VARIABILI ALEATORIE MULTIVARIATE

$$X = [X_1 — X_n] : \Omega \to \mathbb{R}^n$$

VETTORE DI V.A.

MISURA IMMAGINE:

$$P^X(B) = P(\{X \in B\}) \quad , \; B \in \mathcal{B}^n$$

PROBABILITÀ SU $(\mathbb{R}^n, \mathcal{B}^n)$

FUNZIONE DI RIPARTIZIONE (MULTIVARIATA) DI X:

$$F_X(x_1, \dots, x_n) = P(X_1 \le x_1, \dots, X_n \le x_n)$$
$$= P^X(I_{x_1, \dots, x_n})$$

$$I_{x_1, \dots, x_n} = (-\infty, x_1] \times \dots \times (-\infty, x_n]$$

$(x_1 — x_n) \in \mathbb{R}^n$

VARIABILI ALEATORIE MULTIVARIATE

PROPRIETÀ DI F_X

(i) F_X È NON DECRESCENTE

$(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$

$x_i \leq y_i$ PER OGNI i

$$F_X(x_1, \ldots, x_n) \leq F_X(y_1, \ldots, y_n)$$

(ii) F_X CONTINUA "DA SOPRA"

$x_i^{(k)} \downarrow x_i$ $\left(x_i^{(k)} \xrightarrow{k \to +\infty} x_i, \; x_i^{(k)} \text{ NON CRESCENTE}\right)$

$k \to +\infty$

PER $i = 1, \ldots, n$

$$\lim_{k \to +\infty} F_X\left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right) = $$

$$= F_X(x_1, \ldots, x_n)$$

PER OGNI $(x_1, \ldots, x_n)$ E OGNI SUCCESSIONE

VARIABILI ALEATORIE MULTIVARIATE

(iii) $\lim_{x_i \to -\infty} F_X(x_1, \ldots, x_i, \ldots, x_n) = 0$

PER OGNI i, PER OGNI x_j, $j \neq i$

$$\lim_{\substack{x_1 \to +\infty \\ \vdots \\ x_n \to +\infty}} F_X(x_1, \ldots, x_n) = 1$$

(iv) PER OGNI $(a_1, b_1] \times \ldots \times (a_n, b_n] = I$
CON $a_i < b_i$,

$$P(X \in I) = P(a_1 < X_1 \leq b_1, \ldots, a_n < X_n \leq b_n)$$

$$= \Delta^1_{a_1, b_1} \Delta^2_{a_2, b_2} \ldots \Delta^n_{a_n, b_n} F(x_1 \ldots x_n) \geq 0$$

DOVE,

$$\Delta^i_{a,b}\, g(x_1 \ldots x_n) = g(x_1, \ldots, \underset{i}{b}, \ldots x_n) - g(x_1, \ldots, \underset{i}{a}, \ldots x_n)$$

VARIABILI ALEATORIE MULTIVARIATE

(V) $x_i^{(q)} \uparrow x_i$ $\left(x_i^{(q)} \xrightarrow{q \to +\infty} x_i\right.$, $x_i^{(q)}$ NON DECRESCENTE$\left.\right)$

PER $i = 1 \dots n$

$$\lim_{q \to +\infty} F_X(x_1^{(q)} \dots x_n^{(q)}) =$$

$$= P(X_1 < x_1, \dots, X_n < x_n)$$

$$= P^X(\mathring{I}_{x_1 \dots x_n})$$

(VI) $F_X(x_1 \dots x_n) - \lim_{q \to +\infty} F_X(x_1^{(q)}, \dots, x_n^{(q)}) =$

$$= P(X \in \underbrace{FR(I)}_{})$$

$$= I_{x_1 \dots x_n} - \mathring{I}_{x_1 \dots x_n}$$

$= P(X_i \leq x_i$ PER OGNI i,

ESISTE $j : X_j = x_j)$

VARIABILI ALEATORIE MULTIVARIATE

DATA UNA FUNZIONE $F : \mathbb{R}^n \to [0,1]$

CHE SODDISFA (ii), (iii), (iv)
($\exists$ F. DI RIPARTIZIONE MULTIVARIATA)

