05BQX Metodi Matematici per l'ingegneria 2011-2012

Marina Santacroce

Dipartimento di Scienze Matematiche, Politecnico di Torino

7. Complementi sulle leggi congiunte, valori attesi e teoremi limite.

Riferimenti: S.Ross *Calcolo delle probabilità* Cap.7-8

Outline

Valore atteso di una funzione di due variabili aleatorie

X e Y sono due variabili aleatorie e $g(x, y)$ una funzione reale. Allora $g(X, Y)$ è una variabile aleatoria reale, il cui valor medio si calcola conoscendo la legge congiunta di X e Y.

Prop.: Siano X e Y due variabili aleatorie discrete con densità congiunta p (continue con densità congiunta f) e g una funzione reale, allora, se esiste,

$$\boxed{\mathbb{E}(g(X,Y)) = \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} g(x,y)p(x,y)} \quad \left(\mathbb{E}(g(X,Y)) = \int_{\mathbb{R}^2} g(x,y)f(x,y)dxdy\right)$$

Esercizio

- Consideriamo la coppia di v.a. X, Y uniforme sul quadrato $[0,1]\times[0,1]$. Calcoliamo $\mathbb{E}(|X-Y|)$ che rappresenta la distanza media tra due punti scelti a caso in modo indipendente in $[0,1]$.

Valore atteso della somma di variabili aleatorie

Prendiamo il caso particolare

- $g(x,y) = x + y \quad \Rightarrow \quad \boxed{\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)}$

Dimostro nel caso continuo:

$$\begin{aligned}
\mathbb{E}(X+Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x+y)\, f(x,y)\, dxdy \\
&= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f(x,y)\, dy\, dx + \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f(x,y)\, dx\, dy \\
&= \int_{-\infty}^{+\infty} x\, f_X(x)\, dx + \int_{-\infty}^{+\infty} y\, f_Y(y)\, dy = \mathbb{E}(X) + \mathbb{E}(Y)
\end{aligned}$$

- Si verifichi il caso discreto.
- Sia $Z = X + Y$, $f_Z = f_X \star f_Y$. Si verifichi

$$\int z\, f_Z(z)\, dz = \mathbb{E}(X) + \mathbb{E}(Y).$$

Consideriamo n variabili aleatorie $X_1, \ldots, X_n$ allora (per induzione)

$$\mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)$$

- $X \sim \text{Bin}(n, p)$, $\quad X = \sum_{i:1}^{n} X_i$, $\quad$ con $X_i \sim \text{Be}(p)$,

$$\Rightarrow \qquad \mathbb{E}(X) = \sum_{i:1}^{n} \mathbb{E}(X_i) = np$$

- $X \sim \Gamma(n, \lambda)$, $\quad X = \sum_{i:1}^{n} X_i$, $\quad$ con $X_i \sim \text{Exp}(\lambda)$,

$$\Rightarrow \qquad \mathbb{E}(X) = \sum_{i:1}^{n} \mathbb{E}(X_i) = \frac{n}{\lambda}$$

• $X_1, \ldots, X_n$ sono n variabili aleatorie indipendenti identicamente distribuite (in breve i.i.d.) con valore atteso μ. $X_1, \ldots, X_n$ si chiama campione casuale.

$\overline{X} = \frac{1}{n}\sum_{i:1}^{n} X_i$ rappresenta la media campionaria e

$$\mathbb{E}(\overline{X}) = \frac{1}{n}\left(\mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)\right) = \mu$$

Indipendenza e valori attesi

Prop. Se le variabili X e Y sono *indipendenti*, e g, h sono due funzioni reali, allora

$$\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$$

- $$\begin{aligned}\mathbb{E}(g(X)h(Y)) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)h(y)\ f(x,y)\ dxdy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)h(y)\ f_X(x)f_Y(y)\ dxdy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)f_X(x)\ h(y)f_Y(y)\ dxdy \\ &= \int_{-\infty}^{+\infty} g(x)f_X(x)\ dx \int_{-\infty}^{+\infty} h(y)f_Y(y)\ dy \\ &= \mathbb{E}(g(X))\mathbb{E}(h(Y))\end{aligned}$$

- Si dimostri il caso discreto.

