Corso di Laurea in Ingegneria Informatica
Anno Accademico 2017/2018
Probabilità e Statistica - Prova pratica

Nome ....

N. Matricola .... Ancona, 16 febbraio 2018

1. (6 punti) Un azienda che produce relè elettrici possiede tre impianti di fabbricazione, diciamo A, B e C, che forniscono rispettivamente il 50%, il 30% ed il 20% della produzione. L'esperienza suggerisce che la probabilità che un relè risulti difettoso è pari a 0,02 se viene prodotto dall'impianto A, 0,05 se prodotto dall'impianto B e 0,01 se prodotto dall'impianto C.

   - Selezionando casualmente un relè dalla produzione dell'ultimo mese, si calcoli la probabilità che questo risulti difettoso.
   - Un relè selezionato casualmente viene trovato difettoso. Si calcoli la probabilità che il relè sia stato fabbricato dall'impianto B.

Svolgimento. Introduciamo gli eventi:

$A =$"il relè viene dall'impianto A";
$B =$"il relè viene dall'impianto B";
$C =$"il relè viene dall'impianto C";
$D =$"il relè è difettoso".

Allora dai dati abbiamo:

$$P(A) = 0.5; \quad P(B) = 0.3; \quad P(C) = 0.2;$$
$$P(D|A) = 0.02; \quad P(D|B) = 0.05; \quad P(D|C) = 0.05.$$

- La risposta alla prima domanda è:

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

Sostituendo i valori abbiamo $P(D) = 0.035$.

- Per la seconda domanda dobbiamo calcolare $P(B|D)$. Usando la formula di Bayes abbiamo:

$$P(B|D) = \frac{P(D|B)\,P(B)}{P(D)} = 0.43$$

2. (9 punti) Il numero di macchine che transitano in una giornata nel mese di ottobre nel centro di una città è rappresentato da una variabile casuale normale X di media μ_X e deviazione standard σ_X. Il livello giornaliero in millilitri di pioggia nello stesso centro cittadino durante il mese di ottobre invece descritto dalla variabile casuale Y con media μ_Y e deviazione standard σ_Y. La concentrazione di anidride carbonica (in microgrammi per millilitro) presente nell'atmosfera in una giornata del mese di ottobre è a sua volta rappresentata da una variabile Z legata ad X e Y dalla relazione $Z = a\,X - b\,Y - z_0$. Sapendo che $Cov(X,Y) = \gamma$,

   - determinare la media e la varianza di Z in funzione dei parametri introdotti sopra;
   - siano ora: $\mu_X = 2500$, $\mu_Y = 0.5$, $\sigma_X = 300$, $\sigma_Y = 0.25$, $a = 0.003$, $b = 7.72$, $z_0 = 3.6$, $\gamma = 97.85$; calcolare la probabilità che, in una giornata di ottobre, la concentrazione di anidride carbonica superi la soglia critica di 0,06.

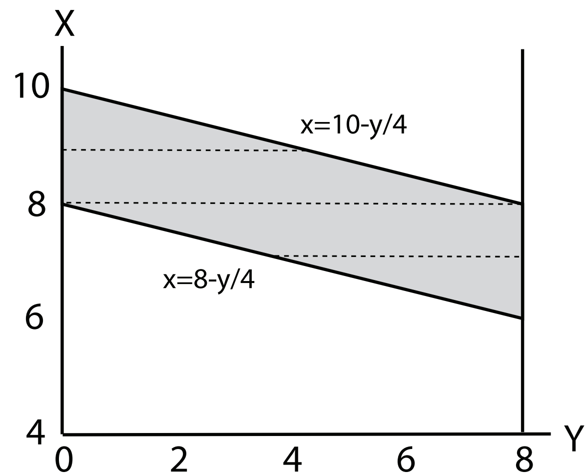
Svolgimento.

- Per rispondere alla prima domanda non sono necessarie le ipotesi sulla natura delle distribuzioni di X e Y, ma sono sufficienti le proprietà degli operatori di media e varianza. Abbiamo così:

$$\begin{aligned}
&\mu_Z \equiv E[Z] = a\,E[X] - b\,E[Y] - z_0 == a\,\mu_X - b\,\mu_Y - z_0\\
&\sigma_Z^2 \equiv Var(Z) = Var(a\,X - b\,Y - z_0) = Var(a\,X - b\,Y) =\\
&\quad = Var(a\,X) + Var(b\,Y) - 2\,Cov(a\,X, b\,Y) =\\
&\quad = a^2\,Var(X) + b^2\,Var(Y) - 2\,a\,b\,Cov(X,Y) = a^2\,\sigma_X^2 + b^2\,\sigma_Y^2 - 2\,a\,b\,\gamma
\end{aligned}$$

- Con i valori dati abbiamo innanzitutto $\mu_Z = 0.04$ e $\sigma_Z^2 = 0.0025$. Dobbiamo supporre che $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$. Con ciò abbiamo:

$$\begin{aligned}
&P(Z > 0.06) = P\left(\frac{Z-\mu_Z}{\sigma_Z} > \frac{0.06-0.04}{0.05}\right) = P\left(\frac{Z-\mu_Z}{\sigma_Z} > 0.4\right) =\\
&\quad = 1 - \Phi(0.4) = 0.3446
\end{aligned}$$



3. (9 punti) Il tempo di impiego della CPU da parte di una certa applicazione del sistema di un computer è rappresentato dalla variabile Y, uniforme nell'intervallo $[0, 8]$ (in ore). La durata di una carica della batteria è a sua volta rappresentata da una variabile casuale X, uniforme nell'intervallo $[8 - y/4, 10 - y/4]$.