1) ESISTE UN UNICA PROB P
SU $(\mathbb{R}^n, \mathcal{B}^n)$ TALE CHE

$$P(I_{x_1 \ldots x_n}) = F(x_1 \ldots x_n)$$

2) ESISTE UN VETTORE X DI
V.A. SU UNO SPAZIO $(\Omega, \mathcal{F}, P)$
TALE CHE $F_X = F$

VARIABILI ALEATORIE MULTIVARIATE

DISTRIBUZIONI MARGINALI

$X = [X_1 \text{———} X_m]$ $\quad F_X$ F.D.R.

DISTRIBUZIONE (MARGINALE) DI X_i?

$$\lim_{\substack{x_j \to +\infty \\ j \neq i}} F_X(x_1 \text{———} x_m) = F_{X_i}(x_i) \qquad x_i \in \mathbb{R}$$

IN MANIERA SIMILE LA MARGINALE MULTIVARIATA DI $X_J = [X_j, \; j \in J]$ CON $J \subset \{1 \text{———} m\}$ SI OTTIENE PRENDENDO IL LIMITE

$$\lim_{\substack{x_j \to +\infty \\ j \notin J}} F_X(x_1 \text{———} x_m)$$

VARIABILI ALEATORIE MULTIVARIATE

CASO DISCRETO

X_1 VALORI $x_1^{(1)}, x_2^{(1)}, \ldots$

$\vdots$ FINITI O NUMERABILI

X_n " $x_1^{(n)}, x_2^{(n)}, \ldots$

MASSA DI PROBABILITÀ CONGIUNTA

$$p_{i_1 i_2, \ldots, i_n} = P(X_1 = x_{i_1}^{(1)}, \ldots, X_n = x_{i_n}^{(n)})$$

$$\begin{cases} p_{i_1, i_2, \ldots, i_n} \geq 0 & \text{PER OGNI } i_1 \ldots i_n \\ \sum_{i_1, i_2, \ldots, i_n} p_{i_1, i_2, \ldots, i_n} = 1 \end{cases}$$

DATA $p_{i_1 \ldots i_n}$ CHE SODDISFA LE PROPRIETÀ SOPRA,

$$F_X(x_1 \ldots x_n) = \sum_{x_{i_1}^{(1)} \leq x_1, \ldots, x_{i_n}^{(n)} \leq x_n} p_{i_1, \ldots, i_n}$$

È UNA F.D.R.

MASSA MARGINALE:

$$P(X_i = x_i) = p_i^{(i)} = \sum_{\underbrace{i_1 \ldots i_n}_{\text{TRANNE } i}} p_{i_1 \ldots i \ldots i_n}$$

VARIABILI ALEATORIE MULTIVARIATE

CASO CONTINUO

F_X CONTINUA $\iff$ $P(\cancel{X} \in FR(I)) = 0$

PER OGNI RETTANGOLO

$I = (-\infty, x_1] \times \dots \times (-\infty, x_n]$

SI ASSEGNA USUALMENTE UNA DENSITÀ CONGIUNTA

$f_X : \mathbb{R}^n \to [0, +\infty)$ MISURABILE

TALE CHE

$$F_X(x_1 \dots x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(u_1 \dots u_n)\, du_1 \dots du_n$$

OGNI FUNZIONE $f : \mathbb{R}^n \to \mathbb{R}$ TALE CHE

$$\begin{cases} f \geq 0 \\ \int_{\mathbb{R}^n} f(x_1 \dots x_n)\, dx_1 \dots dx_n \end{cases}$$

È UNA DENSITÀ CONGIUNTA E

$F_X(x_1 \dots x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(u_1 \dots u_n)\, du_1 \dots du_n$ È UNA F.D.R.

