Concetti chiave e regole

Variabili aleatorie discrete

Una variabile aleatoria discreta è una grandezza che può assumere valori diversi x_i al verificarsi di determinati eventi, incompatibili e complementari, che hanno ciascuno una probabilità stabilita p_i.
Di ogni variabile aleatoria discreta X si può costruire la **funzione distribuzione di probabilità** $p = p(x_i)$ associando a ciascun valore x_i la propria probabilità p_i.
Sommando le probabilità $p(x_i)$ dalla prima fino alla i-esima, si ottiene la probabilità che la variabile X assuma valori minori o uguali a x_i. La funzione che si ottiene al variare di i fra 1 e n si chiama **funzione di ripartizione**; essa è quindi definita dalla relazione:

$$F(x_i) = p(X \leq x_i) = \sum_{k=1}^{i} p_k$$

Valori di sintesi

I valori di sintesi di una variabile aleatoria X sono il **valore atteso** (o **speranza matematica**) e la **varianza**, che sono così definiti:

- valore atteso simboli usati: μ $E(X)$ espressione: $\sum_{i=1}^{n} x_i p(x_i)$
- varianza simboli usati: σ^2 $V(X)$ espressione: $\sum_{i=1}^{n} (x_i - \mu)^2 \cdot p_i$

Per il calcolo della varianza si può usare anche la formula $V(X) = E(X^2) - [E(X)]^2$.

Spesso si usa anche un altro indice, detto **scarto quadratico medio** o **deviazione standard**, che viene indicato con il simbolo σ ed è la radice quadrata della varianza.

Distribuzioni di probabilità discrete

Alcune distribuzioni discrete di probabilità rivestono particolare importanza perchè sono in grado di descrivere fenomeni con caratteristiche particolari.
Fra queste abbiamo visto la **distribuzione binomiale** nella quale la variabile X conta il numero di successi di un evento A nell'ambito di un esperimento aleatorio; se p è la probabilità che A si verifichi e $q = 1 - p$ quella che non si verifichi, allora la probabilità che su n ripetizioni dell'esperimento ci siano x successi è

$$p(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$$

La funzione distribuzione di probabilità binomiale ha quindi espressione:

$$f(x) = \begin{cases} \binom{n}{x} \cdot p^x \cdot q^{n-x} & x = 0,\ 1,\ 2,\ \ldots,\ n \\ 0 & \text{altrove} \end{cases}$$

Per questa distribuzione si ha che $E(X) = np$ $\quad V(X) = npq = np(1 - p)$.

Le variabili aleatorie continue

Una variabile aleatoria è continua se può assumere tutti i valori che appartengono ad un certo intervallo.
Per le variabili aleatorie continue non ha senso calcolare la probabilità che X assuma un particolare valore x_0 perché $p(x_0) = 0$; si deve parlare invece di **funzione densità di probabilità** $f(x)$ che è definita come segue:

- è una funzione non negativa: $f(x) \geq 0 \quad \forall x \in R$
- insieme con l'asse delle ascisse racchiude un'area uguale a 1: $\int_{-\infty}^{+\infty} f(x)dx = 1$

Con queste condizioni, la probabilità che X assuma valori compresi fra a e b è:

$$p(a < X \leq b) = \int_a^b f(x)dx$$

Il valore atteso e la varianza si calcolano rispettivamente con le formule:

- $\mu = E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)dx$
- $\sigma^2 = V(X) = \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx - \mu^2$

La distribuzione normale

Fra tutte le funzioni densità di probabilità, quella normale è fra le più importanti perché approssima in modo soddisfacente tutte le situazioni in cui la maggior parte dei valori di X si concentra attorno ad uno particolare.
La **funzione densità di probabilità normale**, detta anche **gaussiana**, ha espressione

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

essendo μ la media della distribuzione e σ la deviazione standard.

La probabilità che la variabile X assuma valori minori o uguali a k è data dalla misura dell'area delimitata della curva gaussiana per $x \leq k$.