CORRELAZIONE E COMBINAZIONI LINEARI di *n* variabili aleatorie: definizioni e proprietà

Nota Bene: Quanto segue è una sintesi schematica (anche se con molti dettagli) di quanto svolto dai docenti in aula sia a lezione sia ad esercitazione. Pertanto quanto segue **non sostituisce** tutto ciò che nei minimi dettagli è stato visto e dimostrato a lezione o ad esercitazione durante lo svolgimento del corso (vedansi, p. es., le copie dei lucidi presso Yellow Print, Castellanza).

COPPIA (X,Y) di due v.a. X e Y discrete CONGIUNTE

Due v.a. X e Y discrete si dicono **congiunte** (o considerate congiuntamente), e si scrivono allora (X,Y), quando è data la **probabilità congiunta**

$$p(x,y) = P(X = x, Y = y) \quad \forall (x,y), x \in S_X, y \in S_Y$$

cioè la probabilità di osservare ciascuna coppia (x,y) composta congiuntamente da uno dei valori possibili di una variabile, ovvero $x \in S_X$, e da uno dei valori possibili dell'altra variabile, ovvero $y \in S_Y$. Le probabilità congiunte $p(x,y)$ di tutte le coppie si possono rappresentare agli incroci delle righe e delle colonne di una **tabella a doppia entrata** (Nota bene: esiste una analoga definizione, mutatis mutandis, per X e Y v.a. **continue).**

COVARIANZA $Cov(X,Y)$ di (X,Y)

La Covarianza di (X,Y) è un numero che indica **il grado** (alto, medio, basso, ecc.) e **la direzione** (diretta o inversa) **della relazione, o associazione, lineare esistente** tra le coppie di valori delle due v.a. X e Y.
In quanto segue si considera solo il caso di due v.a. discrete (le definizioni e proprietà valgono anche, mutatis mutandis, per le coppie (X,Y) di v.a. **continue**).

<u>Simboli</u>: $Cov(X,Y) = \sigma_{XY}$.

<u>Formula</u>: $Cov(X,Y) = \sigma_{XY} = E(XY) - E(X)E(Y)$ (1)

dove $E(XY) = \sum_{\forall(x,y)} x\,y\,p(x,y)$ si dice **momento primo misto** delle v.a. (X,Y)

Date n coppie **osservate** $(x_1,y_1),(x_2,y_2),\ldots(x_n,y_n)$ di valori possibili delle due v.a. (X,Y), la covarianza **delle n coppie osservate** si calcola con formula simile (mutatis mutandis) alla (1) sopra, ovvero con la

<u>Formula</u>: $Cov(X,Y) = \sigma_{XY} = \frac{1}{n}\sum_{\forall(x_t,x_s)} x_t y_s - \bar{x}_n \bar{y}_n$ (1bis)

Proprietà di simmetria di $Cov(X,Y)$:

$$Cov(X,Y) = Cov(Y,X) \quad (\text{ ovvero } \sigma_{XY} = \sigma_{YX})$$

Disuguaglianza notevole di $Cov(X,Y)$:

$$-\infty < -\sigma_X\sigma_Y \le Cov(X,Y) = \sigma_{XY} \le \sigma_X\sigma_Y < \infty$$

COEFFICIENTE DI CORRELAZIONE LINEARE $\rho(X,Y)$ di (X,Y)

<u>Simboli</u>: $\rho(X,Y) = \rho_{XY}$

<u>Formula</u>: $\rho(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$

Le proprietà di $\rho(X,Y)$ seguono immediatamente da quelle di $Cov(X,Y)$, ovvero si ha:

Proprietà di simmetria di $\rho(X,Y)$**:**

$$\rho(X,Y) = \rho(Y,X) \quad (\text{ovvero } \rho_{XY} = \rho_{YX})$$

Disuguaglianza notevole di $\rho(X,Y)$**:**

$$-1 \le \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \rho_{XY} \le 1$$

Nota Bene: Rispetto alla disuguaglianza di $Cov(X,Y)$, la disuguaglianza di $\rho(X,Y)$ ha il vantaggio che i suoi due estremi sono **finiti e non cambiano** per **qualsiasi** coppia di v.a. mentre quelli di $Cov(X,Y)$ possono cambiare a seconda della coppia di v.a. considerate. Per il resto $\rho(X,Y)$ dà le **stesse** informazioni che dà $Cov(X,Y)$, ovvero è un numero che indica **il grado** (alto, medio, basso, ecc.) e **la direzione** (diretta o inversa) **della relazione, o associazione, lineare esistente** tra le coppie di valori delle due v.a. X e Y. In particolare, **se** $Cov(X,Y) = 0$ **allora** si ha anche $\rho_{XY} = 0$ e viceversa (le definizioni e proprietà di cui sopra valgono anche, mutatis mutandis, per le coppie (X,Y) di v.a. **continue**).