VARIABILI ALEATORIE MULTIVARIATE

CASO CONTINUO

COME RICAVARE LA DENSITÀ DALLA F. ~~DE~~ DI R. ?

$$f_X(x_1 \ldots x_n) = \frac{\partial F}{\partial x_1 \partial x_2 \ldots \partial x_n}(x_1 \ldots x_n)$$

VARIABILI ALEATORIE MULTIVARIATE

$$[X_1 \_\_\_ X_n] = X$$

$g : \mathbb{R}^n \longrightarrow \mathbb{R}$ MISURABILE

$$g(X) = g(X_1 \_\_ \_X_n) \in \mathcal{L}^1 \quad \circ \quad g \geq 0$$

$$E\left[g(X_1 \_\_\_, X_n)\right] = \int_{\mathbb{R}^n} g(u_1 \_\_ u_n)\, dP^X(u_1 \_\_ u_n)$$

$$dF_X(u_1 \_\_ u_n)$$

$$= \sum_{i_1, i_2 \_\_ i_n} g(x^{(1)}_{i_1} \,\text{--}\, x^{(n)}_{i_n})\, p_{i_1, i_2 \,\text{--}\, i_n} \quad \text{CASO DISCRETO}$$

$$= \int_{\mathbb{R}^n} g(u_1 \_\_ u_n)\, f_X(u_1 \_\_ u_n)\, du_1 \_\_ du_n \quad \text{CASO CONTINUO}$$

VARIABILI ALEATORIE MULTIVARIATE

COVARIANZA TRA X E Y V.A.

CON $X, Y \in \mathcal{L}^2$

$$COV(X,Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

PROPRIETÀ:

0) $COV(X,Y) = E(XY) - E(X)E(Y)$

1) $COV(X,Y) = COV(Y,X)$ SIMMETRIA

2) $COV(a_1 X_1 + a_2 X_2, Y) = a_1 COV(X_1, Y) + a_2 COV(X_2, Y)$

BILINEARITÀ

3) $COV(X,X) = VAR(X) \geq 0$

4) $COV(X,c) = 0$ $c \in \mathbb{R}$

5) $|COV(X,Y)| \leq \sqrt{VAR[X] \cdot VAR[Y]}$

DISUGUAGLIANZA DI CAUCHY-SCHWARTZ

6) $X_1 \ldots X_n$ V.A. $\in \mathcal{L}^2$

$$VAR[X_1 + \ldots + X_n] = \sum_{i=1}^{n} VAR[X_i] + \sum_{\substack{1 \ldots n \\ i \neq j}} COV(X_i, X_j)$$

VARIABILI ALEATORIE MULTIVARIATE

COEFFICIENTE DI CORRELAZIONE (DI PEARSON)

SE $VAR(X) > 0$, $VAR(Y) > 0$

$$\rho_{X,Y} = CORR(X,Y) = \frac{COV(X,Y)}{\sqrt{VAR(X)\, VAR(Y)}}$$

PROPRIETÀ

1) $\rho_{X,Y} = \rho_{Y,X}$

2) $\rho_{\alpha X + \beta, Y} = \rho_{X,Y}$

3) $\rho_{X,X} = 1$, $\rho_{X,-X} = -1$

4) $|\rho_{X,Y}| \leq 1$ $\quad |\rho_{X,Y}| = 1$ SE E SOLO SE ~~XXX~~ $Y = \alpha X + \beta$

$\alpha > 0$ SE $\rho_{X,Y} = 1$

$\alpha < 0$ SE $\rho_{X,Y} = -1$

VARIABILI ALEATORIE MULTIVARIATE

5) REGRESSIONE DI Y SU X:

