Probabilità e Statistica (*Ing. Amb. Terr. - Latina* - 17/1/2008)

(il punteggio massimo si ottiene risolvendo correttamente 5 esercizi)

1. Tizio può abbinare la possibile vincita di una somma S a uno solo tra due esperimenti aleatori: (i) 3 lanci di una moneta ('non truccata'), in cui la somma S viene vinta se esce sempre Testa o sempre Croce (evento A); (ii) 3 estrazioni con restituzione da un'urna contenente una proporzione p di palline bianche e $q = 1 - p$ di palline nere, in cui la somma S viene vinta se esce sempre pallina bianca oppure sempre pallina nera (evento B). A quale esperimento conviene abbinare la vincita?

   Esperimento preferibile ?

2. Con riferimento all'esercizio precedente, si assuma che Tizio abbia scelto in modo aleatorio, con uguale probabilità, l'abbinamento a uno dei due esperimenti e che il risultato di tale scelta non sia noto. Calcolare al variare del parametro p il valore massimo della probabilità condizionata γ che l'abbinamento sia stato fatto al primo esperimento (evento H), supposto che Tizio abbia vinto la somma S (evento E).

   $\gamma =$

3. Fissato un valore $a \in (0,1)$, siano: A il pentagono di vertici i punti $(0,0), (a,0), (1,1-a), (1,1), (0,1)$; T_1 il triangolo di vertici i punti $(0,0), (a,0), (0,1)$; T_2 il triangolo di vertici i punti $(1,1-a), (1,1), (0,1)$; T_3 il triangolo di vertici i punti $(a,0), (1,1-a), (0,1)$. Un punto aleatorio $Q = (X,Y)$ ha una distribuzione di probabilità uniforme su A. Definiti gli eventi $E_i = (Q \in T_i)$, $i = 1,2,3$, calcolare: (i) $\alpha = P(E_1 \mid E_1 \vee E_2)$; (ii) il valore di a tale che $P(E_i) = \frac{1}{3}$.

   $\alpha =$ $\qquad$ $a =$

4. Con riferimento all'esercizio 3, calcolare per $a = \frac{1}{2}$ la densità $f_1(x)$ e la mediana M di X.

   $f_1(x) =$ $\qquad$ $M =$

5. La densità congiunta di un vettore aleatorio (X,Y) è $f(x,y) = e^{-x-y}$ per $x > 0, y > 0$, con $f(x,y) = 0$ altrove. Calcolare, per ogni $z > 0$, la funzione di sopravvivenza $S(z)$ del numero aleatorio $Z = \frac{Y}{X}$.

   $S(z) =$

6. Con riferimento all'esercizio precedente, calcolare per ogni $z > 0$ la funzione di rischio $h(z)$ di Z. Stabilire inoltre se, fissati tre valori positivi z, z_1, z_2, con $z_1 < z_2$, e posto
$$p_1 = P(Z > z_1 + z \mid Z > z_1)\,, \quad p_2 = P(Z > z_2 + z \mid Z > z_2)\,,$$
   risulta $p_1 < p_2$ oppure no.

   $h(z) =$ $\qquad$ $p_1 < p_2$?

7. La distribuzione iniziale $\beta(\theta)$ di un parametro aleatorio $\Theta \in [0,1]$ è di tipo beta di parametri $r_0 = s_0 = 2$. Le componenti di un campione casuale $X = (X_1, \ldots, X_6)$, subordinatamente a ogni fissato valore θ, hanno una distribuzione binomiale di parametri $n = 1$, $p = \theta$. Avendo osservato un campione $x = (x_1, \ldots, x_6)$, con $x_1 + \cdots + x_6 = 3$, calcolare la previsione di Θ condizionata ad x.

   $\mathbb{P}(\Theta \mid x) =$

Probabilità e Statistica (Ing. Amb. Terr. - Latina)

Soluzioni della prova scritta del 17/1/2008.

