6 Le variabili casuali congiunte

6.1 Definizioni e proprietà

Nei casi trattati, gli esiti di un esperimento erano considerati realizzazioni di una singola variabile casuale. Tuttavia, in alcune situazioni può essere necessario o desiderabile ottenere esiti simultanei da più variabili casuali.
È possibile estendere il concetto di variabile casuale, di funzione di ripartizione, di funzione di densità di una variabile casuale al caso $n-$dimensionale.

Definizione: Chiamiamo variabile casuale $n-$dimensionale una funzione $X = (X_1, \dots, X_n)$: $\Omega \to \mathbb{R}^n$ tale che $\forall (r_1, \dots, r_n) \in \mathbb{R}^n$

$$\{w \in \Omega : X_1(w) \leq r_1, \dots, X_n(w) \leq r_n\} \text{ è un evento.}$$

Quindi una variabile casuale $n-$dimensionale è una $n-$upla di variabili casuali, le quali associano un numero ad un risultato.

Definizione: Chiamiamo funzione di ripartizione congiunta di $X = (X_1, ..., X_n)$, la funzione

$$F_{X_1,...,X_n} : \mathbb{R}^n \to [0, 1]$$

tale che $\forall x = (x_1, ..., x_n)$

$$F_{X_1,...,X_n}(x_1, ..., x_n) = P[X_1 \leq x_1, ..., X_n \leq x_n].$$

Come nel caso unidimensionale, per definire la funzione di densità congiunta bisogna distinguere il caso discreto dal caso continuo.

Definizione: Chiamiamo funzione di densità discreta congiunta di $X = (X_1, ..., X_n)$, variabile casuale discreta $n-$dimensionale, la funzione:

$$f_{X_1,...,X_n}(x_1, ..., x_n) = P[X_1 = x_1, ..., X_n = x_n].$$

Proprietà

- $f_{X_1,...,X_n}(x_1, ..., x_n) \geq 0 \quad \forall (x_1, ..., x_n)$
- $\sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} f_{X_1,...,X_n}(x_1, ..., x_n) = 1$
- $\forall A \subset \mathbb{R}^n \; P[(x_1, ..., x_n) \in A] = \sum_{A} f_{X_1,...,X_n}(x_1, ..., x_n)$

Definizione: La v.c. $n-$dimensionale $X = (X_1, ..., X_n)$ è detta continua se e soltanto se esiste una funzione

$$f_{X_1,...,X_n}(x_1, ..., x_n) : \mathbb{R}^n \to \mathbb{R}$$

integrabile $\forall x_k$ tra $-\infty$ e $+\infty$, tale che:

$$F_{X_1,...,X_n}(x_1, ..., x_n) =$$

$$\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_{X_1,...,X_n}(t_1, ..., t_n) dt_1 \cdots dt_n$$

$\forall x = (x_1, ..., x_n) \in \mathbb{R}^n$ ed $f_{X_1,...,X_n}(x_1, ..., x_n)$ è detta funzione di densità continua congiunta.

Proprietà

- $f_{X_1,...,X_n}(x_1, ..., x_n) \geq 0 \quad \forall (x_1, ..., x_n)$
- $\int_{\mathbb{R}^n} f_{X_1,...,X_n}(x_1, ..., x_n) = 1$
- $\forall A \subset \mathbb{R}^n$ regione regolare

$$P[(x_1, ..., x_n) \in A] = \int_A f_{X_1,...,X_n}(x_1, ..., x_n) dx_1 \cdots dx_n$$

Nel caso $n = 2$ elenchiamo le proprietà della funzione di ripartizione congiunta relativa alla variabile casuale bidimensionale (X, Y).

Proprietà di $F_{X,Y}(x,y)$

- $\lim_{x \to -\infty} F_{X,Y}(x,y) = 0 \quad \forall y$
- $\lim_{y \to -\infty} F_{X,Y}(x,y) = 0 \quad \forall x$
- $\lim_{\substack{x \to +\infty \\ y \to +\infty}} F_{X,Y}(x,y) = 1$
- se $x_1 < x_2$ e $y_1 < y_2$ allora

  $P[x_1 < X \leq x_2, y_1 < Y \leq y_2] = F_{X,Y}(x_2,y_2)+$

  $+F_{X,Y}(x_1,y_1) - F_{X,Y}(x_1,y_2) - F_{X,Y}(x_2,y_1) \geq 0$
- $\lim_{h \to 0^+} F_{X,Y}(x+h,y) = \lim_{h \to 0^+} F_{X,Y}(x,y+h) =$
  $= F_{X,Y}(x,y)$ (continuità a destra)

Dalla funzione di densità congiunta $f_{X,Y}(x,y)$ delle v. c. congiunte X, Y è possibile ricavare le funzioni di densità $f_X(\cdot)$ di X e $f_Y(\cdot)$ di Y, dette funzioni di densità marginali.

