Parte 4
Variabili casuali multiple

Variabile casuale multipla

Interesse esteso alle n ($n \geq 2$) v.c. $X_1, X_2, \cdots, X_n$

$(X_1, X_2, \cdots, X_n)$ v.c. **multipla** (o multivariata) ad n **componenti** (o dimensioni)

Vettore di n variabili casuali (**vettore casuale**)

Le singole v.c. componenti $X_1, X_2, \cdots, X_n$ sono le v.c **marginali**

Dato un **esperimento casuale** (come il campionamento casuale)

La v.c multipla $(X_1, X_2, \cdots, X_n)$ è una *funzione* (non necessariamente *biunivoca*) che associa agli **eventi elementari** (elementi dello spazio campionario Ω) una $n-$pla ordinata di **numeri reali** $(x_1, x_2, \cdots, x_n)$

Traduzione degli eventi elementari in **vettori numerici**

$x_1, x_2, \cdots, x_n$ valori, rispettivamente, di $X_1, X_2, \cdots, X_n$

Dominio	Funzione	Codominio
Spazio campionario Ω	$\longrightarrow$ Variabile casuale **multipla**	Spazio vettoriale $\mathbb{R}^n$

$(x_1, x_2, \cdots, x_n)$ **realizzazione** numerica della v.c. multipla $(X_1, X_2, \cdots, X_n)$

L'**insieme** delle possibili **realizzazioni** numeriche di una v.c. multipla
è lo **spazio campionario** *generato* dalla variabile casuale (sottoinsieme di $\mathbb{R}^n$)

Supporto della v.c. multipla (*immagine* della funzione)

Le **realizzazioni** della v.c. multipla sono i *nuovi* **eventi elementari** di interesse
(anche denominate *determinazioni* o *vettori numerici* della v.c. multipla)

$(X_1, X_2, \cdots, X_n)$ v.c. multipla **discreta** o **continua** se, rispettivamente,
le *v.c. marginali* sono *tutte* v.c **discrete** o *tutte* v.c. **continue**

La nozione di v.c. multipla è alla base del concetto di **campione casuale**
(una *speciale* v.c. multipla) su cui si fondano le *metodologie inferenziali* (ndr)

$(X_1, X_2, \cdots, X_n)$ v.c. **multipla** ad n **componenti** $(n \geq 2)$

$(X_1, X_2, \cdots, X_n)$ v.c. multipla **discreta**
$X_1, X_2, \cdots, X_n$ v.c. *marginali* (tutte **discrete**) componenti della v.c. multipla
$f(x_1), f(x_2), \cdots, f(x_n)$
Funzioni di **probabilità** *marginali* delle v.c. $X_1, X_2, \cdots, X_n$
$f(x_1, x_2, \cdots, x_n)$
Funzione di **probabilità** *congiunta* della v.c. multipla $(X_1, X_2, \cdots, X_n)$

$$f(x_1, x_2, \cdots, x_n) = P\big((X_1 = x_1) \cap (X_2 = x_2) \cap \cdots \cap (X_n = x_n)\big)$$

$P(\cdot)$ è la **funzione di probabilità** secondo la **definizione assiomatica**

$$f(x_1, x_2, \cdots, x_n) \geq 0 \quad \forall (x_1, x_2, \cdots, x_n) \qquad \sum_{x_1, x_2, \cdots, x_n} f(x_1, x_2, \cdots, x_n) = 1$$

$(X_1, X_2, \cdots, X_n)$ v.c. multipla **continua**
$X_1, X_2, \cdots, X_n$ v.c. *marginali* (tutte **continue**) componenti della v.c. multipla
$f(x_1), f(x_2), \cdots, f(x_n)$
Funzioni di **densità** *marginali* delle v.c. $X_1, X_2, \cdots, X_n$
$f(x_1, x_2, \cdots, x_n)$
Funzione di **densità** *congiunta* della v.c. multipla $(X_1, X_2, \cdots, X_n)$

$f(x_1, x_2, \cdots, x_n) \geq 0 \quad \forall (x_1, x_2, \cdots, x_n)$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, \cdots, x_n)\, dx_1\, dx_2 \cdots dx_n = 1$

$f(x_1, x_2, \cdots, x_n)$ consente, mediante opportuna **integrazione** *multipla*,
di calcolare la probabilità di prefissate *regioni* del supporto

$(X_1, X_2, \cdots, X_n)$ v.c. **multipla** (*discreta* o *continua*) ad n **componenti** $(n \geq 2)$

Distribuzione di probabilità *congiunta* della v.c. multipla $(X_1, X_2, \cdots, X_n)$

Fornita dalla funzione **congiunta** $f(x_1, x_2, \cdots, x_n)$

Funzione di **probabilità** (caso *discreto*) o funzione di **densità** (caso *continuo*)

Indipendenza delle n v.c. *marginali* $X_1, X_2, \cdots, X_n$

$X_1, X_2, \cdots, X_n$ sono v.c. **indipendenti se se solo se**

$f(x_1, x_2, \cdots, x_n) = f(x_1) \times f(x_2) \times \cdots \times f(x_n)$	$\forall (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$

La funzione **congiunta** (di *probabilità* o di *densità*) si ottiene **moltiplicando** tra di esse le funzioni **marginali** (di *probabilità* o di *densità*)

Condizione **necessaria e sufficiente** per l'indipendenza delle v.c. $X_1, X_2, \cdots, X_n$

Esperimento casuale composto da n *sottoprove* **indipendenti**

Lanci ripetuti di monete/dadi *non truccati*, altri giochi di sorte ripetuti in *condizioni simili*, simbolica estrazione da un'urna *con riposizione*, **campionamento casuale**