Covarianza

Def. Date X e Y variabili aleatorie reali, la *covarianza* di X e Y è

$$\boxed{\mathrm{COV}(X,Y)=\mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)}$$

Se la $\mathrm{COV}(X,Y)=0$ le variabili X e Y si dicono *non correlate* o *scorrelate*.

Esempio: $X\sim \mathrm{U}([-1,1])$ e sia $Y=X^2 \;\Rightarrow\; \mathrm{COV}(X,Y)=0$.

Prop.: Se X e Y sono indipendenti, allora sono non correlate, i.e. $\mathrm{COV}(X,Y)=0$.

Infatti,

$$\mathrm{COV}(X,Y)=\mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]=\mathbb{E}[(X-\mathbb{E}(X))]\mathbb{E}[(Y-\mathbb{E}(Y))]=0$$

Oss. *Come si vede dall'esempio, in generale non vale il viceversa*!!!

In generale variabili aleatorie *non correlate* sono *molto lontane* dall'essere *indipendenti*!

Fanno *eccezione*:

- X e Y normale bivariata;
- X e Y Bernoulli.

In questi due casi la non correlazione implica l'indipendenza.

Proprietà della Covarianza e Varianza della somma

Verificare le seguenti uguaglianze.

a) $\text{COV}(X, Y) = \text{COV}(Y, X)$
b) $\text{COV}(X, X) = Var(X)$
c) $\text{COV}(X, a) = 0$, per a costante;
d) $\text{COV}(aX, Y) = a\text{COV}(X, Y)$, $\quad \text{COV}(X, bY) = b\text{COV}(X, Y)$, per due costanti a, b;
e) $\text{COV}(aX + bZ + c, Y) = a\text{COV}(X, Y) + b\text{COV}(Z, Y)$ per tre costanti a, b, c.

La covarianza è *bilineare*, i.e. scelte le costanti a_i, b_j si ha

$$\text{COV}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \text{COV}(X_i, Y_j)$$

Da questa formula si ricava immediatamente

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i) + \sum_{i \neq j} \text{COV}(X_i, X_j)$$

Esempio: *Problema delle corrispondenze (cappelli, tango, etc.)* rivisitato....

Studiamo i casi notevoli:

- $X_1, \ldots X_n$ sono *non correlate* $\Rightarrow \mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$
- $X_1, \ldots X_n$ sono *indipendenti* $\Rightarrow \mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$;
- $X_1, \ldots, X_n$ sono i.i.d. con varianza σ^2. Allora

$$\mathrm{Var}(\sum_{i:1}^{n} X_i) = n\sigma^2$$

e la varianza della media campionaria $\overline{X} = \frac{1}{n}\sum_{i:1}^{n} X_i$ è

$$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}.$$

Oss: Se $X_1, \ldots X_n$ sono i.i.d. e $X_i \sim \mathrm{N}(\mu, \sigma^2) \Rightarrow \overline{X} \sim \mathrm{N}(\mu, \frac{\sigma^2}{n})$

- S^2 indica la varianza campionaria *corretta*,

$$\mathbb{E}(S^2) = \mathbb{E}\left(\sum_{i:1}^{n} \frac{\left(X_i - \overline{X}\right)^2}{n-1}\right) = \sigma^2.$$

Correlazione

Def. Date X e Y variabili aleatorie reali, il *coefficiente di correlazione* tra X e Y è definito, se $Var(X)Var(Y) \neq 0$, da

$$\boxed{\rho(X,Y) = \frac{\mathrm{COV}(X,Y)}{\sqrt{Var(X)Var(Y)}}}$$

Oss. Si osservi che $\rho(X,Y) = 0 \iff X$ e Y sono *non correlate*.