(X,Y) con X e Y CORRELATIVAMENTE INDIPENDENTI

Le v.a. X e Y (discrete o continue) di (X,Y) sono **correlativamente indipendenti** o **non correlate** se

$$\rho_{XY} = 0, \text{ ovvero } Cov(X,Y) = 0$$

(X,Y) con X e Y STOCASTICAMENTE INDIPENDENTI

Le v.a. discrete X e Y della coppia (X,Y) si dicono **stocasticamente indipendenti** se

$$p(x,y) = p(x)\,p(y) \quad \forall (x,y), x \in S_X, y \in S_Y$$

(Nota bene: esiste un'analoga definizione, mutatis mutandis, per X e Y v.a. **continue**). **Conseguenza dell'indipendenza stocastica**: Se le v.a. X e Y della coppia (X,Y) sono stocasticamente indipendenti, allora:

$$Cov(X,Y) = \rho_{XY} = 0$$

Ovvero, se le v.a. X e Y della coppia (X,Y) sono stocasticamente indipendenti, allora X e Y sono correlativamente indipendenti. **Nota bene:** non vale in generale la proposizione inversa.

(**Nota Bene:** in quanto segue gli indici $s,t = 1,2,\ldots n$ possono essere sostituiti da qualsiasi altri indici p. es. $i, j = 1,2,\ldots n$)

v.a. Y COMBINAZIONE LINEARE di *n* v.a. X_t $(t = 1,2,\ldots n)$

Una v.a. Y è **combinazione lineare** di *n* v.a. X_t ($t = 1,2,\ldots n$) **discrete o continue** considerate congiuntamente $(X_1, X_2, \ldots X_n)$ se è data dalla formula

$$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n = \sum_{t=1}^{n} a_t X_t$$

che indica che tutti i valori possibili $y \in S_Y$ della v.a. Y si ottengono dalla formula (o funzione) seguente

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = \sum_{t=1}^{n} a_t x_t \quad \forall (x_1, x_2, \ldots, x_n)$$

al variare dei valori $(x_1, x_2, \ldots, x_n)$ delle *n* v.a. X_t ($t = 1,2,\ldots n$).

CASI PARTICOLARI NOTEVOLI di v.a. combinazioni lineari

Se $\boxed{a_t = 1\ \forall t} \Rightarrow$ *v.a.* **SOMMA** *delle v.a.* X_t :

$$Y = X_1 + X_2 + \cdots + X_n = \boxed{\sum_{t=1}^{n} X_t}$$

Se $\boxed{a_t = 1/n\ \ \forall t} \Rightarrow$ *v.a.* **MEDIA CAMPIONARIA** $\bar{X}_n$ *delle v.a.* X_t :

$$Y = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n = \sum_{t=1}^{n} \frac{1}{n} X_t = \boxed{\frac{\sum_{t=1}^{n} X_t}{n} = \bar{X}_n}$$

VALORE ATTESO $E(Y)$ di una v.a. Y combinazione lineare

di n v.a. X_t ($t = 1,2,\ldots n$) [con $E(X_t) = \mu_t$]: $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n = \sum_{t=1}^{n} a_t X_t$

$E(Y)$ **è la combinazione lineare dei valori attesi** $E(X_t) = \mu_t$

con gli <u>stessi coefficienti</u> a_t ($t = 1,2,\ldots n$), ovvero

$$E(Y) = a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n) = \sum_{t=1}^{n} a_t E(X_t) = \sum_{t=1}^{n} a_t \mu_t$$

Valore atteso della SOMMA e della MEDIA CAMPIONARIA

v.a. **SOMMA** *delle v.a.* X_t : $Y = X_1 + X_2 + \cdots + X_n = \sum_{t=1}^{n} X_t$

$$\boxed{E(Y)} = \sum_{t=1}^{n} E(X_t) = \boxed{\sum_{t=1}^{n} \mu_t}$$, **inoltre** $\boxed{\textbf{se } \mu_t = \mu\ \forall t \textbf{ allora } \boxed{E(Y) = n\mu}}$

v.a. **MEDIA CAMPIONARIA** $\bar{X}_n$ *delle v.a.* X_t : $Y = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n = \sum_{t=1}^{n} \frac{1}{n} X_t = \bar{X}_n$

$$\boxed{E(\bar{X}_n)} = \sum_{t=1}^{n} \frac{1}{n} E(X_t) = \sum_{t=1}^{n} \frac{1}{n} \mu_t = \boxed{\frac{1}{n}\sum_{t=1}^{n} \mu_t}$$, **inoltre** $\boxed{\textbf{se } \mu_t = \mu\ \forall t \textbf{ allora } \boxed{E(\bar{X}_n) = \mu\left(= \mu_t\ \ \forall t\right)}}$

<u>Nota Bene</u>: il risultato $E(\bar{X}_n) = \mu\left(= \mu_t\ \ \forall t\right)$ di cui sopra significa che la media campionaria $\bar{X}_n$ ha la proprietà di **conservare** (o di **non alterare o non "distorcere"**) il valore atteso $\mu_t = \mu$ comune a tutte le n v.a. X_t ($t = 1,2,\ldots n$). Tale proprietà si dice proprietà di **non distorsione** della media campionaria $\bar{X}_n$.

