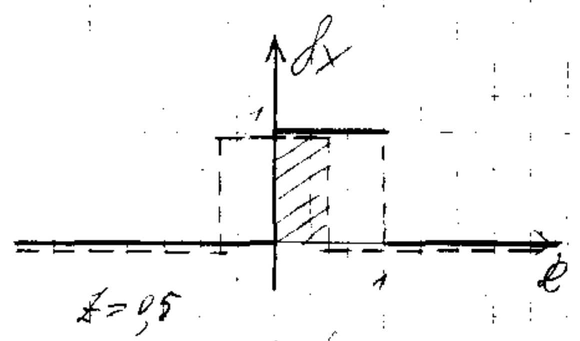
Occorre tener conto di entrambe le condizioni
$0\le k\le 1$
e
$z-1\le k\le z$

Visualizziamo graficamente il risultato per alcuni valori di z.



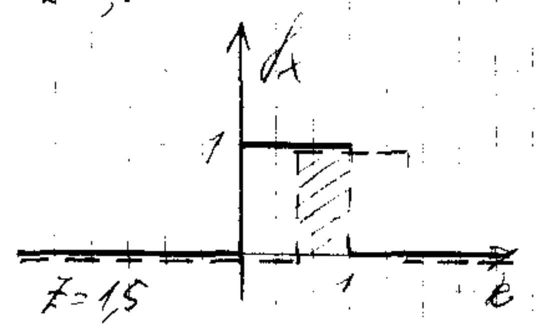
$z=0 \Rightarrow f_X(z-k)=f_X(-k)$ risulta la $f_X(k)$ ribaltata e traslata di z.

$\int_{-\infty}^{+\infty} f_X(k)\, f_X(z-k)\, dk = 0$



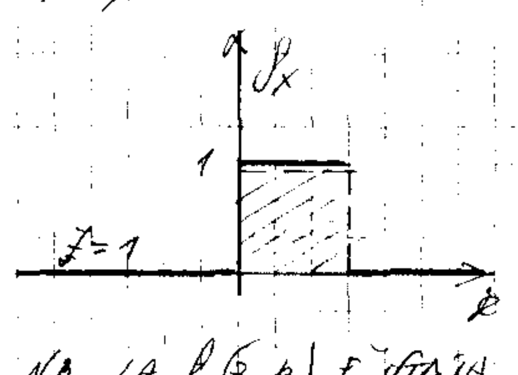
Per $z=0,5$ le due funzioni si sovrappongono parzialmente

$\int_{-\infty}^{+\infty} f_X(k)\, f_X(z-k)\, dk = \frac{1}{2} = f_Z(0,5)$



L'integrale del prodotto delle due funzioni è l'area tratteggiata

Per $z=1,5$ $f_Z(1,5)=\frac{1}{2}$



Per $z=1$ $f_Z(1)=1$ e si ha il valore massimo della convoluzione

N.B. LA $f_X(z-k)$ È STATA LEGGERMENTE TRASLATA VERSO IL BASSO PER RAGIONI GRAFICHE

102 caso h) esponenziale

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrove} \end{cases} \qquad E[X] = E[Y] = \frac{1}{\lambda}$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_X(z-x)\, dx$$

Osserviamo che

$$f_X(z-x) = \begin{cases} \lambda e^{-\lambda(z-x)} & z-x \geq 0 \Rightarrow x \leq z \\ 0 & z-x < 0 \Rightarrow x > z \end{cases}$$

L'integrale diventa

$$f_Z(z) = \int_0^z \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(z-x)}\, dx =$$

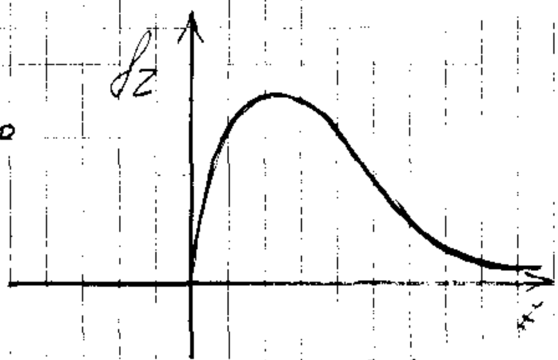
$$= \lambda^2 \int_0^z e^{-\lambda(x+z-x)}\, dx = \lambda^2 \int_0^z e^{-\lambda z}\, dx =$$

$$= \lambda^2 e^{-\lambda z} \int_0^z dx = \lambda^2 e^{-\lambda z} \cdot z \qquad \text{per } z \geq 0$$