Prop. $\boxed{-1 \leq \rho(X,Y) \leq +1.}$

Siano $\sigma_X = \sqrt{Var(X)}$ e $\sigma_Y = \sqrt{Var(Y)}$.

$$0 \leq Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = Var\left(\frac{X}{\sigma_X}\right) + Var\left(\frac{Y}{\sigma_Y}\right) + 2\mathrm{COV}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) =$$

$$= 1 + 1 + 2\frac{\mathrm{COV}(X,Y)}{\sigma_X\sigma_Y}) = 2(1+\rho(X,Y)) \quad \Rightarrow \quad \rho(X,Y) \geq -1$$

Analogamente da $Var(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}) \geq 0$ si ricava che $\rho(X,Y) \leq 1$.

Oss. $\rho(X,Y) = \pm 1$ corrisponde rispettivamente a $Var(\frac{X}{\sigma_X} \mp \frac{Y}{\sigma_Y}) = 0$, i.e. a $\mathbb{P}\left(\frac{X}{\sigma_X} \mp \frac{Y}{\sigma_Y} = c\right) = 1$, quindi $Y = a + bX$ con $b = \pm\frac{\sigma_Y}{\sigma_X}$ e $a = \mp\sigma_Y c$.

- Se $\rho(X,Y) = \pm 1$, le variabili aleatorie si dicono *perfettamente correlate* (Y è funzione lineare di X).

Alcune disuguaglianze

Prop. (Disuguaglianza di Markov) Sia X una variabile aleatoria non negativa. Allora, per ogni $a > 0$

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Supponiamo che X sia continua, dato che $X \geq 0$ con probabilità 1,

$$\mathbb{E}(X) = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \geq a \int_a^\infty f(x)dx = a\,\mathbb{P}(X \geq a).$$

Oss. La disuguaglianza fornisce un *limite superiore* per la probabilità di $X \geq a$ in termini di $\mathbb{E}(X)$.

Cor: (Disuguaglianza di Chebyshev) Sia X una variabile aleatoria di media μ e varianza σ^2 finite. Allora per ogni $k > 0$

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

Consideriamo l'evento $\{|X - \mu| \geq k\} = \{(X - \mu)^2 \geq k^2\}$ e applichiamo la disuguaglianza di Markov alla variabile aleatoria (non negativa) $(X - \mu)^2$, notando che $\mathbb{E}((X - \mu)^2) = \sigma^2$.

Oss. Se X ha $Var(X) = 0 \Rightarrow \mathbb{P}(X = \mathbb{E}(X)) = 1$.

Legge dei grandi numeri

Sono leggi dei grandi numeri tutti i teoremi che stabiliscono le condizioni sotto le quali la "media campionaria" converge al suo valore atteso in un qualche modo.

Teorema: "Legge (debole) dei grandi numeri"[1][2]

Sia $X_1, X_2, \dots$ una successione di variabili aleatorie i.i.d., con valore atteso finito $\mathbb{E}(X_i) = \mu$. Allora per ogni $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) = 0.$$

Sia $Var(X_i) = \sigma^2 < \infty$. Richiamiamo

$$\mathbb{E}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \mu \;\text{ e }\; Var\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}.$$

Dalla disuguaglianza di Chebyshev segue che

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2},$$

da cui si ottiene il risultato prendendone il limite per $n \to \infty$.

[1] l'aggettivo debole si riferisce al tipo di convergenza (convergenza in probabilità)

[2] Dimostrata da J. Bernoulli nel 1713 nel caso di variabili bernoulliane, questa dimostrazione si deve a Khintchine

Teorema del limite centrale

Nel teorema del limite centrale si dimostra che la distribuzione di somme di variabili aleatorie opportunamente standardizzate tende alla normale standard.

Teorema: "Teorema del limite centrale"[3]

Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie i.i.d., con valore atteso $\mathbb{E}(X_i) = \mu$ e varianza $Var(X_i) = \sigma^2$ finiti. Allora, la distribuzione di

$$\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}}$$

tende a quella di una normale standard, i.e.

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}} \leq z\right) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du, \qquad z \in \mathbb{R}.$$

Oss. Il teorema del limite centrale si può dimostrare sotto ipotesi molto più deboli.

Questo risultato giustifica l'evidenza empirica per cui molti dati reali osservati presentano una distribuzione a campana come se in effetti provenissero da una normale.

[3] Il teorema del limite centrale con variabili bernoulliane è noto come Teorema di De Moivre Laplace