1. Definiti gli eventi $T_i =$ *"esce Testa nell'i-mo lancio"*, $C_i =$ *"esce Croce nell'i-mo lancio"*, $E_i =$ *"esce pallina bianca nell'i-ma estrazione"*, $i = 1, 2, 3$, si ha

$$P(A) = P(T_1T_2T_3 \vee C_1C_2C_3) = P(T_1)P(T_2)P(T_3) + P(C_1)P(C_2)P(C_3) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}\,,$$

$$P(B) = P(E_1E_2E_3 \vee E_1^cE_2^cE_3^c) = P(E_1)P(E_2)P(E_3) + P(E_1^c)P(E_2^c)P(E_3^c) = p^3 + q^3 = 3p^2 - 3p + 1\,,$$

con

$$P(B) - P(A) = 3p^2 - 3p + \frac{3}{4} = 3\left(p - \frac{1}{2}\right)^2 \geq 0\,, \ \ \forall p\,.$$

Conviene, quindi, scegliere il secondo esperimento se $p \neq \frac{1}{2}$; gli esperimenti sono equivalenti se $p = \frac{1}{2}$.

2. Si ha: $E = AH \vee BH^c$; quindi

$$P(E|H) = P(A|H) = P(A)\,, \quad P(E|H^c) = P(B|H^c) = P(B)\,;$$

allora, tenendo conto che $P(H) = P(H^c)$, si ottiene

$$\gamma = P(H|E) = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|H^c)P(H^c)} = \frac{P(E|H)}{P(E|H) + P(E|H^c)} = \frac{P(A)}{P(A) + P(B)} =$$

$$= \frac{\frac{1}{4}}{\frac{1}{4} + 3p^2 - 3p + 1} = \frac{\frac{1}{4}}{3\left(p - \frac{1}{2}\right)^2 + \frac{1}{2}} = \frac{1}{12\left(p - \frac{1}{2}\right)^2 + 2}\,.$$

Essendo $\left(p - \frac{1}{2}\right)^2 \geq 0$, segue $\gamma \leq \frac{1}{2}$, con $\gamma = \frac{1}{2}$ per $p = \frac{1}{2}$.
(*Nota: se* $p = \frac{1}{2}$ *si ha* $P(H|E) = P(H)$; *se* $p \neq \frac{1}{2}$ *si ha* $P(H|E) < P(H)$).

3. Indicando con $\mu(S)$ l'area di un fissato insieme S, si ha

$$\mu(A) = 1 - \frac{(1-a)^2}{2} = \frac{1 + 2a - a^2}{2}\,, \quad \mu(T_1) = \mu(T_2) = \frac{a}{2}\,.$$

Inoltre, $f(x, y) = k = \frac{1}{\mu(A)} = \frac{2}{1+2a-a^2}$ per $(x, y) \in A$, con $f(x, y) = 0$ altrove. Allora

$$P(E_1) = \int\int_{T_1} f(x,y)dxdy = k\,\mu(T_1) = \frac{ka}{2}\,, \ \ P(E_2) = \int\int_{T_2} f(x,y)dxdy = k\,\mu(T_2) = \frac{ka}{2} = P(E_1)\,.$$

Pertanto

$$\alpha = \frac{P[E_1 \wedge (E_1 \vee E_2)]}{P(E_1 \vee E_2)} = \frac{P(E_1)}{P(E_1) + P(E_2)} = \frac{1}{2}\,, \ \forall a \in (0, 1)\,.$$

Inoltre, tenendo conto che $0 < a < 1$, si ha

$$P(E_1) = P(E_2) = P(E_3) = \frac{1}{3} \iff \frac{ka}{2} = \frac{1}{3} \iff 1 + 2a - a^2 = 3a \iff a = \frac{\sqrt{5} - 1}{2}\,.$$

4. Osservando che per $a = \frac{1}{2}$ si ha $\mu(A) = \frac{7}{8}$ e che l'equazione della retta passante per i punti $(\frac{1}{2}, 0), (1, \frac{1}{2})$ è $y = x - \frac{1}{2}$, segue

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_0^1 \frac{8}{7}\,dy = \frac{8}{7}\,, \quad 0 \le x \le \frac{1}{2}\,;$$

$$f_1(x) = \int_{x-\frac{1}{2}}^1 \frac{8}{7}\,dy = \frac{8}{7}(\frac{3}{2} - x)\,, \quad \frac{1}{2} < x \le 1\,; \qquad f_1(x) = 0\,, \quad x \notin [0,1]\,.$$