- Determinare la legge (marginale) della variabile X;
- calcolare media e varianza di X;
- calcolare la probabilità che la carica della batteria duri più di 7 ore e meno di 9 ore.

Svolgimento. I dati del problema dicono che la variabile Y ha come legge

$$\begin{aligned} f_Y(y) &= \frac{1}{8}, \quad 0 \le y \le 8 \\ &= 0 \quad \text{altrimenti} \end{aligned}$$

mentre l'altra indicazione dice che

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{1}{2}, \quad 8 - y/4 \le x \le 10 - y/4 \\ &= 0 \qquad \text{altrimenti} \end{aligned}$$

La densità congiunta $f_{XY}(x, y) = f_{X|Y}(x|y)\, f_Y(y)$ è allora data da

$$\begin{aligned} f_{XY}(x, y) &= \frac{1}{16}, \quad 0 \le y \le 8, \quad 8 - y/4 \le x \le 10 - y/4 \\ &= 0 \quad \text{altrimenti} \end{aligned}$$

La regione dove $f_{XY}(x,y)$ è diversa da zero è la regione ombreggiata nella figura ed è la regione compresa tra le rette di equazione $x = 8 - y/4$ e $x = 10 - y/4$ con $0 \leq y \leq 8$. Le due rette si possono anche scrivere come $y = 4\,(8 - x)$ e $y = 4\,(10 - x)$, che torneranno utili fra poco.

- La densità marginale della variabile X,

$$f_X(x) \equiv \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

  è quindi il risultato di un'integrazione della densità congiunta $f_{XY}(x,y)$ rispetto a y. Come si vede dalla figura, gli estremi dell'integrale su y sono $[0, 4\,(10-x)]$ se $x \in [8,10]$ e $[4\,(8-x), 8]$ se $x \in [6,8]$. Abbiamo pertanto:

$$\begin{aligned} f_X(x) &= \int_0^{4\,(10-x)} \frac{1}{16}\,dy = \frac{1}{4}\,(10-x), \quad \text{se} \quad 8 \leq x \leq 10 \\ f_X(x) &= \int_{4\,(8-x)}^{8} \frac{1}{16}\,dy = \frac{1}{4}\,(x-6), \quad \text{se} \quad 6 \leq x \leq 8 \end{aligned}$$

- Per la media e la varianza di X abbiamo:

$$\begin{aligned} E[X] &= \int_6^{10} x\, f_X(x)\,dx = \int_6^8 x\,\frac{x-6}{4}\,dx + \int_8^{10} x\,\frac{10-x}{4}\,dx = 8 \\ E[X^2] &= \int_6^{10} x^2\, f_X(x)\,dx = \int_6^8 x^2\,\frac{x-6}{4}\,dx + \int_8^{10} x^2\,\frac{10-x}{4}\,dx = \frac{194}{3} \\ Var(x) &= E[X^2] - E[X]^2 = \frac{194}{3} - 64 = \frac{2}{3} \end{aligned}$$

- $$P(7 < X < 9) = \int_7^9 f_X(x)\,dx = \int_7^8 \frac{x-6}{4}\,dx + \int_8^9 \frac{10-x}{4}\,dx = \frac{3}{4}$$

4. (6 punti) Si vuole stimare il reddito medio mensile delle famiglie di una certa regione. A tale scopo viene selezionato un campione di 10 famiglie, ottenendo un reddito medio mensile di 1500 euro con una varianza campionaria di 40000 euro^2. Determinare gli intervalli di confidenza per il reddito medio al 10%, 5% e 1%.

Svolgimento. Il campione ha rango $n = 10$, quindi dovremo usare i quantili della legge di Student a 9 gradi di libertà. Abbiamo $\overline{X}_n = 1500$ per la media campionaria e $S_n^2 = 40000$ per la varianza campionaria. I valori che ci interessano, con $\alpha = 0.1$, 0.05 e 0.01 sono:

α	$\alpha/2$	$t_{\alpha/2}(9)$
0.1	0.05	1.833
0.05	0.025	2.262
0.01	0.005	3.250

Gli intervalli di confidenza sono dati dalla formula

$$\left[\overline{X}_n - \frac{S_n}{\sqrt{n}}\, t_{\alpha/2}(n-1), \overline{X}_n + \frac{S_n}{\sqrt{n}}\, t_{\alpha/2}(n-1)\right].$$

Sostituendo i valori dati otteniamo:

α	intervallo
0.1	$[1384, 1615]$
0.05	$[1357, 1643]$
0.01	$[1294, 1706]$