2. Variabili aleatorie e valore atteso

Indice

Variabili aleatorie

Le quantità di interesse determinate dall'esito di un esperimento casuale sono dette **variabili aleatorie**. Poiché il valore di una variabile aleatoria dipende dall'esito dell'esperimento, è possibile associare una probabilità ai suoi possibili valori.

La **funzione di ripartizione** F (o **CDF**, Cumulative Distribution Function) di una variabile aleatoria X è definita, per ogni numero reale i, come:

$$F(i) = P(X \leq i)$$

Quindi, $F(i)$ esprime la **probabilità** che X assuma un valore **minore o uguale** a i. Si utilizza la notazione $X \sim F$ per indicare che F è la funzione di ripartizione di X.

Variabili discrete e continue

Le variabili aleatorie che possono assumere un numero **finito** o **numerabile** di valori sono dette **discrete**, mentre quelle che possono assumere un insieme continuo (**infinito non numerabile**) di valori sono dette **continue**.

Se X è una variabile aleatoria **discreta**, la sua **funzione di massa di probabilità** (o **PMF**, Probability Mass Function) è definita come:

$$p(a) = P(X = a)$$

Nota: Poiché X deve assumere uno dei valori $x_1, x_2, \ldots$, la funzione di massa di probabilità deve soddisfare la seguente condizione:

$$\sum_i p(x_i) = 1$$

Se X è una variabile aleatoria **continua**, definita su tutto $\mathbb{R}$, per ogni insieme misurabile B di numeri reali vale:

$$P(X \in B) = \int_B f(x)\,dx$$

La funzione f è detta **funzione di densità di probabilità** (o **PDF**, Probability Density Function).

Nota: Analogamente alla funzione di massa di probabilità, la funzione di densità deve soddisfare la condizione:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Quando si conosce la **funzione di massa** di una variabile aleatoria discreta, la **funzione di densità** di una variabile continua, o la **funzione di ripartizione** di una variabile aleatoria qualsiasi, si hanno informazioni sufficienti per calcolare la probabilità di qualsiasi evento dipendente da quella variabile. In tal caso, si dice che è nota la **distribuzione** della variabile considerata. Inoltre, dire che due variabili aleatorie X e Y hanno la stessa distribuzione equivale a dire che le loro funzioni sopra citate sono identiche.

Coppie e vettori di variabili aleatorie

Siano X e Y due variabili aleatorie relative allo stesso esperimento. La **funzione di ripartizione congiunta** di X e Y è definita come:

$$F(a,b) = P(X \leq a, Y \leq b)$$

Nota: La virgola nell'argomento di P rappresenta l'**intersezione** tra i due eventi.

Se X e Y sono variabili aleatorie discrete, la loro **funzione di massa congiunta** è:

$$p(x_i, y_j) = P(X = x_i, Y = y_j)$$

Due variabili aleatorie sono **congiuntamente continue** se esiste una funzione $f(x,y)$ non negativa tale che, per ogni sottoinsieme misurabile C del piano cartesiano, vale:

$$P\big((X,Y) \in C\big) = \iint_{(x,y)\in C} f(x,y)\,dx\,dy$$

La funzione $f(x,y)$ è detta **densità congiunta** di X e Y.

Spiegazione: Se X rappresenta l'altezza di una persona e Y il suo peso, la densità congiunta $f_{X,Y}(a,b)$ indica quanto è probabile che una persona scelta a caso abbia altezza a e peso b. Maggiore è il valore di f in un punto (a,b), maggiore è la probabilità di osservare quella particolare combinazione di altezza e peso. Integrando f su un'area specifica, otteniamo la probabilità che altezza e peso cadano in quell'intervallo. Ad esempio, integrando su altezze tra 160 cm e 170 cm e pesi tra 60 kg e 70 kg, si ottiene la probabilità che una persona abbia altezza e peso in quegli intervalli.

Due variabili aleatorie relative allo stesso esperimento si dicono **indipendenti** se, per ogni coppia di insiemi misurabili A e B, vale:

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$$

In caso contrario, si dicono **dipendenti**. Questa definizione è equivalente alla richiesta che, per ogni coppia di numeri reali a e b:

$$P(X \leq a, Y \leq b) = P(X \leq a) \cdot P(Y \leq b)$$

ovvero, che la **funzione di ripartizione congiunta** sia uguale al prodotto delle **funzioni di ripartizione marginali**:

$$F(a, b) = F_X(a) \cdot F_Y(b)$$

Se le variabili considerate sono discrete, l'indipendenza è equivalente a richiedere che la **funzione di massa congiunta** sia uguale al prodotto delle **funzioni di massa marginali**:

$$p(a, b) = p_X(a) \cdot p_Y(b)$$

Valore atteso

Se X è una variabile aleatoria **discreta** che può assumere i valori $x_1, x_2, \ldots$, il **valore atteso** di X, indicato con $E[X]$, è (se esiste) dato da:

$$E[X] = \sum_i x_i \cdot P(X = x_i)$$

$E[X]$ è anche detto **media** o **aspettazione** di X, e rappresenta la media ponderata dei valori possibili di X, usando come pesi le probabilità associate.

