Variabili Aleatorie Multivariate

Variabili Aleatorie Multivariate

- $(\Omega, \mathscr{F}, P)$ spazio di probabilità
  $X : \Omega \to \mathbb{R}^n$ vettore di variabili aleatorie

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

  - per ogni $(x_1, \ldots, x_n) \in \mathbb{R}^n$ poniamo

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

    iper-rettangolo chiuso illimitato
  - X è $\mathscr{F}/\mathscr{B}^n$ misurabile se e solo se ogni X_i è $\mathscr{F}/\mathscr{B}$ misurabile

Variabili Aleatorie Multivariate

- Misura immagine di X: per ogni $B \in \mathscr{B}^n$,

$$P_X(B) = P(X \in B) = P(X^{-1}(B))$$

probabilità su $(\mathbb{R}^n, \mathscr{B}^n)$

- Funzione di ripartizione congiunta di X:

$$\begin{aligned} F_X(x_1, \ldots, x_n) &= P_X(I_{x_1,\ldots,x_n}) \\ &= P(X \in I_{x_1,\ldots,x_n}) \\ &= P(X_1 \le x_1, \ldots, X_n \le x_n) \end{aligned}$$

per ogni $(x_1, \ldots, x_n) \in \mathbb{R}^n$

Variabili Aleatorie Multivariate

- Proprietà di F_X
  1. F_X non decrescente: se $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ con $x_i \leq y_i$ per ogni i, e almeno una disuguaglianza stretta,
  $$F_X(x_1, \dots, x_n) \leq F_X(y_1, \dots, y_n)$$
  2. F_X continua "da sopra": per ogni $(x_1, \dots, x_n) \in \mathbb{R}^n$
  $$\lim_{y_1 \downarrow x_1, \dots, y_n \downarrow x_n} F_X(y_1, \dots, y_n) = F_X(x_1, \dots, x_n)$$
  3. limiti: per ogni i fissato e per ogni $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \mathbb{R}^{n-1}$
  $$\lim_{x_i \to -\infty} F_X(x_1, \dots, x_n) = 0$$
  inoltre
  $$\lim_{x_1 \uparrow +\infty, \dots, x_n \uparrow +\infty} F_X(x_1, \dots, x_n) = 1$$

Variabili Aleatorie Multivariate

- Proprietà di F_X

4. per ogni $a_1 < b_1, \dots, a_n < b_n$

$$\Delta^{(1)}_{a_1,b_1}\Delta^{(2)}_{a_2,b_2}\dots\Delta^{(n)}_{a_n,b_n}F_X \geq 0$$

dove

$$\Delta^{(i)}_{a,b}g(x_1,\dots,x_n) = g(x_1,\dots,b,\dots,x_n) - g(x_1,\dots,a,\dots,x_n)$$

(in posizione i-esima)

5. per ogni iper-rettangolo limitato $I = (a_1, b_1] \times \dots \times (a_n, b_n]$ con $a_i < b_i$ per ogni i,

$$\begin{aligned} P(X \in I) =& P(a_1 < X_1 \leq b_1, \dots, a_n < X_n \leq b_n) \\ =& \Delta^{(1)}_{a_1,b_1}\Delta^{(2)}_{a_2,b_2}\dots\Delta^{(n)}_{a_n,b_n}F_X \end{aligned}$$

Variabili Aleatorie Multivariate

- Proprietà di F_X

6. per ogni $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$
\begin{aligned}
& P(X \in \mathrm{Int}(I_{x_1,\ldots,x_n})) = \\
=& P(X_1 < x_1, \ldots, X_n < x_n) \\
=& \lim_{y_1 \uparrow x_1, \ldots, y_n \uparrow x_n} F_X(y_1, \ldots, y_n)
\end{aligned}
$$

7. per ogni $(x_1, \ldots, x_n) \in \mathbb{R}^n$

$$
\begin{aligned}
& P(X \in \mathrm{Fr}(I_{x_1,\ldots,x_n})) = \\
=& P(X_i \leq x_i \text{ per ogni } i, \text{ esiste } j \text{ tale che } X_j = x_j) \\
=& F_X(x_1, \ldots, x_n) - \lim_{y_1 \uparrow x_1, \ldots, y_n \uparrow x_n} F_X(y_1, \ldots, y_n)
\end{aligned}
$$

dove $\mathrm{Fr}(B) = B - \mathrm{Int}(B)$

Variabili Aleatorie Multivariate

- Teorema: per ogni funzione $F : \mathbb{R}^n \to [0,1]$ che soddisfa (2), (3), (4) (nota come funzione di ripartizione congiunta)