Se $z < 0 \Rightarrow f_Z(z) = 0$

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & z \geq 0 \\ 0 & z < 0 \end{cases}$$

La funzione prende il nome di ERLANG-2



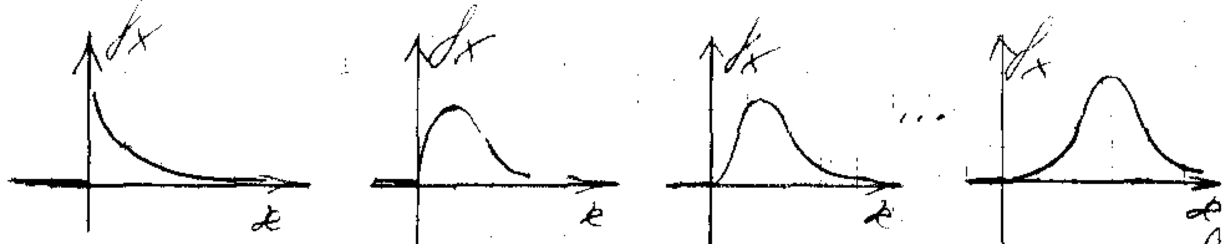
ERLANG-2

(AGNER KRARUP ERLANG, LONBORG, 1878 – COPENAGHEN, 1929)

La Erlang-2 risulta la somma di due v.c. esponenziali.
Per la somma di k variabili esponenziali si considera la Erlang-k che risulta

$$\text{Erlang-}k = \sum_{i=1}^{k} X_i$$, con X_i v.c. esponenziali indipendenti con $E[X_i] = \frac{1}{\lambda}$

All'aumentare di k le Erlang tendono a diventare delle Gaussiane.



ERLANG-1 ERLANG-2 ERLANG-3 ERLANG-k

Si è visto che né nel caso, né nel caso si sommando due funzioni dello stesso tipo si ottiene una d.d.f. di tipo diverso, anzi molto diverso.
Tra le funzioni che sommate, mantengano una d.d.f. che ha la stessa forma di quelle iniziali, citiamo le Gaussiane.

Momenti di variabili congiunte

Media di $g(X,Y)$

Per definizione, se $Z = g(X,Y)$ si ha

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

Sembra che si debba ricorrere al calcolo di $f_Z(z)$

Teorema (che permette di evitare il calcolo di $f_Z(z)$)

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

Conseguenza: linearità dell'operatore di aspettazione

$$E[X+Y] = E[X] + E[Y]$$

Dimostrazione:

$$E[X+Y] = \text{(applicando il teorema)} =$$

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

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

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

Osserviamo che

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

Integrando la densità congiunta lungo una delle due variabili si ottiene la densità marginale dell'altra variabile. (pag. 91)

Analogamente

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

Otteniamo:

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

$$= E[X] + E[Y]$$

-) <u>Proprietà generale</u>

$$E\left[\sum_i a_i X_i\right] = \sum_i a_i E[X_i]$$

Il valore atteso di una combinazione lineare di ~~tot~~ variabili casuali è la combinazione lineare dei valori attesi delle singole variabili.

•) <u>Momenti congiunti</u>

-) <u>Momenti di ordine n</u>

$$m_{kr} = E[X^k Y^r] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^k y^r f_{XY}(x,y)\,dx\,dy$$

con la condizione: $k + r = n$

-) <u>Momenti di ordine 1</u>

Dovendo essere $k+r=1$, si hanno due possibilità.
I momenti possono essere solo $E[X]$ ed $E[Y]$

-) <u>Momenti di ordine 2</u>

Sono $E[X^2]$, $E[XY]$, $E[Y^2]$

Il momento incrociato $E[XY]$ è detto anche correlazione.

