

Variabili aleatorie

Daniele Falanga

1 Variabili aleatorie

Una variabile aleatoria è una funzione che associa ad ogni esperimento un valore numerico. Possono essere di due tipi:

- Continue: che associa una serie infinita di valori all'interno di un intervallo
- Discrete: che associa un numero finito e numerabile di valori

Definizione 1.1. *La funzione di ripartizione F di una variabile aleatoria X , è definita, per ogni numero reale x , tramite:*

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

La funzione $F(x)$ esprime la probabilità che la variabile aleatoria X assuma un valore minore o uguale ad x

La funzione di ripartizione è una descrizione completa della distribuzione di probabilità di una variabile casuale. In sostanza, la funzione di ripartizione di una variabile aleatoria ci dice come è ripartita in un intervallo la probabilità della variabile.

2 Variabili aleatorie discrete e continue

Definizione 2.1. *Se X è una variabile aleatoria discreta, la sua funzione di massa di probabilità si definisce in questo modo*

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

Fornisce il valore della probabilità quando la variabile X assume un determinato valore nell'intervallo. Infatti i suoi valori possibili sono:

$$\begin{aligned} p(x_i) &> 0, \quad i = 1, 2, \dots \\ p(x) &= 0, \quad \text{tutti gli altri valori di } x \end{aligned}$$

Infatti deve valere necessariamente questa relazione:

$$\sum_{i=1}^{\infty} = 1$$

Per una variabile aleatoria discreta, la funzione di ripartizione può essere riscritta in funzione della funzione di massa:

$$F(a) = \sum_{x \leq a} p(x)$$

Caso differente per le variabili aleatorie **continue**.

Definizione 2.2. Si definisce densità della variabile aleatoria continua X

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

La sua densità deve rispettare:

$$1 = P(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f(x)dx$$

La relazione che lega la funzione di ripartizione F alla densità f è la seguente:

$$F(a) = P(X \in (-\infty, a]) = \int_{-\infty}^a f(x)dx$$

se si derivano entrambi i membri si ottiene che la densità è la derivata della funzione di ripartizione:

$$\frac{d}{da} F(a) = f(a)$$

3 Coppie e vettori di variabili aleatorie

Non sempre una sola variabile aleatoria va bene, quindi si tende ad usarne 2 o più.

Definizione 3.1. Siano X e Y due variabili aleatorie di uno stesso esperimento. Si dice funzione di ripartizione congiunta delle variabili X ed Y la seguente funzione:

$$F(x, y) := P(X \leq x, Y \leq y)$$

3.1 Distribuzione congiunta per variabili aleatorie discrete

La funzione di massa di due variabili aleatorie si definisce:

Definizione 3.2. Se X e Y sono variabili aleatorie discrete che assumono valori x_1, x_2, \dots e y_1, y_2, \dots rispettivamente, la funzione:

$$p(x_i, y_j) := P(X = x_i, Y = y_j), \quad i = 1, 2, \dots, \quad j = 1, 2, \dots$$

è la funzione di massa di probabilità congiunta

Le funzioni di massa di probabilità individuali, si possono ricavare da quella congiunta, se per esempio volessi calcolare la funzione di massa di X al evento x_i mi basta fissare x e unire gli eventi y .

$$P_X(x_i) = \sum_j p(x_i, y_j)$$

In modo del tutto analogo per y .

Non vale il viceversa !!! Non posso calcolare la funzione di massa congiunta da quelle individuali.

3.2 Variabili aleatorie indipendenti

Due variabili sono indipendenti se ogni evento relativo alla prima variabile è indipendente dalla seconda

Definizione 3.3. *Due variabili aleatorie che riguardano lo stesso esperimento casuale si dicono indipendenti se, per ogni coppia di insiemi di numeri reali A e B è soddisfatta l'equazione:*

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

La funzione di ripartizione congiunta è il prodotto delle marginali:

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

Anche la funzione di massa congiunta è il prodotto delle marginali:

$$p(x, y) = p_X(x)p_Y(y)$$

3.3 Distribuzioni condizionali

Definizione 3.4. *Siano X e Y due variabili aleatorie discrete con funzione di massa congiunta $p(\cdot, \cdot)$. Si dice funzione di massa di probabilità condizionata di X dato Y , e si indica con $p_{X|Y}(\cdot|\cdot)$, la funzione a due variabili così definita:*

$$\begin{aligned} p_{X|Y}(x|y) &:= P(X = x|Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{p(x, y)}{p_Y(y)} \end{aligned}$$

4 Valore Atteso

È uno dei concetti fondamentali della teoria della probabilità.

Definizione 4.1. *Sia X una variabile aleatoria discreta che può assumere i valori x_1, x_2, \dots ; il valore atteso di X , se esiste (deve convergere la sommatoria) è definito come segue:*

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

In pratica, la media pesata che i dei possibili valori di X , dove il peso è la probabilità che X assuma proprio quel valore.

Nota 4.1. *Il concetto di valore atteso è analogo a quello del baricentro in fisica*

4.1 Proprietà del valore atteso

Data una variabile aleatoria X e una sua distribuzione, in molti casi è utile voler calcolare il valore atteso di una sua funzione $g(X)$, anch'essa una variabile aleatoria.