IL PROBLEMA

$$\min_{a,b} E\left[\left(Y-(aX+b)\right)^2\right]$$

HA SOLUZIONE $(\hat{a}, \hat{b})$ TALE CHE

$$\hat{a} = \rho_{X,Y} \sqrt{\frac{VAR[Y]}{VAR[X]}}$$

INOLTRE

$$R^2 = 1 - \frac{E\left[\left(Y-(\hat{a}X+\hat{b})\right)^2\right]}{VAR[Y]} =$$

$$= 1 - \frac{\text{VAR. NON SPIEGATA}}{\text{VAR. TOTALE}} = \rho_{X,Y}^2$$

VARIABILI ALEATORIE MULTIVARIATE

$X = [X_1 \text{ —— } X_n] \qquad X_i \in \mathcal{L}^2 \;\; \forall i$

MATRICE DI VARIANZE-COVARIANZE

$$\mathrm{COV}[X] = \begin{bmatrix} \mathrm{cov}(X_1,X_1) & \cdots & \mathrm{cov}(X_1,X_n) \\ & & \\ \mathrm{cov}(X_n,X_1) & & \mathrm{cov}(X_n,X_n) \end{bmatrix}_{n \times n}$$

MATRICE DI CORRELAZIONE

$$\mathrm{CORR}[X] = \begin{bmatrix} \rho_{X_1,X_1} = 1 & \cdots & \rho_{X_1,X_n} \\ \vdots & & \vdots \\ \rho_{X_n,X_1} & & \rho_{X_n,X_n} = 1 \end{bmatrix}_{n \times n}$$

PROPRIETÀ DI COV(X)

0) $\mathrm{cov}(X) = E[XX^T] - E[X]\,E[X]^T$

1) $\mathrm{cov}(X)$ è SIMMETRICA

2) ~~[illegible]~~ LINEARITÀ

$$\underbrace{Y}_{m \times 1} = \underbrace{A}_{m \times n} \underbrace{X}_{n \times 1} + \underbrace{b}_{m \times 1} \qquad E[Y] = A \cdot E[X] + b$$

$$\mathrm{cov}[Y] = A\,\mathrm{cov}[X]\,A^T$$

VARIABILI ALEATORIE MULTIVARIATE

3) $COV[X]_{ii} = VAR(X_i)$

4) $COV(X)$ È DEFINITA NON NEGATIVA (≡ SEMI-DEFINITA POSITIVA)

PER OGNI $x \in \mathbb{R}^n$

$$x^T \, COV[X] \, x \geq 0$$

PROPRIETÀ DI CORR[X]

1) $CORR[X]$ È SIMMETRICA

2) SEMI DEFINITA POSITIVA

PER OGNI $x \in \mathbb{R}^n$

$$x^T \, CORR[X] \, x \geq 0$$

3) $CORR[X]_{ii} = 1$

4) NORMALIZZAZIONE

$$|CORR[X]_{ij}| \leq 1 \quad \forall i, j$$

VARIABILI ALEATORIE MULTIVARIATE

DA COV(X) A CORR(X)

$$CORR[X] = \sqrt{DIAG(COV(X))}^{-1/2} \cdot COV(X) \sqrt{DIAG(COV(X))}^{-1/2}$$

DA CORR(X) A COV(X)

$$COV[X] = DIAG(COV(X))^{1/2} \cdot CORR(X) \cdot DIAG(COV(X))^{1/2}$$

SE COV(X) (O CORR(X)) NON È DI RANGO PIENO ($DET(COV(X)) = 0$)

$\exists\ (\alpha_1 \ldots \alpha_n) \in \mathbb{R}^n \neq (0 \ldots 0)$:

$\alpha^T X$ · OTTENERE

ESISTE UNA V.A. COMBINAZIONE LINEARE DELLE ALTRE

VARIABILI ALEATORIE MULTIVARIATE

$X_1 \ldots X_n$ V.A. SU $\Omega, \mathcal{F}, P$

SONO INDIPENDENTI SE

$\sigma(X_1), \ldots, \sigma(X_n)$ SONO

INDIPENDENTI, CIOÈ

$$P(X_1 \in B_1, \ldots, X_n \in B_n) =$$

$$= P(X_1 \in B_1) \cdot \ldots \cdot P(X_n \in B_n)$$

PER OGNI $B_1, \ldots, B_n \in \mathcal{B}$

VARIABILI ALEATORIE MULTIVARIATE

TEOREMA
LE SEGUENTI SONO EQUIVALENTI:

1) $X_1 \ldots X_n$ SONO INDIPENDENTI

2) $F_X(x_1 \ldots x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)$

PER OGNI $x_1 \ldots x_n$

3) $E[g_1(X_1) \cdot \ldots \cdot g_n(X_n)] =$

$= E[g_1(X_1)] \cdot \ldots \cdot E[g_n(X_n)]$

PER OGNI $g_1 \ldots g_n : \mathbb{R} \to \mathbb{R}$, MISURABILI TALI CHE LE SPERANZE SONO DEFINITE

4) $p_{i_1, \ldots, i_n} = p^{(1)}_{i_1} \cdot p^{(2)}_{i_2} \cdot \ldots \cdot p^{(n)}_{i_n}$

PER OGNI $i_1 \ldots i_n$

CASO DISCRETO

5) $f_X(x_1 \ldots x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)$

PER OGNI $x_1 \ldots x_n$

CASO CONTINUO

VARIABILI ALEATORIE MULTIVARIATE

X, Y v.a. INDIPENDENTI $\in \mathcal{L}^2$

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

$$\Rightarrow \operatorname{cov}(X,Y) = 0$$

$$\Rightarrow \rho_{X,Y} = 0 \qquad \text{SE } \operatorname{VAR}(X) > 0,\ \operatorname{VAR}(Y) > 0$$

NON VALE IL VICEVERSA
(A MENO DEL CASO NORMALE)

ESEMPIO: $X \sim N(0,1)$

$Y = X^2$

$\operatorname{cov}(X,Y) = 0$

VARIABILI ALEATORIE MULTIVARIATE

CALCOLO DI SPERANZE ITERATE (TEOREMA DI FUBINI)

X, Y V.A. INDIPENDENTI

$g : \mathbb{R}^2 \to \mathbb{R}$ MISURABILE

$g(X,Y) \in \mathcal{L}^1$, o $g \geq 0$

ALLORA

$$E[g(X,Y)] = E\left[E[g(x,Y)]_{x=X}\right]$$

IN PARTICOLARE

$$P((X,Y) \in B) = E\left[P((x,Y) \in B)\big|_{x=X}\right]$$

$B \in \mathcal{B}^2$

VARIABILI ALEATORIE MULTIVARIATE

INDIPENDENZA PER UNA FAMIGLIA DI V.A.

$(X_\alpha)_{\alpha \in I}$ V.A. SU $(\Omega, \mathcal{F}, P)$

SONO INDIPENDENTI SE

$(\sigma(X_\alpha))_{\alpha \in I}$ SONO INDIPENDENTI

PER OGNI $\alpha_1 \ldots \alpha_n \in I$ DISTINTI

PER OGNI n

$$P(X_{\alpha_1} \in B_1, \ldots X_{\alpha_n} \in B_n) =$$

$$P(X_{\alpha_1} \in B_1) \cdot \ldots \cdot P(X_{\alpha_n} \in B_n)$$

VARIABILI ALEATORIE MULTIVARIATE

INDIPENDENZA PER INSIEMI DI V.A.

$(\mathcal{A}_\alpha)_{\alpha \in I}$ $\quad \mathcal{A}_\alpha$ = INSIEME DI V.A. IN $(\Omega, \mathcal{F}, P)$

(AD ESEMPIO OGNI $\mathcal{A}_\alpha$ CONTIENE UN VETTORE DI V.A., DI LUNGHEZZA DIVERSA)

GLI INSIEMI $(\mathcal{A}_\alpha)_{\alpha \in I}$ SONO INDIPENDENTI SE

$(\sigma(X, X \in \mathcal{A}_\alpha))_{\alpha \in I}$ SONO INDIPENDENTI

VARIABILI ALEATORIE MULTIVARIATE

$(X_\alpha)_{\alpha \in I}$ V.A. INDIPENDENTI

SE $(I_\beta)_{\beta \in J}$ È UNA PARTIZIONE DI I

$$\mathcal{A}_\beta = \{ X_\alpha \, , \, \alpha \in I_\beta \}$$

ALLORA SONO INSIEMI INDIPENDENTI DI V.A.