-) Momenti centrali di ordine n

$$\mu_{kr} = E\left[(X-E[X])^k (Y-E[Y])^r\right] \quad \text{con } k+r=n$$

I momenti di v.c. congiunte sono importanti poiché permettono di capire come si influenzano le variabili e quali sono le relazioni che intercorrono tra di esse.

•) Covarianza e coefficiente di correlazione

$E[X^2]$ è il valore quadratico medio di X e fornisce un valore medio di quanto la variabile si scosti dall'origine.

Continua a valere la relazione

$$Var[X] = E[X^2] - E[X]^2$$

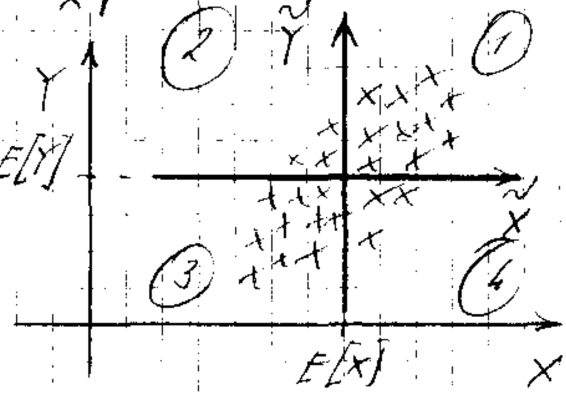
$$\mu_{11} = E\left[(X-E[X])(Y-E[Y])\right] = Cov[X,Y] \quad \text{(covarianza)}$$

μ_{11} si può anche indicare con σ_{XY}.

Interpretazione

Si consideri una coppia di v.c. X e Y (ad es. peso ed altezza) e si raccolgano dei dati inserendoli poi in un grafico.

Siano $E[X]$ ed $E[Y]$ i valori attesi delle due variabili.



SCATTER PLOT

(GRAFICO DI DISPERSIONE)

Ogni punto del grafico rappresenta la coppia peso-altezza di una persona campionata.

Definiamo due nuove variabili: $\tilde{X} = X - E[X]$, $\tilde{Y} = Y - E[Y]$

Ne consegue che la covarianza risulta:

$$\tilde{\sigma}_{XY} = E[\tilde{X}\tilde{Y}]$$

Ci chiediamo quale sia il significato di una covarianza positiva.

$$\tilde{\sigma}_{XY} > 0$$

Una covarianza positiva significa che, mediamente, il prodotto

$$\tilde{X}\tilde{Y} > 0$$

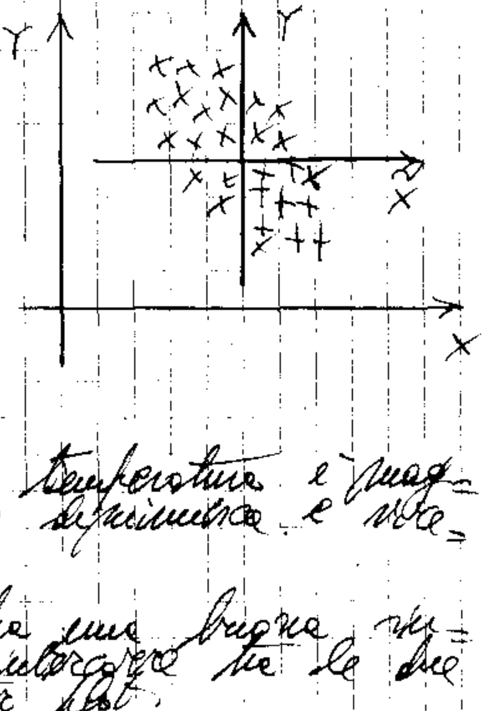
Ne consegue che le due variabili $\tilde{X}$ e $\tilde{Y}$ hanno lo stesso segno almeno mediamente ed i loro valori cadono prevalentemente nei quadranti 1 e 3.

Ne consegue anche che, se X è superiore alla propria media, ci si deve aspettare che anche Y lo sia.

Vale anche il viceversa: se X è ~~superiore~~ inferiore alla propria media, ci si aspetta che anche Y stia sotto la propria media.