Definizione: Le funzioni di densità marginali di X e di Y sono:

$$f_X(x_k) = \sum_j f_{X,Y}(x_k, y_j)\,, \quad f_Y(y_k) = \sum_j f_{X,Y}(x_j, y_k)$$

se X, Y sono congiunte discrete,

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)dy\,, \;\; f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)dx$$

se X, Y sono congiunte continue.

Nota bene: Dalla funzione di densità congiunta è possibile sempre ricavare le funzioni di densità marginali, ma non vale il viceversa.

Ricordando la definizione di probabilità condizionata

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) > 0,$$

se ora A è l'evento definito da $X = x$ e B è l'evento definito da $Y = y$, allora

$$P[X = x|Y = y] = \frac{P[X = x, Y = y]}{P[Y = y]}, \; P[Y = y] > 0$$

ed è possibile dare la definizione formale di funzione di densità condizionata.

Definizione: Date X, Y variabili casuali congiunte (discrete o continue) con funzione di densità congiunta $f_{X,Y}$, chiamiamo funzione di densità condizionata di X, dato $Y = y$, la funzione:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad f_Y(y) > 0.$$

Analogamente si definisce la funzione di densità condizionata di Y, dato $X = x$, la funzione:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \qquad f_X(x) > 0.$$

Nota bene:

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) \cdot f_Y(y),$$
$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \cdot f_X(x).$$

6.2 Indipendenza

Definizione: Data $X = (X_1, ..., X_n)$ variabile casuale $n-$dimensionale (discreta o continua) con funzione di densità congiunta $f_{X_1,...,X_n}$, $X_1, ..., X_n$ sono variabili casuali indipendenti se e solo se

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

cioè la funzione di densità congiunta si scrive come prodotto delle funzioni di densità marginali.
Nel caso bidimensionale si ha quindi che X, Y sono indipendenti se e solo se $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$.
Osservazione
Se $X_1, ..., X_n$ sono variabili casuali indipendenti e se $g_1, ..., g_n$ sono n funzioni tali che $Y_k = g_k(X_k)$,

$k = 1, .., n$, sono variabili casuali, allora $Y_1, ..., Y_n$ sono indipendenti.

6.3 Estensione del concetto di valore atteso

Vogliamo ora estendere il concetto di valore atteso di una variabile casuale unidimensionale al caso $n-$dimensionale.
Definizione: Sia $X = (X_1, ..., X_n)$ una variabile casuale $n-$dimensionale con funzione di densità congiunta $f_{X_1,...,X_n}(x_1, ..., x_n)$ e sia $g : \mathbb{R}^n \to \mathbb{R}$ una funzione definita su $X_1, ..., X_n$, cioè $g = g(X_1, ..., X_n)$ (a sua volta è una variabile casuale).
Definiamo valore atteso di g la quantità:

$$E[g] = \sum g(x_1, ..., x_n) f_{X_1,...,X_n}(x_1, ..., x_n)$$

se X è discreta,

$$E[g] = \int_{\mathbb{R}^n} g(x_1, ..., x_n) f_{X_1,...,X_n}(\cdot) dt_1 \cdots dt_n$$

se X è continua.

Nel caso $n = 2$, se $g(X, Y) = (X - \mu_X)(Y - \mu_Y)$
$\Rightarrow E[g] = E[(X - \mu_X)(Y - \mu_Y)]$

Definizione: Date le variabili casuali X e Y, definiamo covarianza di X e Y la quantità:

$$\sigma_{X,Y} = \text{cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)]$$

Proprietà della covarianza

1. $\text{cov}[X, Y] = E[XY] - \mu_X \mu_Y$
2. $\text{cov}[aX, bY] = ab\,\text{cov}[X, Y] \qquad \forall a, b \in \mathbb{R}$
3. $\text{cov}[X, Y] = \text{cov}[Y, X]$
4. $\text{cov}[X, X] = \text{var}[X]$
5. $\text{cov}[X + Y, Z] = \text{cov}[X, Z] + \text{cov}[Y, Z]$

La covarianza tra due variabili casuali descrive la possibilità che tra le due variabili possa esistere una relazione di tipo lineare. Se la covarianza è nulla, tra X ed Y è possibile che esista una relazione di tipo non lineare.
La covarianza dipende dall'unità di misura utilizzata per misurare X e Y; perciò si preferisce usare un altro indice, indipendente dall'unità di misura.