  - esiste un unica probabilità P su $(\mathbb{R}^n, \mathscr{B}^n)$ tale che

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

  - esiste un vettore di va X su uno spazio di probabilità $(\Omega, \mathscr{F}, P)$ tale che $F_X = F$

Variabili Aleatorie Multivariate

- Distribuzioni marginali
  - la marginale univariata di X_i si ottiene con

$$F_{X_i}(x) = \lim_{x_j \to +\infty, j \neq i} F_X(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$$

  - se $J \subset \{1, \ldots, n\}$, la marginale congiunta di

$$X_J = [X_i, i \in J]$$

    si ottiene con

$$F_{X_J}(x_i, i \in J) = \lim_{x_j \to +\infty, j \notin J} F_X(x_1, \ldots, x_n)$$

Variabili Aleatorie Multivariate

- Caso discreto:
  - tutte le variabili sono discrete: X_i ha valori $x_1^{(i)}, \dots, x_m^{(i)}, \dots$
  - massa di probabilità congiunta: per ogni $j_1, \dots, j_n$

$$p_{j_1,\dots,j_n} = P(X_1 = x_{j_1}^{(1)}, \dots, X_n = x_{j_n}^{(n)})$$

  - basta assegnare

$$p_{j_1,\dots,j_n} \geq 0 \qquad \text{per ogni } j_1, \dots, j_n$$
$$\sum_{j_1,\dots,j_n} p_{j_1,\dots,j_n} = 1$$

    e la fdr congiunta è

$$F_X(x_1, \dots, x_n) = \sum_{x_{j_1}^{(1)} \leq x_1, \dots, x_{j_n}^{(n)} \leq x_n} p_{j_1,\dots,j_n}$$

Variabili Aleatorie Multivariate

- Proprietà della massa congiunta
  - [masse marginali] si ottengono sommando rispetto agli indici delle variabili rimanenti: la massa di prob. di X_i è

$$P(X_i = x_i) = \sum_{\underbrace{j_1, \ldots, j_n}_{\text{tutte tranne } i}} p_{j_1,\ldots,i,\ldots,j_n}$$

- Esempio. Distribuzione multinomiale: n prove indipendenti, ognuna delle quali può avere k esiti mutualmente esclusivi con prob. $p_1, \ldots, p_k \geq 0$, $p_1 + \ldots + p_k = 1$; $X_i =$ numero di volte in cui si verifica l'esito i

$$P(X_1 = j_1, \ldots, X_k = j_k) = \frac{n!}{j_1! \ldots j_k!} p_1^{j_1} \cdots p_k^{j_k} \qquad \text{se } j_1 + \ldots + j_k = n$$

0 altrimenti

le marginali sono Binomiali(n, p_i)

Variabili Aleatorie Multivariate

- Caso continuo: la fdr congiunta F_X è continua,

$$P(X \in \mathrm{Fr}(I_{x_1,\dots,x_n})) = 0$$

per ogni $(x_1, \dots, x_n) \in \mathbb{R}^n$

  - ogni marginale è continua (non viceversa!)
  - tipicamente, nel caso assolutamente continuo, si assegna una densità congiunta $f_X : \mathbb{R}^n \to \mathbb{R}$, integrabile, tale che

$$f_X(x_1, \dots, x_n) \geq 0 \text{ per ogni } (x_1, \dots, x_n) \in \mathbb{R}^n$$
$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathrm{d}x_1 \dots \mathrm{d}x_n = 1$$

e

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

è una fdr multivariata

Variabili Aleatorie Multivariate

- Proprietà della densità congiunta

  - [densità marginali] si ottengono integrando rispetto alle variabili rimanenti: la densità marginale di X_i è

$$f_{X_i}(x) = \int_{\mathbb{R}^{n-1}} f_X(x_1, \ldots, x, \ldots, x_n) \underbrace{\mathrm{d}x_1, \ldots \mathrm{d}x_n}_{\text{tutte tranne } x}$$

  - [densità da fdr]:

$$f_X(x_1, \ldots, x_n) = \frac{\partial F_X}{\partial x_1 \cdots \partial x_n}(x_1, \ldots, x_n) \qquad \text{q.o.}$$

Variabili Aleatorie Multivariate

- Proprietà della distribuzione congiunta:
  - [integrazione rispetto alla misura immagine] se $g : \mathbb{R}^n \to \mathbb{R}$ è $\mathscr{B}^n/\mathscr{B}$ misurabile tale che $g(X) \in L^1$ o $g \geq 0$

$$E[g(X)] = \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) \mathrm{d}P_X(x_1, \ldots, x_n)$$

  - [caso discreto]

$$E[g(X)] = \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) f_X(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