Tutto ciò va inteso in senso probabilistico, anche se ci si aspetta che un individuo con un peso superiore alla media della popolazione abbia pure un'altezza superiore alla media.

Consideriamo un altro scatter plot, immaginando di monitorare la temperatura media giornaliera ed il consumo di gasolio per il riscaldamento.



È un caso di covarianza negativa

$$\tilde{\sigma}_{XY} < 0$$

È logico pensare che, quando la temperatura è maggiore, il consumo di gasolio diminuisce e viceversa.

Dal valore della covarianza si ha una buona informazione sulla relazione che intercorre tra le due variabili ed anche sullo scatter plot.

–) Proprietà

–) $\tilde{\sigma}_{XY} = E[XY] - E[X]\cdot E[Y]$

$E[XY]$ è > 0 quando X e Y hanno lo stesso segno.

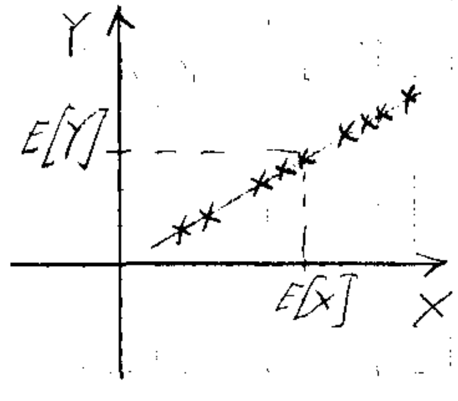
-) $(Cov[X,Y])^2 \le Var[X]\,Var[Y]$

Si può anche scrivere

$$|\sigma_{XY}| \le \sigma_X \sigma_Y$$

-) $\sigma_{XY}^2 = \sigma_X^2 \sigma_Y^2$ se e solo se vale la relazione

$$P\left(Y = \frac{\sigma_{XY}}{\sigma_X^2} \cdot (X - E[X] + E[Y])\right) = 1$$



La relazione vale quando, nello scatter plot, i dati cadono su una retta, e viceversa.

Si dice che tra X e Y esiste un legame lineare con probabilità pari ad 1.

Ci chiediamo se dallo scatter plot si possa ricavare il valore della covarianza.

Ci chiediamo anche se, dal valore della covarianza, si possa capire come sono le relazioni tra le due variabili in maniera precisa.

La covarianza è un numero che dipende dalle unità di misura.

Inoltre non è generalmente possibile ricavare il valore della covarianza dallo scatter plot.

C'è un'altra grandezza che risolve tutti i problemi mantenendo i pregi della covarianza ed è il coefficiente di correlazione

-) <u>Coefficiente di correlazione</u>

Definizione:

$$r_{XY} := \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Dalla seconda proprietà (la prima in questa pagina) si ricava

$$|r_{XY}| \le 1$$

Dalla terza proprietà si ha:

$$|r_{XY}| = 1$$

se e solo se esiste un legame lineare con probabilità 1.

Il coefficiente di correlazione è normalizzato, poiché è sempre compreso tra -1 e 1.

r_{XY} è adimensionale e non dipende dalle unità di misura utilizzate.

Se $|r_{XY}|=1$ i dati sullo scatter plot stanno su una retta con il coefficiente angolare avente il segno di r_{XY}.

Il coefficiente di correlazione è molto più leggibile della covarianza.

Osservazione: le v.c. X e Y hanno la stessa covarianza e lo stesso coefficiente di correlazione delle v.c.

$$(X+a) \quad , \quad (Y+b)$$

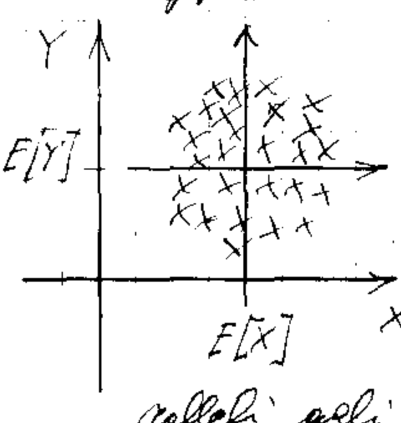
Traslando due v.c. covarianza e coefficiente di correlazione non cambiano.