Nota: $E[X]$ non è necessariamente uguale a uno dei valori possibili di X.

Se X è una variabile aleatoria **continua** con densità $f(x)$, il valore atteso di X è (se esiste):

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f(x) \, dx$$

In pratica, il concetto di valore atteso è analogo a quello fisico del baricentro, ma applicato alle distribuzioni di probabilità.

Per $n = 1, 2, \ldots$, la quantità $E[X^n]$ è detta, se esiste, **momento di ordine** n della variabile X. In particolare:

$$E[X^n] = \begin{cases} \sum_i x_i^n \cdot p(x_i), & \text{se } X \text{ è discreta} \\ \int_{-\infty}^{+\infty} x^n \cdot f(x)\,dx, & \text{se } X \text{ è continua} \end{cases}$$

Di conseguenza, $E[X]$ è il momento di primo ordine di X. Inoltre, vale la proprietà:

$$E[X+Y] = E[X] + E[Y]$$

Varianza

Data una variabile aleatoria X con media $\mu = E[X]$, la **varianza** di X è, se esiste, definita come:

$$\mathrm{Var}(X) = E\big[(X-\mu)^2\big]$$

Esiste una formula alternativa per calcolare la varianza:

$$\mathrm{Var}(X) = E[X^2] - [E[X]]^2$$

La quantità $\sigma = \sqrt{\mathrm{Var}(X)}$ è detta **deviazione standard** della variabile aleatoria X.

Covarianza e varianza della somma di variabili aleatorie

Date le variabili aleatorie X e Y con medie μ_X e μ_Y, la **covarianza** di X e Y è, se esiste, la quantità:

$$\mathrm{Cov}(X,Y) = E\big[(X-\mu_X)(Y-\mu_Y)\big]$$

Anche in questo caso, esiste una formula alternativa per calcolare la covarianza:

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]\cdot E[Y]$$

Inoltre, se X e Y sono variabili aleatorie indipendenti, allora:

$$E[XY] = E[X]\cdot E[Y]$$

e quindi:

$$\mathrm{Cov}(X,Y) = 0$$

Data la somma di due variabili aleatorie X e Y, vale:

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y)$$

Se X e Y sono indipendenti, allora:

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Inoltre, date le variabili aleatorie $X_1, X_2, \ldots, X_n$:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{1\le i<j\le n} \mathrm{Cov}(X_i, X_j)$$

Se $X_1, X_2, \ldots, X_n$ sono tutte indipendenti, allora:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Casi particolari. Di seguito sono mostrati alcuni casi particolari:

- $\mathrm{Var}(aX + bY) = a^2 \,\mathrm{Var}(X) + b^2 \,\mathrm{Var}(Y) + 2ab \,\mathrm{Cov}(X, Y)$
- $\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2 \,\mathrm{Cov}(X, Y)$
- $\mathrm{Var}(a + bX) = b^2 \,\mathrm{Var}(X)$

In generare, si può mostrare che un valore positivo di $Cov(X, Y)$ indica che X ed Y tendenzialmente assumono valori grandi o piccoli contemporaneamente. La "forza" della relazione tra X ed Y è misurata più propriamente dal **coefficiente di correlazione lineare**, un numero puro (senza unità di misura) che tiene conto anche delle deviazioni standard di X ed Y. Esso si indica con $Corr(X, Y)$ ed è definito come

$$Corr(X, Y) := \frac{Cov(X, Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

Si può inoltre dimostrare che questa quantità è sempre compresa tra -1 e $+1$.

Legge debole dei grandi numeri

Disuguaglianza di Markov. Se X è una variabile aleatoria che non è mai negativa, allora per ogni $a > 0$

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Disuguaglianza di Chebyshev. Se X è una variabile aleatoria con media μ e varianza σ^2, allora per ogni $r > 0$

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

Legge debole dei grandi numeri. Sia $X_1, \ldots, X_n$ una successione di variabili aleatorie i.i.d (indipendenti e identicamente distribuite), tutte con media $E[X_i] = \mu$. Allora, per ogni $\varepsilon > 0$,

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right) \to 0$$

quando $n \to \infty$.