Ogni sottoprova *genera* una variabile casuale

Le n v.c. $X_1, X_2, \cdots, X_n$ *generate* dalle n sottoprove (indipendenti) dell'esperimento casuale sono v.c. **indipendenti**

Non vale necessariamente il viceversa

v.c. *indipendenti* non sono necessariamente generate da esperimenti casuali con sottoprove *indipendenti* (ndr)

Campionamento casuale

Alla base dei metodi di inferenza statistica

La v.c. multipla $(X_1, X_2, \cdots, X_n)$ *generata* dalle n *estrazioni campionarie* è a componenti **indipendenti** (le n v.c. marginali)

$X_1, X_2, \cdots, X_n$ v.c. **indipendenti**	**Implicazioni**	
Le n v.c. sono anche indipendenti *a gruppi* (ottenuti in tutti i modi possibili)	$n-1$	Indipendenti ad $n-1$ ad $n-1$
	$n-2$	Indipendenti ad $n-2$ ad $n-2$
	$\vdots$	
	3	Indipendenti a tre a tre
	2	Indipendenti a due a due

La funzione **congiunta** (di probabilità o di densità) di *ogni gruppo* di v.c.
è il **prodotto** delle corrispondenti funzioni **marginali** (di probabilità o di densità)

Ad esempio $f(x_1, x_2) = f(x_1) \times f(x_2)$

L'indipendenza di un gruppo di v.c. **non implica** l'indipendenza *globale* delle n v.c.

Indipendenza di **funzioni** di variabili casuali

$X_1, X_2, \cdots, X_n$ v.c. **indipendenti**	
Siano $g_1(X_1), g_2(X_2), \cdots, g_n(X_n)$ n v.c. **funzioni** delle v.c. $X_1, X_2, \cdots, X_n$	
Anche $g_1(X_1), g_2(X_2), \cdots, g_n(X_n)$ sono v.c. **indipendenti**	
Sono indipendenti, ad esempio (ndr):	
Le n v.c. **scarto**	$(X_1 - \mu_{X_1}), (X_2 - \mu_{X_2}), \cdots, (X_n - \mu_{X_n})$
Le n v.c. **standardizzate**	$\left(\frac{X_1 - \mu_{X_1}}{\sigma_{X_1}}\right), \left(\frac{X_2 - \mu_{X_2}}{\sigma_{X_2}}\right), \cdots, \left(\frac{X_n - \mu_{X_n}}{\sigma_{X_n}}\right)$

Media del **prodotto** (momento misto) di variabili casuali **indipendenti**

<table>
<tr><td colspan="3">$X_1, X_2, \cdots, X_n$ v.c. indipendenti</td></tr>
<tr><td colspan="3">$E(X_1 X_2 \cdots X_n) = E(X_1)E(X_2)\cdots E(X_n)$</td></tr>
<tr><td colspan="3">La media del prodotto (momento misto) di n v.c. indipendenti
è uguale al prodotto delle n medie (momenti primi)</td></tr>
<tr><td colspan="2">$E(X_1 X_2 \cdots X_n) \neq E(X_1)E(X_2)\cdots E(X_n)$</td><td>Se le n v.c non sono indipendenti</td></tr>
<tr><td rowspan="2">$n = 2$</td><td colspan="2">$E(X_1 X_2) = E(X_1)E(X_2)$</td></tr>
<tr><td>$E\left(\frac{X_1}{X_2}\right) = E(X_1)E\left(\frac{1}{X_2}\right) \neq \frac{E(X_1)}{E(X_2)}$</td><td>La media del rapporto tra due v.c.
non è il rapporto delle due medie</td></tr>
</table>

Media della **somma** di variabili casuali (e di una loro **combinazione lineare**)

$X_1, X_2, \cdots, X_n$ v.c. **non necessariamente** indipendenti

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

$n = 2$

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

La **media** della **somma** di n v.c. (che siano indipendenti o meno)
è uguale alla **somma** delle n **medie**

$$g(X_1, X_2, \cdots, X_n) = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

Combinazione lineare (è una v.c.)
delle n v.c. $X_1, X_2, \cdots, X_n$

$$E(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n) = a_1 E(X_1) + a_2 E(X_2) + \cdots + a_n E(X_n)$$

La **media** di una **combinazione lineare** di n v.c. (che siano indipendenti o meno)
è uguale alla **combinazione lineare** delle n **medie**

$a_i = \frac{1}{n} \quad \forall i$

$$g(X_1, X_2, \cdots, X_n) = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

Media aritmetica delle n v.c.
(in seguito **media campionaria**)

$$E\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) = \frac{1}{n}\big(E(X_1) + E(X_2) + \cdots + E(X_n)\big)$$

La **media** della **media aritmetica** di n v.c. (che siano indipendenti o meno)
è uguale alla **media aritmetica** delle n **medie**

Varianza della **somma** di variabili casuali **indipendenti** (e di una loro **combinazione lineare**)

$X_1, X_2, \cdots, X_n$ v.c. **indipendenti**

$$V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$$

$n = 2$

$$V(X_1 + X_2) = V(X_1) + V(X_2)$$

La **varianza** della **somma** di n v.c. **indipendenti** è uguale alla **somma** delle n **varianze**

$$V(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1{}^2V(X_1) + a_2{}^2V(X_2) + \cdots + a_n{}^2V(X_n)$$

La **varianza** di una **combinazione lineare** di n v.c. **indipendenti** è uguale alla **combinazione lineare** delle n **varianze**, con pesi pari ai **quadrati** dei coefficienti

$$V\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) = \frac{1}{n^2}\big(V(X_1) + V(X_2) + \cdots + V(X_n)\big)$$

La **varianza** della **media aritmetica** di n v.c. **indipendenti** è uguale alla **somma** delle n **varianze**, rapportata al **quadrato** di n