Quesito: può dirsi la stessa cosa per le v.c.?

$$(\alpha X+a) \quad e \quad (\beta Y+b)$$

La risposta verrà fornita alla prossima lezione.

Definizione: se $\sigma_{XY}=0$, X e Y si dicono incorrelate.



Ci sarebbe la possibilità di usare il coefficiente di correlazione per definire l'incorrelazione ma c'è la difficoltà di considerare che uno dei termini al denominatore può essere nullo.

Se lo scatter plot è a forma di palla oppure di ellisse con gli assi paralleli agli assi coordinati allora il coefficiente di correlazione è nullo e le variabili sono incorrelate.

In questo caso è molto difficile fare previsioni sul comportamento di una variabile in base all'altra.

Nota bene: incorrelazione e indipendenza che nel linguaggio italiano sono quasi sinonimi, nella nostra disciplina non lo sono affatto.

Teorema: se X e Y sono v.c. indipendenti, sono anche incorrelate.

Non vale il viceversa. Indipendenza implica incorrelazione, se sono incorrelate non è detto che siano indipendenti.

Dimostrazione

Mostriamo che $E[XY] = E[X]E[Y]$ e ciò implica che $\operatorname{cov}[X,Y] = \underline{E[XY]} - E[X]E[Y] = 0$ (pag. 106)

$$E[XY] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f_{XY}(x,y)\, dx\, dy = \text{(per l'indipendenza, pag. 95)}$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f_X(x) f_Y(y)\, dx\, dy =$$

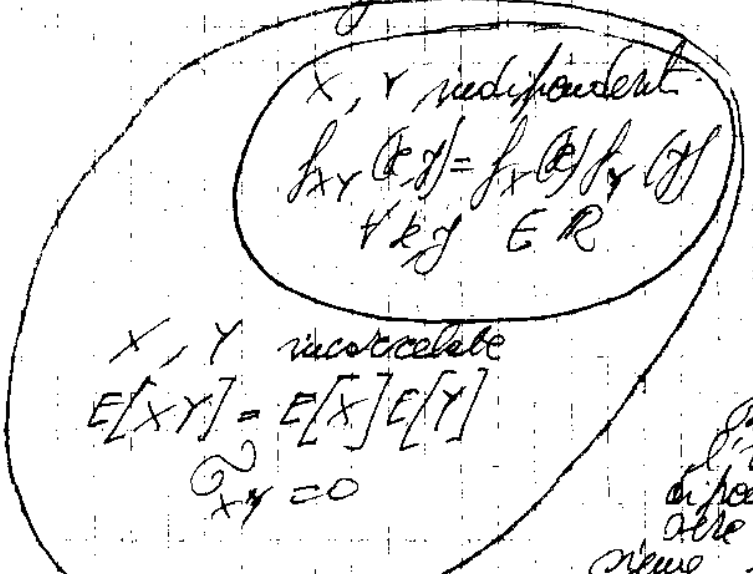
$$= \int_{-\infty}^{+\infty} y\, f_Y(y) \int_{-\infty}^{+\infty} x\, f_X(x)\, dx \cdot dy =$$

$$= \int_{-\infty}^{+\infty} y\, f_Y(y)\, E[X]\, dy =$$

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

In generale l'incorrelazione non garantisce l'indipendenza.

Usando i diagrammi di Venn:



l'insieme delle coppie X, Y indipendenti è minore dell'insieme delle coppie incorrelate e sta all'interno dell'insieme delle coppie incorrelate.

Essere all'interno dell'insieme delle coppie indipendenti garantisce di essere all'interno dell'insieme delle coppie incorrelate.

L'indipendenza è una condizione più forte dell'incorrelazione.

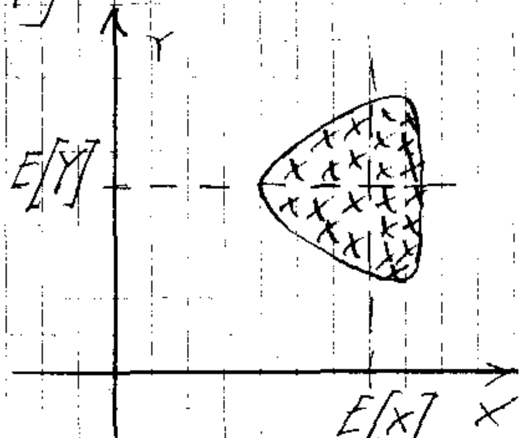
La condizione $f_{XY}(x,y) = f_X(x)\, f_Y(y)$ deve valere per ogni coppia x, y cioè per ogni scelta di X e Y.

