Esercizi di Calcolo delle Probabilità con Elementi di Statistica Matematica

Lucio Demeio
Dipartimento di Scienze Matematiche
Università Politecnica delle Marche

1. **Esercizio.** *Siano X ed Y due variabili casuali discrete che possono assumere i valori $\{-1, 0, 1\}$ e la cui densità congiunta $p(x, y)$ è data dalla seguente tabella:*

	x		
y	-1	0	1
-1	a	b	a
0	b	0	b
1	a	b	a

(i) I parametri a e b possono essere scelti indipendentemente, o sono legati da qualche relazione?

(ii) calcolare le densità marginali p_X e p_Y;

(iii) verificare l'indipendenza (o meno) delle due variabili.

Soluzione.

(i) La somma delle densità congiunte su tutti i valori possibili deve essere uguale a 1 per la condizione di normalizzazione, dunque $4a + 4b = 1$ e pertanto a e b devono obbedire alla relazione $a + b = 1/4$.

(ii) La densità marginale rispetto ad X è data da $p_X(x) = \sum_y p(x, y)$. Dunque

$$p_X(-1) = 2a + b, \qquad p_X(0) = 2b, \qquad p_X(1) = 2a + b$$

ed analogamente

$$p_Y(-1) = 2a + b, \qquad p_Y(0) = 2b, \qquad p_Y(1) = 2a + b.$$

(iii) le due variabili non sono indipendenti, in quanto, ad esempio, $p(-1, -1) = a \neq p_X(-1)\, p_Y(-1) = (2a + b)^2$.

2. **Esercizio.** *Sia X una variabile casuale continua distribuita uniformemente sull'intervallo $[0, 1]$ ed Y una seconda variabile continua distribuita con legge esponenziale di*

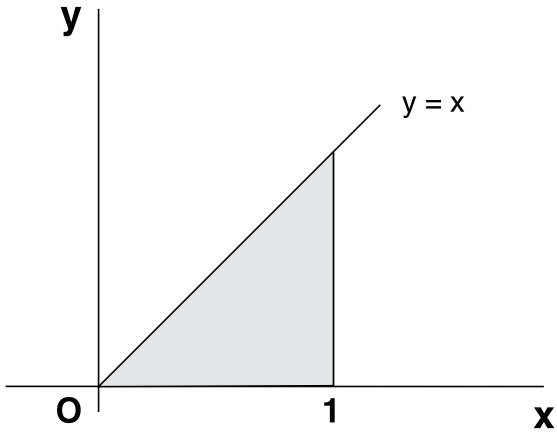

Figura 1: La regione $x \geq y$.

parametro $\lambda = 2$. Determinare la probabilità che X assuma valori maggiori o uguali ad Y, nell'ipotesi che le due variabili siano indipendenti.

Soluzione.

Abbiamo dal testo del problema:

$f_X(x) = 1$ se $0 \leq x \leq 1$ ed $f_X(x) = 0$ altrimenti;
$f_Y(y) = \lambda\, e^{-\lambda y}$ se $y \geq 0$ ed $f_Y(y) = 0$ altrimenti;
$f(x,y) = f_X(x)\, f_Y(y)$.

Sia A la regione del piano cartesiano dove $x \geq y$ e contemporaneamente $f(x,y) \neq 0$, vale a dire

$A = \{(x,y) \in \mathbb{R}^2 : x \geq y \;\&\; f(x,y) \neq 0\}$ (vedi figura 1).

Dunque la probabilità richiesta è

$$
\begin{aligned}
P(\{X \geq Y\}) &= \int\int_{\mathbb{R}^2} f(x,y)\, dx\, dy = \int_0^1 dx \int_0^x dy\, \lambda\, e^{-\lambda y} = \int_0^1 dx\, [-e^{-\lambda y}]_0^x = \\
&= \int_0^1 dx\, [1 - e^{-\lambda x}] = 1 - [-\frac{e^{-\lambda x}}{\lambda}]_0^1 = 1 + \frac{1}{\lambda}(e^{-\lambda} - 1) = 1 + \frac{e^{-\lambda} - 1}{\lambda} = \\
&(\lambda = 2) \to \frac{1 + e^{-2}}{2}
\end{aligned}
$$

3. **Esercizio.** *Un gruppo di n automobili arriva ad un incrocio, dove ciascuna automobile può continuare dritta, svoltare a destra oppure a sinistra. Da osservazioni statistiche precedenti, si sa che la probablità di svoltare a destra è p, la probabilità di svoltare a sinistra è q, mentre r è la probabilità di continuare dritta. Si supponga inoltren che ciascun automobilista si comporti indipendentemente dagli altri. Siano X ed Y le variabili casuali*

che indicano il numero di macchine che svoltano rispettivamente a destra ed a sinistra. Determinare la probabilità congiunta di X e Y.