Proposizione 4.1. *Se X è una variabile aleatoria discreta con funzione di massa di probabilità p , allora, per ogni funzione g :*

$$E[g(x)] = \sum_x g(x)p(x)$$

4.1.1 Valore atteso di una somma di variabili aleatorie

La versione in due dimensioni della funzione di variabile aleatoria:

$$E[g(X, Y)] = \sum_x \sum_y g(x, y)p(x, y)$$

Si può applicare anche a $g(X, Y) = X + Y$:

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

5 Varianza

Data una variabile aleatoria X , è utile, oltre calcolare il la sua media ($E[X]$), vedere quanto i dati si discostano dalla media, ed è possibile farlo tramite la varianza.

Definizione 5.1. *Sia X una variabile aleatoria con media μ . la varianza di X , si denota con $Var(X)$, è la quantità:*

$$Var(X) := E[(X - \mu)^2]$$

Come formula alternativa, ed anche più utile nel nostro caso:

$$Var(X) := E[X^2] - E[X]^2$$

Una utile identità per la covarianza:

$$Var(aX + b) = a^2 Var(X)$$

Si definisce **deviazione standard** della variabile X :

$$\sqrt{Var(X)}$$

La varianza, in termini fisici, corrisponde al momento di inerzia.

6 La covarianza e la varianza della somma di variabili aleatorie

Da vedere se le formule a pagina 135 servono

Attenzione !!!

Per la varianza, la media delle somme, non coincide con la somma delle medie.

La covarianza indica la relazione lineare tra due variabili aleatorie, in altre parole, indica se le variazioni di una variabile sono associate a quelle dell' altra variabile.

Se la covarianza è positiva vuol dire che sono direttamente proporzionali, altrimenti sono inversamente proporzionali

Definizione 6.1 (Covarianza). *Siano assegnate due variabili aleatorie X ed Y di media μ_x e μ_y rispettivamente. La loro covarianza, che si indica con $Cov(X, Y)$ è la quantità:*

$$Cov(X, Y) := E[(X - \mu_x)(Y - \mu_y)]$$

Come formula alternativa:

$$Cov(X, Y) := E[XY] - E[X]E[Y]$$

Teorema 6.1. *Se X e Y sono variabili aleatorie indipendenti, allora:*

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

La loro covarianza, essendo indipendenti sarà uguale a 0

Come detto prima, la covarianza di due variabili aleatorie, se è positiva, vuol dire che le variabili sono direttamente proporzionali, viceversa, inversamente proporzionali.

La forza di correlazione tra le due variabili, è misurata più propriamente dal **coefficiente di correlazione** definito come segue:

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

L'indice di correlazione è **sempre** compreso tra -1 e 1.

7 Legge debole dei grandi numeri

Prima di esporre la legge debole dei grandi numeri vanno preliminarmente esposte le due proposizioni seguenti:

Proposizione 7.1 (Disuguaglianza di Markov). *Se X è una variabile aleatoria non negativa, allora per ogni $a > 0$:*

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

La dimostrazione non credo si debba fare, ma la metto comunque

Dimostrazione.

$$\begin{aligned} E[X] &:= \int_0^\infty x f(x) dx \\ &= \int_0^a f(x) dx + \int_a^\infty f(x) dx \\ &\geq \int_a^\infty x f(x) dx \quad \text{Perchè il primo addendo è positivo} \\ &\geq \int_a^\infty a f(x) dx \quad \text{Perchè } x \text{ è maggiore di } a \text{ nella regione di integrazione} \\ &= a \int_a^\infty f(x) dx \\ &= a P(X \geq a) \end{aligned}$$

□

In altre parole, la probabilità che una variabile casuale non negativa superi un certo valore a è limitata superiormente dal rapporto tra il valore atteso della variabile casuale e il valore a

Proposizione 7.2 (Disuguaglianza di Chebyshev). *Se X è una variabile aleatoria con media μ e varianza σ^2 , allora per ogni $r > 0$:*

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

oppure in modo analogo:

$$P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}$$

Dimostrazione. Gli eventi $\{|X - \mu| \geq r\}$ e $\{(X - \mu)^2 \geq r^2\}$ coincidono e si dicono equiprobabili.

Visto che $(X - \mu)^2$ è una variabile aleatoria non negativa, possono applicare Markov: con $a = r^2$

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

□

La disuguaglianza di Chebyshev fornisce una misura di quanto una variabile casuale si discosta dalla sua media. Indica che, a prescindere dalla forma precisa della distribuzione, la maggior parte dei valori della variabile casuale si trova entro un certo numero di deviazioni standard dalla media. Più grande è il valore di k , più stretto diventa l'intervallo in cui la maggior parte dei valori si concentra.

Legge debole dei grandi numeri

Teorema 7.1. *Sia X_1, X_2, \dots una successione di variabili aleatorie indipendenti e identicamente distribuite, tutte con media $E[X_i] = \mu$. Allora per ogni ϵ :*

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \epsilon\right) \rightarrow 0 \quad n \rightarrow \infty$$

Dimostrazione. Supponendo che le X abbiano varianza σ^2 finita:

$$E\left[\frac{X_1 + \dots + X_n}{n}\right] = \mu \quad \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

La seconda si trova in questo modo:

$$\begin{aligned} \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right) &= \frac{1}{n^2} \text{Var}(X_1 + \dots + X_n) \\ &= \frac{\text{Var}(X_1) + \dots + \text{Var}(X_n)}{n^2} \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

Segue, applicando a Chebyshev la variabile aleatoria $(X_1, \dots, X_n)/n$ che:

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

□