05BQX Metodi Matematici per l'ingegneria 2011-2012

Marina Santacroce

Dipartimento di Scienze Matematiche, Politecnico di Torino

6. Leggi congiunte di variabili aleatorie

Riferimenti: S.Ross *Calcolo delle probabilità* Cap.6

Outline

Distribuzione congiunta di due variabili aleatorie

Siano X e Y *due* variabili aleatorie reali.

Def. La *funzione di distribuzione congiunta* di X e Y è una funzione $F_{XY} : \mathbb{R}^2 \to [0;1]$ definita da:

$$F_{XY}(x,y) = \mathbb{P}(X \leq x, Y \leq y) \qquad x, y \in \mathbb{R}.$$

Nota Bene: Dalla funzione di distribuzione congiunta delle variabili X e Y si ricavano le distribuzioni di X e di Y, che chiamiamo funzioni di distribuzione *marginali*:

$$F_X(x) = \lim_{y \to +\infty} F(x,y) \qquad F_Y(y) = \lim_{x \to +\infty} F(x,y)$$

Oss. Inoltre, a partire da F_{XY} si possono calcolare, con il principio di inclusione-esclusione, le probabilità di altre regioni di forma rettangolare:

$$\mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1)$$

Funzioni di densità congiunte: variabili discrete

Consideriamo due variabili aleatorie reali discrete X e Y.

Def. Definiamo la funzione di *densità (discreta) congiunta* di X e Y la funzione $p : \mathbb{X} \times \mathbb{Y}$[1] $\to [0,1]$

$$p(x,y) = \mathbb{P}(X = x, Y = y) \qquad (x,y) \in \mathbb{X} \times \mathbb{Y}.$$

Oss. $p(x,y) \geq 0$ per ogni (x,y) e $\displaystyle\sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} p(x,y) = 1.$

Le *densità marginali di X e di Y* sono rispettivamente:

$$p_X(x) = \textstyle\sum_{y\in\mathbb{Y}} p(x,y) \quad \text{con} \quad x \in \mathbb{X},$$
$$p_Y(y) = \textstyle\sum_{x\in\mathbb{X}} p(x,y) \quad \text{con} \quad y \in \mathbb{Y}.$$

▷
$$\begin{aligned} p_X(x) =& \mathbb{P}(X = x) = \mathbb{P}(X = x, \bigcup_{y\in\mathbb{Y}} (Y = y)) = \\ & \mathbb{P}(\bigcup_{y\in\mathbb{Y}} (X = x, Y = y)) = \sum_{y\in\mathbb{Y}} p(x,y) \end{aligned}$$

[1] $\mathbb{X}$ e $\mathbb{Y}$ sono rispettivamente il supporto della X e della Y

Esempio

Esempio: Si consideri **un'estrazione di 3 palline** da un'urna composta da 3 palline rosse, 4 blu e 5 bianche. Siano X e Y rispettivamente il numero di palline rosse estratte e di palline blu estratte. Supposto che tutte le estrazioni siano equiprobabili, determinare la **distribuzione congiunta di X e Y**.

- ▷ Le variabili X e Y possono assumere i valori $\{0, 1, 2, 3\}$.

  Determiniamo la densità congiunta:

  $p(i, j) = \mathbb{P}(X = i, Y = j)$ per ogni i e $j \in \{0, 1, 2, 3\}$ tali che $i + j \leq 3$.

  $p(0, 0) = \mathbb{P}(\{\text{estrazione 3 palline bianche}\}) = \frac{\binom{5}{3}}{\binom{12}{3}}$

  $p(0, 1) = \mathbb{P}(\{\text{estrazione 1 pallina blu e 2 palline bianche}\}) = \frac{\binom{4}{1}\binom{5}{2}}{\binom{12}{3}}$.

⇒ In generale,

$$p(i, j) = \frac{\binom{3}{i}\binom{4}{j}\binom{5}{3-(i+j)}}{\binom{12}{3}}$$

- ◇ Verificare che $p_X(1) = \sum_{j:0}^{3} p(1, j)$

  osservando che $p_X(1) = \mathbb{P}(\{\text{1 pallina rossa estratta}\}) = \frac{\binom{3}{1}\binom{9}{2}}{\binom{12}{3}}$

Funzioni di densità congiunte: variabili continue

Def. Due variabili aleatorie reali X e Y si dicono *congiuntamente (assolutamente) continue* se esiste una funzione f, che chiamiamo *funzione di densità congiunta* $f : \mathbb{R} \times \mathbb{R} \to [0; +\infty)$ tale che, tutte le volte che l'integrale è ben definito,

$$\mathbb{P}((X, Y) \in C) = \iint_C f(x, y)dxdy.$$

Quindi, f deve soddisfare

$$\int_{\mathbb{R}^2} f(x, y)dxdy = 1,$$

inoltre, $F_{XY}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)dudv$[2]

Oss. Derivando la funzione di distribuzione F_{XY}

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y).$$

Se prendiamo dx e dy "piccoli"

$$\mathbb{P}(x \leq X < x+dx, y \leq Y < y+dy) = \int_x^{x+dx} \int_y^{y+dy} f(u, v)dudv \approx f(x, y)dxdy.$$

[2] $C = (-\infty, x] \times (-\infty, y]$.