Soluzione. Intanto, dobbiamo avere $p + q + r = 1$, ed X ed Y sono variabili discrete, con $x, y = 0, 1, 2, ..., n$. In questo caso, è più semplice determinare la densità discreta, $\bar{p}_{X|Y}(x|y)$ e da questa risalire alla densità congiunta dalla relazione $p(x, y) = \bar{p}_{X|Y}(x|y)\, p_Y(y)$. La densità marginale $p_Y(y)$ rappresenta la probabilità che y automobili svoltino a sinistra, su un totale di n macchine. È dunque lo stesso problema di determinare il numero di teste su n lanci di una moneta, e quindi

$$p_Y(y) = \binom{n}{y} q^y (1-q)^{n-y}$$

La densità condizionale $\bar{p}_{X|Y}(x|y)$ rappresenta la probabilità che x automobili svoltino a destra sapendo già che y automobili hanno svoltato a sinistra. La probabilità in questo caso non è più p, perchè il numero di macchine si è ridotto ad $n - y$. La probabilità di svolta a destra è quindi ora $p/(r + p)$ e pertanto

$$\bar{p}_{X|Y}(x|y) = \binom{n-y}{x} \left(\frac{p}{r+p}\right)^x \left(1 - \frac{p}{r+p}\right)^{n-y-x}, \quad x = 0, 1, ..., n-y;\ y = 0, 1, ..., n.$$

La densità congiunta è data dal prodotto $\bar{p}_{X|Y}(x|y)\, p_Y(y)$.

4. **Esercizio.** *Un test diagnostico di una certa malattia ha accuratezza η, con $0 < \eta < 1$, sia quando le persone sono malate che quando sono sane. Se p è la frazione della popolazione che presenta la malattia, determinare la probabilità che una persona sia malata se risulta positiva al test. Calcolare quindi tale probabilità nel caso $\eta = 95\%$ e $p = 0.005$.*

Soluzione. Prendendo una persona a caso dalla popolazione, indichiamo con M l'evento "la persona è malata", con S "la persona è sana", con T "la persona è positiva al test" e con N "la persona è negativa al test". In termini di probabilità, i dati del problema ci dicono che

$$P(T|M) = P(N|S) = \eta, \qquad P(M) = p$$

da cui deduciamo che $P(T|S) = P(N|M) = 1 - \eta$, e che $P(S) = 1 - p$, mentre la probabilità richiesta è $P(M|T)$. Usando la formula di Bayes abbiamo che

$$P(M|T) = \frac{P(T|M)\, P(M)}{P(T)},$$

ma dobbiamo ancora determinare $P(T)$. Questo lo determiniamo dalla formula delle probabilità totali, quindi

$$P(T) = P(T|M)P(M) + P(T|S)P(S) = \eta\, p + (1-\eta)(1-p)$$

e dunque

$$P(M|T) = \frac{\eta\, p}{\eta\, p + (1-\eta)(1-p)}.$$

Per $\eta = 0.95$ e $p = 0.005$ otteniamo $P(M|T) = 0.087$.

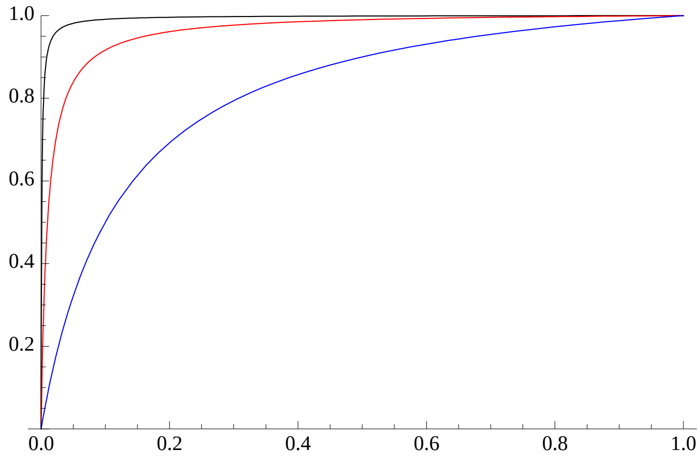

Figura 2: La probabilità $P(M|T)$ in funzione di p per $\eta = 0.9$ (blu), 0.99 (rossa), 0.999 (nera).

5. **Esercizio.** *Siano X ed Y due variabili aleatorie continue che prendono valori nell'intervallo $[0, 1]$ e sia $f(x, y) = a\,(x + 2\,y)$ la densità congiunta. Si chiede di:*

   (i) calcolare a ed usare il valore ottenuto nelle domande successive;

   (ii) calcolare le densità marginali;

   (iii) dire se le variabili X ed Y sono indipendenti.