  - [caso continuo con densità]

$$E[g(X)] = \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) f_X(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Variabili Aleatorie Multivariate

- $(\Omega, \mathscr{F}, P)$ spazio di probabilità
  - $(X_\alpha)_{\alpha \in I}$ va sono indipendenti se sono indipendenti le σ-algebre
    $$(\sigma(X_\alpha))_{\alpha \in I}$$
  - per ogni $\alpha \in I$, $\mathscr{X}_\alpha$ è un insieme di va
    $(\mathscr{X}_\alpha)_{\alpha \in I}$ sono indipendenti se sono indipendenti le σ-algebre
    $$(\sigma(X, X \in \mathscr{X}_\alpha))_{\alpha \in I}$$
  - le va in ogni $\mathscr{X}_\alpha$ potrebbero non essere indipendenti; in particolare ogni $\mathscr{X}_\alpha$ potrebbe essere un vettore di va

Variabili Aleatorie Multivariate

- $(\Omega, \mathscr{F}, P)$ spazio di probabilità, $X = [X_1, \ldots, X_n]^\mathsf{T}$ vettore di va; le seguenti proprietà sono equivalenti

1. $X_1, \ldots, X_n$ sono indipendenti:

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

per ogni $B_1, \ldots, B_n \in \mathscr{B}$

2. per ogni $(x_1, \ldots, x_n) \in \mathbb{R}^n$

$$F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$$

3. per ogni $g_1, \ldots, g_n : \mathbb{R} \to \mathbb{R}$ $\mathscr{B}/\mathscr{B}$ misurabili tali che tutte le variabili sono integrabili,

$$E[g_1(X) \cdots g_n(X_n)] = E[g_1(X)] \cdots E[g_n(X_n)]$$

Variabili Aleatorie Multivariate

- $(\Omega, \mathscr{F}, P)$ spazio di probabilità, $X = [X_1, \dots, X_n]^\mathsf{T}$ vettore di va; le seguenti proprietà sono equivalenti
  1. $X_1, \dots, X_n$ sono indipendenti:
  4. [caso discreto] per ogni $j_1, \dots, j_n$

     $$p_{j_1,\dots,j_n} = p^{(1)}_{j_1} \cdots p^{(n)}_{j_n}$$

     dove $p^{(i)}_j$ è la massa di prob. marginale di X_i
  5. [caso continuo con densità] per ogni $(x_1, \dots, x_n) \in \mathbb{R}^n$

     $$f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

- Esempio. Esponenziale bivariata: per $-1 \le \alpha \le 1$

$$F_{X,Y}(x,y) = (1 - \mathrm{e}^{-x})(1 - \mathrm{e}^{-y})(1 + \alpha \mathrm{e}^{-x-y}), \qquad x \ge 0,\ y \ge 0$$

Variabili Aleatorie Multivariate

- Calcolo di speranze iterate - Teorema di Fubini

  - X, Y va indipendenti,
    $g : \mathbb{R}^2 \to \mathbb{R}$ $\mathscr{B}^2/\mathscr{B}$ misurabile con $g(X, Y) \in L^1$ oppure $g \geq 0$
    allora
    $$E[g(X, Y)] = E[E[g(x, Y)]|_{x=X}]$$
    in particolare, per ogni $B \in \mathscr{B}^2$
    $$P((X, Y) \in B) = E[P((x, Y) \in B)|_{x=X}]$$

- Esercizio. Calcolare la legge di $X + Y$ se X, Y sono indipendenti e X, $Y \sim \text{Exp}(1)$

Variabili Aleatorie Multivariate

- Covarianza: se $X, Y \in L^2$, si definisce

$$\mathrm{COV}[X,Y] = E[(X - E[X])(Y - E[Y])]$$

- Proprietà: se $X, Y, Z \in L^2$
  - $\mathrm{COV}[X,X] = \mathrm{VAR}[X]$; $\mathrm{COV}[X,Y] = 0$ se Y degenere
  - $\mathrm{COV}[X,Y] = E[XY] - E[X]E[Y]$
  - [simmetria] $\mathrm{COV}[X,Y] = \mathrm{COV}[Y,X]$
  - [bilinearità] per ogni $\alpha, \beta \in \mathbb{R}$

$$\mathrm{COV}[\alpha X + \beta Y, Z] = \alpha \mathrm{COV}[X,Z] + \beta \mathrm{COV}[Y,Z]$$