La condizione in effetti implica un numero infinito di equazioni.

L'incorrelazione equivale alla condizione

$$E[XY] = E[X]\,E[Y]$$

che è una sola equazione.



Esempio

Nello scatter plot raffigurato se la X è superiore al valor medio, non si può sapere se la Y sta sopra o sotto il valor medio, poiché la Y è molto dispersa

Se invece la X è inferiore al valor medio, si sa che la Y è poco dispersa e questa è, comunque, un'informazione.

L'esempio riguarda due v.c. incorrelate, ma non indipendenti.

Proposizione

Se $Z = X \pm Y$ allora $\sigma_Z^2 = \sigma_X^2 \pm 2\sigma_{XY} + \sigma_Y^2$

Ciò significa che la varianza della somma non è, in generale, la somma delle varianze

Dimostrazione

$$\sigma_Z^2 = E\left[\left(Z - E[Z]\right)^2\right] = E\left[\left(X \pm Y - E[X \pm Y]\right)^2\right] =$$

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

$$= E\left[\left(X - E[X]\right)^2\right] \pm 2E\left[\left(X - E[X]\right)\left(Y - E[Y]\right)\right] +$$

$$+ E\left[\left(Y - E[Y]\right)^2\right] =$$ (ricordando la proprietà di pag. 103)

$$= \sigma_X^2 \pm 2\sigma_{XY} + \sigma_Y^2$$ (vedi pag. 66)

-) Proposizione

Se fosse $Z = X \pm Y$, con X e Y incorrelate, allora

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2,$$

essendo $\sigma_{XY} = 0$

Si noti che se X, Y sono incorrelate

$$Var[X - Y] = \sigma_X^2 + \sigma_Y^2$$

e

$$Var[X - Y] \neq \sigma_X^2 - \sigma_Y^2$$

La varianza di Y potrebbe essere maggiore di quella di X. Se nell'ultima relazione valesse l'uguale si potrebbe avere una varianza di $(X - Y)$ negativa.

Vale la proprietà

$$E[Z] = A \cdot E[X] + B \cdot E[Y]$$

con $E[X]$ ed $E[Y]$ vettori.

VARIABILI CASUALI VETTORIALI

Estendiamo al caso di n v.c. quanto detto per due v.c.

Sia data una v.c. X

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

X e' un vettore di numeri

-) d.d.p.

$$f_X(x) = f_X(x_1, x_2, \dots, x_n) := \mathbb{R}^n \mapsto \mathbb{R}^1$$

-) Valore atteso

$$E[X] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_n] \end{bmatrix}$$

Il valore atteso del vettore X e' il vettore dei valori attesi dei singoli componenti.

-) Proprieta'

Siano X e Y dei vettori casuali e definiamo

$$Z = AX + BY$$

Supponiamo che X sia un vettore $(n \times 1)$.

Y puo' avere dimensioni diverse e sia $(m \times 1)$

A e' una matrice $(p \times n)$.

B e' una matrice $(p \times m)$

Di conseguenza Z e' un vettore $(p \times 1)$

-) Matrice varianza di un vettore casuale X

$$Var[X] = E\left[(X - E[X]) \cdot (X - E[X])^T\right]$$

dove

$X - E[X]$ è un vettore $(n \times 1)$
$(X - E[X])^T$ è un vettore $(1 \times n)$
La $Var[X]$ è una matrice $(n \times n)$

$$Var[X] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

dove

$$\sigma_{ij} = E\left[\left(X_i - E[X_i]\right) \cdot \left(X_j - E[X_j]\right)\right] = \begin{cases} Cov[X_i\, X_j] & \text{se } i \neq j \\ Var[X_i] & \text{se } i = j \end{cases}$$

Lungo la diagonale principale ci sono le varianze dei componenti il vettore X.
Fuori dalla diagonale principale ci sono tutte le possibili covarianze.