   Soluzione. *(i)* Il valore di a si determina imponendo la normalizzazione della densità congiunta:

$$1 = \int\int_{\mathbb{R}^2} f(x, y)\, dx\, dy = \int_0^1 dx \int_0^1 dy\, a\,(x + 2\,y) = a \int_0^1 dx\, \left[x\,y + y^2\right]_0^1 =$$
$$= a \int_0^1 dx(x + 1) = a \left[\frac{x^2}{2} + x\right]_0^1 = \frac{3}{2}\, a$$

   e dunque $a = 2/3$ e $f(x, y) = 2\,(x + 2\,y)/3$.

   (ii) Per le densità marginali abbiamo:

$$f_X(x) = \int_0^1 f(x, y)\, dy = \frac{3}{2} \int_0^1 (x + 2\,y)\, dy = \frac{3}{2}\,(x + 1)$$
$$f_Y(y) = \int_0^1 f(x, y)\, dx = \frac{3}{2} \int_0^1 (x + 2\,y) dx = \frac{3}{2} \left(\frac{1}{2} + 2y\right)$$

   (iii) Le variabili X ed Y non sono indipendenti. Infatti:

$$f_X(x)\, f_Y(y) = \frac{9}{4}\,(x + 1) \left(\frac{1}{2} + 2y\right) \neq f(x, y) = \frac{3}{2}\,(x + 2\,y)$$

6. **Esercizio.** *In un certo territorio, tre compagnie telefoniche, A, B e C, si spartiscono il mercato delle comunicazioni in ragione, rispettivamente, del 30%, 45% e 25% degli utenti. La compagnia A copre il 70% del territorio, la compagnia B il 90% e la compagnia C l' 80%.*

(i) *Qual'è la probabilità che un utente a caso non sia coperto?*

(ii) *Qual'è la probabilità che un utente non coperto sia cliente della compagnia B?;*

Soluzione. Scelto un utente a caso, indichiamo con A l'evento "l' utente è abbonato alla compagnia A", ed analogamente per gli eventi B e C. I dati del problema dicono che

$$P(A) = 0.3, \qquad P(B) = 0.45, \qquad P(C) = 0.25$$

Siano inoltre S ed N rispettivamente gli eventi "l' utente è coperto" e "l' utente non è coperto". Dal testo sappiamo che

$$P(S|A) = 0.7, \qquad P(S|B) = 0.9, \qquad P(S|C) = 0.8$$

e di conseguenza $P(N|A) = 1 - P(S|A) = 0.3$, $P(N|B) = 0.1$ e $P(N|C) = 0.2$. *(i)* La probabilità che l'utente non sia coperto, $P(N)$, è data dal teorema delle probabilità totali

$$P(N) = P(N|A)\,P(A) + P(N|B)\,P(B) + P(N|C)\,P(C) = 0.185$$

(ii) Si chiede di determinare $P(B|N)$. Usando il teorema di Bayes,

$$P(B|N) = \frac{P(N|B)\,P(B)}{P(N)} = 0.243$$

7. **Esercizio.** *Sia*

$$F_X(t) = \begin{cases} 0, & t < 0 \\ t + \alpha, & 0 \le t < \beta \\ 1, & t \ge \beta \end{cases}$$

la funzione di ripartizione di una variabile aleatoria X.

(i) *Determinare la condizione su α e β affinchè $F_X(t)$ sia una funzione di partizione;*

(ii) *determinare la condizione su α e β affinchè $F_X(t)$ sia continua per tutte le t positive;*

(iii) *per quali valori di α e β la variabile X è continua?*

(iv) *Determinare tutti i valori che può assumere X per $\alpha = 0.2$ e $\beta = 0.5$; È possibile associare una probabilità non nulla a qualcuno di questi valori? Se si, qual'è questo valore e quanto vale la sua probabilità?*

Soluzione. *(i)* Deve essere $\alpha + \beta \le 1$, altrimenti $F_X(\beta) > 1$.

(ii) L'unico valore di t per cui la F_X potrebbe presentare una discontinuità è $t = \beta$. Affinchè F_X sia ivi continua dobbiamo avere $\beta + \alpha = 1$.

(iii) Rimane la possibile discontinuità in $t = 0$, dove F_X è continua se $\alpha = 0$, e dunque pure $\beta = 0$.

(iv) La variabile X può assumere tutti i valori nell'intervallo $[0, 1]$. Può assumere i valori 0 ed 1 con probabilità non nulla, in particolare $P(\{X = 0\}) = \alpha = 0.2$ e $P(\{X = 1\}) = 1 - (\alpha + \beta) = 0.3$