Inoltre, ricordando che M è il valore tale che $P(X \le M) = \frac{1}{2}$, ovvero $\int_{-\infty}^M f_1(x)\,dx = \frac{1}{2}$, osservando che $\int_0^{\frac{1}{2}} f_1(x)\,dx = \frac{4}{7} > \frac{1}{2}$, segue $M < \frac{1}{2}$; quindi $f_1(x) = \frac{8}{7}$, $\forall\, x \in [0, M]$. Allora, dev'essere $\int_0^M \frac{8}{7}\,dx = \frac{8}{7}\,M = \frac{1}{2}$; pertanto: $M = \frac{7}{16}$.

5. Si ha

$$S(z) = P(Z > z) = P(Y > zX) = \int_0^{+\infty} dx \int_{zx}^{+\infty} e^{-x-y} dy =$$

$$= \int_0^{+\infty} e^{-x} e^{-zx} dx = \frac{1}{1+z} \int_0^{+\infty} (1+z) e^{-(1+z)x} dx = \frac{1}{1+z}\,, \quad z > 0\,.$$

6. Si ha

$$h(z) = \frac{f(z)}{S(z)} = \frac{-S'(z)}{S(z)} = \frac{\frac{1}{(1+z)^2}}{\frac{1}{(1+z)}} = \frac{1}{(1+z)} = S(z)\,, \ z > 0\,.$$

Inoltre

$$p_1 = P(Z > z_1 + z \,|\, Z > z_1) = \frac{P(Z > z_1 + z, Z > z_1)}{P(Z > z_1)} = \frac{P(Z > z_1 + z)}{P(Z > z_1)} =$$

$$= \frac{S(z_1 + z)}{S(z_1)} = \frac{1 + z_1}{1 + z_1 + z}\,; \quad p_2 = P(Z > z_2 + z \,|\, Z > z_2) = \cdots = \frac{1 + z_2}{1 + z_2 + z}\,;$$

$$p_2 - p_1 = \cdots = \frac{z_2 - z_1}{(1 + z_2 + z)(1 + z_2 + z)} \ > \ 0\,, \ \ \forall\, z_1 > 0, \, z_2 > 0, \, z > 0, \ z_1 < z_2\,.$$

Pertanto: $p_1 < p_2$.

Nota: la stessa conclusione segue anche osservando che, essendo $h(z)$ decrescente, si ha $\int_{z_1}^{z_1+z} h(t)\,dt \ > \ \int_{z_2}^{z_2+z} h(t)\,dt$; pertanto

$$p_1 = \frac{S(z_1 + z)}{S(z_1)} = \cdots = e^{-\int_{z_1}^{z_1+z} h(t)dt} \ < \ e^{-\int_{z_2}^{z_2+z} h(t)dt} = \frac{S(z_2 + z)}{S(z_2)} = p_2\,.$$

7. Si ha

$$\beta(\theta) = \frac{\Gamma(r_0 + s_0)}{\Gamma(r_0)\Gamma(s_0)}\,\theta^{r_0 - 1}(1-\theta)^{s_0 - 1} = 6\,\theta(1-\theta)\,, \quad \theta \in [0,1]\,,$$

con $\beta(\theta) = 0$ altrove. Inoltre $f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i}$, $i = 1, \ldots, 6$; quindi

$$\alpha(x|\theta) = \theta^{\sum_i x_i}(1-\theta)^{n - \sum_i x_i} = \theta^3(1-\theta)^3\,.$$

Allora

$$\beta(\theta|x) = k(x)\beta(\theta)\alpha(x|\theta) = k_1(x)\theta^4(1-\theta)^4\,, \quad \theta \in [0,1]\,,$$

con $\beta(\theta|x) = 0$ altrove. Quindi, la distribuzione finale di Θ è di tipo beta di parametri $r_6 = s_6 = 5$, con $k_1(x) = \frac{\Gamma(r_6+s_6)}{\Gamma(r_6)\Gamma(s_6)} = \frac{\Gamma(10)}{\Gamma(5)\Gamma(5)} = \frac{9!}{4!\,4!} = 630$. Pertanto: $\mathbb{P}(\Theta \,|\, x) = \frac{r_6}{r_6+s_6} = \frac{1}{2}$.