Esempio: tempi di vita di sitemi

Si consideri la funzione $f(x,y) = \begin{cases} kx & \text{se } 0 < x < 1, 0 < y < 1 \\ 0 & \text{altrimenti} \end{cases}$

con k costante. Trovare k in modo che f sia la densità congiunta di due variabili aleatorie.

- Un dispositivo è composto da due componenti e smette di funzionare non appena una delle due componenti smette di funzionare. La densità congiunta dei tempi di vita dei due componenti, misurate in ore, è $f(x,y)$ con $0 < x < 1$ e $0 < y < 1$.
  Qual è la probabilità che il dispositivo smetta di funzionare nella prima mezzora?
- Un dispositivo smette di funzionare non appena entrambe le componenti smettono di funzionare. La densità congiunta dei tempi di vita dei due componenti, misurate in ore, è

$$f(x,y) = \frac{x+y}{8} \qquad \text{se } 0 < x < 2 \text{ e } 0 < y < 2$$

  e nulla altrimenti.
  Qual è la probabilità che il dispositivo smetta di funzionare nella prima ora?

Funzioni di densità congiunte variabili continue: esempi

Se X e Y sono congiuntamente assolutamente continue anche la X e la Y sono continue.

- La funzione di densità della X si chiama di *densità marginale di X* ed è $f_X(x) = \int_{\mathbb{R}} f(x,y)dy, \quad x \in \mathbb{R}$

$\Rightarrow$ Si trova differenziando
$F_X(x) = \mathbb{P}(X \leq x, -\infty < Y < +\infty) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f(u,y)dydu$

- Analogamente, si può trovare la *densità marginale di Y*

  $f_Y(y) = \int_{\mathbb{R}} f(x,y)dx, \quad y \in \mathbb{R}$

Esempio:

La densità congiunta di X e Y è data dalla funzione

$$f(x,y)=\begin{cases}2e^{-x}e^{-2y} & 0<x<+\infty \; 0<y<+\infty \\ 0 & \text{altrimenti}\end{cases}$$

Determinare $\mathbb{P}(X>1, Y<1)$, $\mathbb{P}(X<Y)$, $\mathbb{P}(X<a)$ e la densità di $Z=\frac{X}{Y}$.

Variabili aleatorie indipendenti

Nel calcolo delle probabilità le variabili aleatorie indipendenti svolgono un ruolo fondamentale. La definizione di indipendenza tra variabili aleatorie si basa sul concetto di *indipendenza di eventi*.

Def. Due variabili aleatorie sono **indipendenti** se gli eventi determinati da X e gli eventi determinati da Y sono indipendenti, cioè

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \quad \text{per ogni } A \text{ e ogni } B$$

In parole povere, conoscere il valore di una delle variabili non modifica la distribuzione di probabilità dell'altra.

Utilizzando gli assiomi della probabilità, si può dimostrare che due variabili aleatorie X e Y sono indipendenti se e solo se:

$$F_{XY}(x, y) = F_X(x)F_Y(y) \text{ per ogni } x \text{ e } y \in \mathbb{R}.$$ [3]

[3] **Oss.** Se per ogni x e y, $F_{XY}(x, y) = H(x)G(y)$ oppure $f_{XY}(x, y) = g(x)h(y)$ nel caso continuo (l'analogo vale per il caso discreto) $\Rightarrow$ le variabili aleatorie X e Y sono indipendenti.

Indipendenza (Bis)

Inoltre, se le variabili aleatorie sono *discrete*, si può dimostrare che la definizione di indipendenza è equivalente a

$$p_{XY}(x,y) = p_X(x)p_Y(y) \text{ per ogni } x \in \mathbb{X} \text{ e } y \in \mathbb{Y}.$$

Analogamente, nel caso *continuo* si ha che X e Y sono indipendenti se e solo se

$$f_{XY}(x,y) = f_X(x)f_Y(y) \text{ per ogni } (x,y) \in \mathbb{R}^2.$$

Variabili aleatorie indipendenti: esempi

1. Ugo e Zoe si danno un'appuntamento davanti ad un locale, con l'accordo di aspettarsi solo per 10 minuti. Se ognuno di loro arriva in maniera indipendente dall'altro in un istante distribuito uniformemente tra le 21.00 e le 22.00, qual è la probabilità che i due entrino nel locale insieme?
2. La densità congiunta di X e Y è

$$f(x,y) = 6e^{-2x}e^{-3y} \qquad \text{se } x > 0 \text{ e } y > 0$$

e uguale a 0 altrimenti. Le variabili sono indipendenti?

e se la densità congiunta è uguale a

$$f(x,y) = 24xy, \qquad \text{se } 0 < x < 1,\ 0 < y < 1,\ 0 < x + y < 1,$$

e nulla altrove, le variabili sono indipendenti?