Variabili Aleatorie Multivariate

- Proprietà: se X, Y, $Z \in L^2$
  - [disuguaglianza di Cauchy-Schwartz]

$$|\mathrm{COV}[X, Y]| \leq \sqrt{\mathrm{VAR}[X]\mathrm{VAR}[Y]}$$

  - [indipendenza] $\mathrm{COV}[X, Y] = 0$ se X, Y indipendenti (non viceversa!)
  - [varianza della somma]

$$\mathrm{VAR}[X + Y] = \mathrm{VAR}[X] + \mathrm{VAR}[Y] + 2\mathrm{COV}[X, Y]$$

e quindi

$$\mathrm{VAR}[X + Y] = \mathrm{VAR}[X] + \mathrm{VAR}[Y]$$

se X, Y indipendenti

Variabili Aleatorie Multivariate

- Coefficiente di correlazione (di Pearson): se $X, Y \in L^2$, $\sigma_X^2 \sigma_Y^2 > 0$, si definisce

$$\rho_{X,Y} = \mathrm{CORR}[X,Y] = \frac{\mathrm{COV}[X,Y]}{\sqrt{\mathrm{VAR}[X]\mathrm{VAR}[Y]}}$$

- Proprietà: se $X, Y \in L^2$, $\sigma_X^2 \sigma_Y^2 > 0$
  - [simmetria] $\rho_{X,Y} = \rho_{Y,X}$
  - [invarianza di scala e locazione] per ogni $\alpha > 0$, $\beta \in \mathbb{R}$, $\rho_{\alpha X+\beta,Y} = \rho_{X,Y}$
  - [normalizzazione] $|\rho_{X,Y}| \leq 1$
    e $|\rho_{X,Y}| = 1 \iff Y = \alpha X + \beta$ con $\alpha, \beta \in \mathbb{R}$
    $(\mathrm{sgn}(\alpha) = \mathrm{sgn}(\rho_{X,Y}))$

Variabili Aleatorie Multivariate

- Proprietà: se X, $Y \in L^2$, $\sigma_X^2 \sigma_Y^2 > 0$
  - [indipendenza] $\rho_{X,Y} = 0$ se X, Y indipendenti
  - [regressione lineare di Y su X] il problema

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

ha soluzione $\widehat{a}$, $\widehat{b}$ tale che

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

inoltre

$$R^2 = 1 - \frac{E[(Y - (\widehat{a}X + \widehat{b}))^2]}{\mathrm{VAR}[Y]} = \rho_{X,Y}^2$$

Variabili Aleatorie Multivariate

- $X = [X_1, \ldots, X_n]^{\mathsf{T}}$ vettore di va con $X_i \in L^2$ per ogni i; la matrice varianze-covarianze è

$$\mathrm{COV}[X]_{i,j} = \mathrm{COV}[X_i, X_j], \qquad 1 \le i, j \le n$$

- Proprietà di $\mathrm{COV}[X]$
  - [simmetria] $\mathrm{COV}[X]$ è simmetrica
  - [linearità] per ogni A matrice $m \times n$ e $b \in \mathbb{R}^m$

    $$Y = AX + b$$

    vettore di va in $\mathbb{R}^m$ con

    $$E[Y] = AE[X] + b, \qquad \mathrm{COV}[Y] = A\mathrm{COV}[X]A^{\mathsf{T}}$$

Variabili Aleatorie Multivariate

- Proprietà di $\mathrm{COV}[X]$
  - $\mathrm{COV}[X]_{ii} = \mathrm{VAR}[X_i]$
  - $\mathrm{COV}[X]$ è semi definita positiva: per ogni $z \in \mathbb{R}^n$
  $$z^{\mathsf{T}}\mathrm{COV}[X]z \geq 0$$
  - se $\mathrm{COV}[X]$ non è di rango pieno, esiste una combinazione lineare di $X_1, \ldots, X_n$ degenere (relazione lineare tra le variabili)
- Proprietà simili valgono per la matrice di correlazione
$$\mathrm{CORR}[X]_{i,j} = \rho_{X_i,X_j}, \qquad 1 \leq i, j \leq n$$

Variabili Aleatorie Multivariate

- Normale multivariata: $X = [X_1, \ldots, X_n]$ è normale se per ogni $z \in \mathbb{R}^n$
$$z_1 X_1 + \ldots + z_n X_n$$
è normale univariata (una va degenere è normale di varianza 0)
  - tutte le marginali (univariate e multivariate) sono normali
  - ogni trasformazione lineare è normale (univariata o multivariata)
  - posto $\mu = E[X]$ e $\Sigma = \mathrm{COV}[X]$, se Σ è di rango pieno allora la densità congiunta di X è
$$f_X(x) = (2\pi)^{-n/2} \det(\Sigma)^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)^\mathsf{T}\Sigma^{-1}(x-\mu)\right\}$$