Osservazione

•) $\sigma_{ij} = \sigma_{ji}$
la matrice varianza è simmetrica.

•) se $n = 2$
la matrice varianza diventa

$$Var\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}$$ compaiono le varianze delle due v.c. X e Y e la covarianza.

–) Teorema

se $Y = AX \Longrightarrow Var[Y] = A\, Var[X]\, A^T$
X ha dimensioni $(n \times 1)$
A ha dimensioni $(m \times n)$
Di conseguenza Y ha dimensioni $(m \times 1)$

La Var[Y] risulta avere le dimensioni $(m \times m)$
D'altra parte la Var[X] ha dimensioni $(n \times n)$, essendo X un vettore $(n \times 1)$
Moltiplicando A per Var[X] si ottiene una matrice di dimensioni

$$(m \times n) \times (n \times n) = (m \times n)$$

Infine moltiplicando quest'ultima matrice $(m \times n)$ per A^T $(n \times m)$ si ottiene una matrice di dimensioni

$$(m \times n) \times (n \times m) = (m \times m)$$

Dimostrazione

Ricordando la definizione di matrice varianza (p. 112)

$$Var[Y] = E[(Y - E[Y]) \cdot (Y - E[Y])^T] =$$

$$= E[(AX - E[AX]) \cdot (AX - E[AX])^T] =$$

$$= E[A(X - E[X]) \cdot (A(X - E[X]))^T]$$

Ricordando anche che il trasposto del prodotto di due matrici è il prodotto delle trasposte della seconda matrice per la trasposta della prima.

$$(A(X - E[X]))^T = (X - E[X])^T \cdot A^T$$

Per cui:

$$Var[Y] = E[A(X - E[X]) \cdot (X - E[X])^T \cdot A^T] =$$

$$= A \cdot E[(X - E[X]) \cdot (X - E[X])^T] \cdot A^T$$

Ma

$$E[(X - E[X]) \cdot (X - E[X])^T] = Var[X] \text{ (pag. 112)}$$

Quindi:

$$Var[Y] = A \cdot Var[X] \cdot A^T$$

- Teorema

Var[X] ≥ 0, nel senso che è una matrice semidefinita positiva

Ricordiamo che, se S è una matrice simmetrica, $S = S^T$ è ≥ 0, cioè è semidefinita positiva, se è vero che

$$x^T \cdot S\, x \geq 0 \qquad \forall x \in \mathbb{R}^n$$

con S di dimensioni $(n \times n)$

Si dimostra che la matrice varianza è sempre ~~de-finita~~ semidefinita positiva.

-) Definizione

Le v.c. vettoriali X e Y sono indipendenti se, essendo

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix},$$

la d.d.p. congiunta

$$f_Z\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = f_X(x)\, f_Y(y)$$

Si ottiene anche in questo caso, che per avere l'indipendenza, la d.d.p. congiunta deve risultare il prodotto delle d.d.p. singole, con la differenza che X e Y sono vettori.

-) Definizione

$$Cov[X, Y] = E\left[(X - E[X]) \cdot (Y - E[Y])^T\right]$$

X ha dimensioni $(n \times 1)$

Y ha dimensioni $(m \times 1)$

Y^T ha dimensioni $(1 \times m)$

Perciò la covarianza ha dimensioni $(n \times 1) \cdot (1 \times m) =$
$= (n \times m)$

-) Definizione

Se $Cov[X, Y] = 0$, allora X e Y si dicono incorrelate

È la stessa definizione di pag. 108 con la differende che, in questo caso, la covarianza è una matrice.

-) Teorema
Il fatto che X e Y siano indipendenti implica che X e Y siano incorrelate

-) Proposizione
Se $Z = X \pm Y$,
allora $Var[Z] = Var[X] \pm Cov[XY] \pm Cov[YX] + Var[Y]$
Occorre mettere entrambe le cov. poiché non è detto che siano matrici simmetriche.

-) Proposizione
Se $Z = X \pm Y$, con X e Y incorrelate
allora
$Var[Z] = Var[X] + Var[Y]$