Somma di variabili aleatorie indipendenti

- Consideriamo X e Y variabili aleatorie reali *indipendenti*, ci poniamo il problema di determinare la distribuzione della variabile aleatoria somma di $Z = X + Y$.

Caso continuo Se X e Y sono due variabili aleatorie continue di densità f_X e f_Y, allora

$$
\begin{aligned}
F_Z(z) = F_{X+Y}(z) = \mathbb{P}(X + Y \leq z) &= \iint_{x+y\leq z} f_X(x) f_Y(y) dxdy \\
&= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{z-y} f_X(x) dx \right) f_Y(y) dy \\
&= \int_{-\infty}^{+\infty} F_X(z - y) f_Y(y) dy
\end{aligned}
$$

da cui, derivando,

$$
\boxed{f_Z(z) = f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy = f_X \star f_Y(z)}
$$

Questa operazione si chiama *convoluzione*.

Convoluzione: caso discreto

Caso discreto Se X e Y sono variabili aleatorie a *valori interi* di densità (discreta) p_X e p_Y, allora per ogni z intero

$$\begin{aligned} p_{X+Y}(z) &= \mathbb{P}(X+Y=z) \\ &= \sum_y \mathbb{P}(X+Y=z, Y=y) \\ &= \sum_y \mathbb{P}(X+y=z, Y=y) \\ &= \sum_y \mathbb{P}(X=z-y)\mathbb{P}(Y=y) \\ &= \boxed{\sum_y p_X(z-y)p_Y(y) = p_X \star p_Y(z)} \end{aligned}$$

Anche questa operazione si chiama *convoluzione*.

Convoluzione: esempi

1. Si calcoli la convoluzione di due variabili aleatorie con legge uniforme su $(0; 1)$.
2. Si verifichino i seguenti casi notevoli di convoluzione:
   - $\mathrm{Binom}(n, p) \star \mathrm{Binom}(m, p) = \mathrm{Binom}(n + m, p)$
   - $\mathrm{Poisson}(\lambda_1) \star \mathrm{Poisson}(\lambda_2) = \mathrm{Poisson}(\lambda_1 + \lambda_2)$
   - $\mathrm{Exp}(\lambda) \star \Gamma(n, \lambda) = \Gamma(n + 1, \lambda)$
   - $\mathrm{N}(\mu_1, \sigma_1^2) \star \mathrm{N}(\mu_2, \sigma_2^2) = \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Densità e valore atteso condizionati

Def. Siano X e Y due variabili aleatorie discrete con densità congiunta $p(x,y)$. Per ogni y tale che $p_Y(y) > 0$[4] si definisce *densità condizionata di* X *dato* $Y = y$, la funzione

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}, \ x \in \mathbb{X}.$$

Prop. La funzione di densità condizionata di X dato $Y = y$, $p_{X|Y}(\cdot|y)$, è una densità discreta per la X:

$$p_{X|Y}(x|y) \geq 0 \qquad \text{e} \qquad \sum_{x \in \mathbb{X}} p_{X|Y}(x|y) = 1.$$

Oss. Si noti che date p_Y e $p_{X|Y}$, la legge congiunta di X e Y è data da:

$$p_{XY}(x,y) = p_{X|Y}(x|y) p_Y(y)$$

Def. Definiamo valore atteso di X condizionato a $Y = y$

$$\mathbb{E}(X|Y = y) = \sum_x x \, p_{X|Y}(x|y)$$

[4] y è nel supporto di Y: $y \in \mathbb{Y}$

Densità e valore atteso condizionati caso continuo

Def. Se X e Y sono congiuntamente continue con densità f caso si definisce, per ogni y tale che $f_Y(y) \neq 0$ la *densità di* X *condizionata a* $Y = y$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} \quad x \in \mathbb{R}$$

Prop. La funzione $f_{X|Y}(\cdot|y)$ è una funzione di densità continua, i.e.

$$f_{X|Y}(x|y) \geq 0, x \in \mathbb{R} \qquad \int_{\mathbb{R}} f_{X|Y}(x|y)dx = 1$$

Oss. Si noti che

$$f_{XY}(x,y) = f_{X|Y}(x|y)f_Y(y)$$

Def. Si definisce valore atteso di X condizionato a $Y = y$

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{+\infty} x \, f_{X|Y}(x|y) \, dx.$$

Si noti che $\mathbb{E}(X|Y = y)$ è una funzione di y.

Prop. Date due variabili aleatorie X e Y, si verifichi che

$$\mathbb{E}(X) = \mathbb{E}\left(\mathbb{E}(X|Y)\right).$$

- ▷ Si consideri $g(Y) = \mathbb{E}(X|Y)$ e si calcoli $\mathbb{E}(g(Y))$.