8.2 Distribuzioni di probabilità

VARIABILI ALEATORIE DISCRETE

Nell'ambito del calcolo delle probabilità si introduce un nuovo concetto di "variabile", quello di variabile aleatoria, che è fondamentale in svariate applicazioni.

Intuitivamente una variabile aleatoria è una variabile i cui valori sono numeri reali determinati dall'esito di un esperimento aleatorio. Per esempio sono varabili aleatorie:

- T che rappresenta il numero di testa uscite nel lancio di 100 monete,
- H che rappresenta l'altezza di una persona scelta a caso nella popolazione,
- N che rappresenta in numero di autovetture che giungono in un giorno ad un casello autostradale.

Una variabile aleatoria si può anche interpretare come una funzione, per esempio se consideriamo la variabile aleatoria X "somma dei numeri ottenuti nel lancio dei due dadi", essa si può interpretare come la funzione che associa ad ogni possibile esito del lancio la somma dei due numeri ottenuti.

Questa osservazione consente di dare una definizione più formale di variabile aleatoria:

Definizione Si chiama *variabile aleatoria*, una funzione che associa ad ogni possibile esito di un esperimento aleatorio un numero reale.

Se Ω è lo spazio campione di un esperimento e X è una variabile aleatoria legata all'esperimento, allora $X: \Omega \longrightarrow \mathbb{R}$

Esempio

Consideriamo l'esperimento che consiste nel lancio di tre monete eque, indichiamo con X il numero di "testa" che si ottiene. Rappresentiamo la variabile aleatoria X.

X rappresenta il numero di teste uscite lanciando tre monete, quindi X può assumere solo i valori $\{0,1,2,3\}$, lo spazio campione $\Omega = \{(CCC),(CCT),(CTC),(CTT),(TCC),(TCT),(TTC),(TTT)\}$, possiamo rappresentare X come segue:

$$\begin{array}{rcl} X: \quad \Omega & \longrightarrow & \{0,1,2,3\} \\ (CCC) & \longmapsto & 0 \\ (TCC),(CTC),(CCT) & \longmapsto & 1 \\ (TTC),(CTT),(TCT) & \longmapsto & 2 \\ (TTT) & \longmapsto & 3 \end{array}$$

Una variabile aleatoria X, che assume un numero finito o numerabile di valori si dice *discreta*.

L'evento: "X assume il valore x_i" (che rappresenta un evento essendo la controimmagine di x_i e quindi un sottoinsieme di Ω), si rappresenta con il simbolo $X = x_i$.

Possiamo anche calcolare la probabilità di $X = x_i$, essa è uguale alla somma delle probabilità degli eventi elementari la cui immagine tramite X è uguale ad x_i.

Esempio: Considerando l'esempio precedente, osserviamo che:

- La probabilità che $X = 0$ è uguale alla probabilità dell'evento elementare CCC, quindi vale $\frac{1}{8}$.
- La probabilità che $X = 1$ è uguale alla somma delle probabilità degli eventi elementari TCC, CTC, CCT, quindi vale $\frac{3}{8}$.
- La probabilità che $X = 2$ è uguale alla somma delle probabilità degli eventi elementari TTC, CTT, TCT, quindi vale $\frac{3}{8}$.

- La probabilità che $X = 3$ è uguale alla probabilità dell'evento elementare TTT, quindi vale $\frac{1}{8}$.

Formalmente scriviamo:

$$p(X = 0) = \frac{1}{8}, \quad p(X = 1) = \frac{3}{8}, \quad p(X = 2) = \frac{3}{8}, \quad p(X = 3) = \frac{1}{8}$$

Supponendo ancora che X sia una variabile aleatoria discreta che assume i valori $x_1, x_2, \dots x_n$, possiamo associare a ciascuno degli eventi, $X = x_1, \dots, X = x_n$, la rispettiva probabilità, si definisce così una nuova funzione:

Definizione: Sia X una variabile aleatoria discreta che assume i valori $x_1, x_2, \dots, x_n$ rispettive probabilità $p_1, p_2, \dots, p_n$; si chiama *distribuzione di probabilità (o densità),* della variabile aleatoria X, la funzione che associa a ciascun x_i la rispettiva probabilità p_i.

La distribuzione di probabilità di una variabile aleatoria discreta si rappresenta tramite una tabella:

x_i	x_1	x_2	…	x_n
$p(X = x_i)$	p_1	p_2	…	p_n

Poiché gli eventi $X = x_i$, sono disgiunti si ha che $p_1 + p_2 + \dots + p_n = 1$.

Esempio Riprendendo l'esempio precedente si ha che

x_i	0	1	2	3
$p(X = x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

C'è una certa similitudine tra il concetto di distribuzione di probabilità e quello di distribuzione di frequenze relative in statistica, la similitudine continua anche nella definizione di concetti di *media, varianza, scarto quadratico medio.*

Definizioni: Sia X una variabile aleatoria discreta che assume i valori $x_1, x_2, \dots, x_n$, con probabilità rispettive $p_1, p_2, \dots, p_n$.

a) Si chiama *media (o valore atteso),* della variabile aleatoria X, e si indica con $E(X)$, o con la lettera μ, il numero:
$$\mu = E(X) = x_1 p_1 + \dots + x_n p_n$$
b) Si definisce *varianza* di X, e si indica con il simbolo $V(X)$ o con il simbolo σ^2, il numero così definito
$$\sigma^2 = V(X) = \left(x_1 - E(X)\right)^2 \cdot p_1 + \dots + \left(x_n - E(X)\right)^2 \cdot p_n$$
c) Si definisce *scarto quadratico medio (o deviazione standard)* di X, e si indica con il simbolo σ, la radice quadrata della varianza, $\sigma = \sqrt{\sigma^2}$.

Esempio Calcola media, varianza, scarto quadratico medio dell'esempio precedentemente studiato. (quante teste lanciando tre monete).

Esercizio Da un urna contenente 4 palline nere e 5 rosse, si estraggono consecutivamente senza reimmissione 4 palline. Determina la distribuzione di probabilità della variabile casuale $X =$"numero delle palline nere uscite".