VARIANZA $V(Y)$ di una v.a. Y combinazione lineare

di n v.a. X_t ($t = 1,2,\ldots n$) [con $V(X_t) = \sigma_t^2$]: $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n = \sum_{t=1}^{n} a_t X_t$

(A) Varianza $V(Y)$ di una combinazione lineare Y in un caso particolare notevole:

$\boxed{\textbf{Nel caso particolare } Cov(X_t, X_s) = \sigma_{ts} = 0\ \forall (t,s) \textbf{ si ha che}}$

$V(Y)$ **è la combinazione lineare delle varianze** $V(X_t) = \sigma_t^2$

con i <u>coefficienti al quadrato</u> a_t^2 ($t = 1,2,\ldots n$), ovvero

$$V(Y) = a_1^2V(X_1) + a_2^2V(X_2) + \cdots + a_3^2V(X_n) = \sum_{t=1}^{n} a_t^2 V(X_t) = \sum_{t=1}^{n} a_t^2 \sigma_t^2$$

(B) Varianza $V(Y)$ di una combinazione lineare Y nel caso generale:

Nel caso generale si ha che

$V(Y)$ è data dalla **somma di due addendi**

il primo addendo è dato dalla formula del caso (A) $\sum_{t=1}^{n} a_t^2 \sigma_t^2$

il secondo addendo è dato dalla somma di tutte le covarianze $Cov(X_t, X_s) = \sigma_{ts}$

ciascuna moltiplicata per i coefficienti a_t e a_s $\forall(t,s)$ (con $t \neq s$) delle due variabili considerate, ovvero

$$V(Y) = \sum_{t=1}^{n} a_t^2 V(X_t) + \sum_{\forall(t,s),t\neq s} a_t a_s Cov(X_t, X_s) =$$

$$= \sum_{t=1}^{n} a_t^2 \sigma_t^2 + \sum_{\forall(t,s),t\neq s} a_t a_s \sigma_{ts}$$

Con $n = 2$, ovvero per $Y = a_1 X_1 + a_2 X_2$, si ha

$$\boxed{V(Y)} = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + a_1 a_2 \sigma_{12} + a_2 a_1 \sigma_{21} = \boxed{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1 a_2 \sigma_{12}} \; (\sigma_{12} = \sigma_{21})$$

Varianza della SOMMA e della MEDIA CAMPIONARIA nel caso particolare notevole (A)

v.a. **SOMMA** *delle v.a.* X_t: $Y = X_1 + X_2 + \cdots + X_n = \sum_{t=1}^{n} X_t$

$$\boxed{V(Y)} = \sum_{t=1}^{n} V(X_t) = \boxed{\sum_{t=1}^{n} \sigma_t^2}$$, **inoltre** **se** $\sigma_t^2 = \sigma^2 \; \forall t$ **allora** $\boxed{V(Y) = n\sigma^2}$

v.a. **MEDIA CAMPIONARIA** $\bar{X}_n$ *delle v.a.* X_t: $\bar{X}_n = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n = \sum_{t=1}^{n} \frac{1}{n} X_t$

$$\boxed{V(\bar{X}_n)} = \sum_{t=1}^{n} \left(\frac{1}{n}\right)^2 V(X_t) = \sum_{t=1}^{n} \left(\frac{1}{n}\right)^2 \sigma_t^2 = \boxed{\left(\frac{1}{n}\right)^2 \sum_{t=1}^{n} \sigma_t^2}$$

inoltre **se** $\sigma_t^2 = \sigma^2 \; \forall t$ **allora** $\boxed{V(\bar{X}_n) = \frac{\sigma^2}{n} \left(\leq \sigma^2 = \sigma_t^2 \; \forall t\right)}$

Il risultato $V(\bar{X}_n) = \frac{\sigma^2}{n}\left(\leq \sigma^2 = \sigma_t^2 \; \forall t\right)$ significa che la media campionaria $\bar{X}_n$ ha la proprietà data da:

(1) la sua varianza è più piccola della varianza $\sigma_t^2 = \sigma^2$ comune a tutte le n v.a. X_t ($t = 1, 2, \ldots n$)

(2) la sua varianza tende a zero al crescere di n, ovvero $\lim_{n\to\infty} V(\bar{X}_n) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$.

La combinazione lineare $Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n = \sum_{t=1}^{n} a_t X_t$

di n v.a. $X_t \sim N(\mu_t; \sigma_t^2)$ ($t = 1, 2, \ldots n$)

è gaussiana con $E(Y)$ e $V(Y)$ come già specificato sopra

La combinazione lineare $\bar{X}_n = \sum_{t=1}^{n} \frac{1}{n} X_t$ di n v.a. X_t ($t = 1, 2, \ldots n$)

tutte uguali (con $\sigma_t^2 = \sigma^2$ e $\mu_t = \mu \; \forall t$) e stocasticamente indipendenti

è $\bar{X}_n \underset{n\to\infty}{\sim} N\left(\mu; \frac{\sigma^2}{n}\right)$ (Teorema Centrale del Limite)