Definizione: Date le variabili casuali X e Y con covarianza $\text{cov}[X, Y]$ e deviazioni standard rispettivamente σ_X e σ_Y, definiamo coefficiente di correlazione di X e Y il numero:

$$\rho_{X,Y} = \frac{\text{cov}[X,Y]}{\sigma_X \sigma_Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

Proprietà di $\rho_{X,Y}$

1. $|\rho_{X,Y}| \leq 1$
2. $\rho_{X,Y} = 0$ se $\text{cov}[X,Y] = 0$
3. Se $Y = mX + q$ (dipendenza lineare), allora
$$\begin{cases} \rho_{X,Y} = 1 & \text{se } m > 0\,, \\ \rho_{X,Y} = -1 & \text{se } m < 0\,. \end{cases}$$
4. $\forall a, b \in \mathbb{R}^+ \quad \rho_{aX,bY} = \rho_{X,Y}$

Proposizione 1. Siano X, Y variabili casuali indipendenti e g_1, g_2 funzioni tali che $g_1 = g_1(X)$ e $g_2 = g_2(Y)$ siano variabili casuali. Allora vale

$$E[g_1 \cdot g_2] = E[g_1] \cdot E[g_2]$$

Proposizione 2. Se X, Y sono variabili casuali indipendenti allora $\text{cov}[X,Y] = 0$

Osservazione
$\text{cov}[X,Y] = 0 \;\; \not\Rightarrow \;\; X, Y$ indipendenti
cioè se $\text{cov}[X,Y] = 0$

NON VALE

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

Definizione: Due variabili casuali X, Y si dicono non correlate se e solo se $\text{cov}[X,Y] = 0$.
Quindi:

$$\text{INDIPENDENZA} \underset{\not\Leftarrow}{\Rightarrow} \text{NON CORRELAZIONE}$$

Nota bene: Esiste un solo caso in cui $\text{cov}[X,Y] = 0$ $\Rightarrow$ indipendenza. Ciò avviene quando la funzione di densità congiunta di X ed Y è gaussiana.

6.4 Combinazioni lineari di variabili casuali

Date n variabili casuali $X_1, ..., X_n$ si consideri:

$$1) \qquad g(X_1, ..., X_n) = X_1 + \cdots + X_n$$

si può dimostrare che:

- $E[g] = \sum_{i=1}^{n} E[X_i]$

- $\text{var}[g] = \sum_{i=1}^{n} \text{var}[X_i] + 2\sum_{i}\sum_{\substack{j \\ i \neq j}} \text{cov}[X_i, X_j]$

$$2) \qquad g(X_1, ..., X_n) = a_1 X_1 + \cdots + a_n X_n$$

con $a_i \in \mathbb{R}$, per $i = 1, ..., n$.

- $E[g] = \sum_{i=1}^{n} a_i E[X_i]$

- $\text{var}[g] = \sum_{i=1}^{n} a_i^2 \,\text{var}[X_i] + 2\sum_{i}\sum_{\substack{j \\ i \neq j}} a_i a_j \,\text{cov}[X_i, X_j]$

$$3) \qquad g(X_1, ..., X_n) = X_1 + \cdots + X_n$$

con $\text{cov}[X_i, X_j] = 0, i \neq j$, cioè X_i a due a due non correlate

- $E[g] = \sum_{i=1}^{n} E[X_i]$

- $\text{var}[g] = \text{var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \text{var}[X_i]$

$$4) \qquad g(X_1, ..., X_n) = X_1 + \cdots + X_n$$

con X_i a due a due non correlate ed identicamente distribuite cioè $\text{cov}[X_i, X_j] = 0$ per $i \neq j$; $\mu_{X_i} = \mu$ e $\sigma^2_{X_i} = \sigma^2$, $\forall i = 1, ..., n$

- $E[g] = n\mu$
- $\text{var}[g] = n\sigma^2$

5) $$g(X_1, ..., X_n) = \overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$$

con X_i a due a due non correlate ed identicamente distribuite con media μ e varianza σ^2, detta media campionaria

- $E[g] = E[\overline{X}_n] = \mu$
- $\text{var}[g] = \text{var}[\overline{X}_n] = \dfrac{\sigma^2}{n}$

<u>Osservazione</u>
Nei casi 4) e 5) l'ipotesi di non correlazione a due a due può essere sostituita con l'ipotesi più forte di indipendenza (indipendenza $\underset{\not\Leftarrow}{\Rightarrow}$ non correlazione).

In tal caso le variabili casuali $X_1, ..., X_n$ si dicono indipendenti ed identicamente distribuite, abbreviate con la dicitura i.i.d.
Relativamente al caso 5) è possibile provare che se X_i sono i.i.d. con funzione di densità normale $N(\mu, \sigma^2)$, allora

$$\overline{X}_n \text{ è NORMALE } N(\mu, \frac{\sigma^2}{n})$$