Università degli Studi di Firenze

Facolta' di Ingegneria — Facsimile dell'Esame di Statistica — 11-4-2002

NOTA BENE: Nel rispondere alle domande teoriche' e' necessario definire TUTTI i termini introdotti e spiegare le notazioni. Per risolvere i problemi e' necessario spiegare i passaggi.

1. Una compagnia di assicurazione ha tre tipi di clienti: ad alto rischio, a medio rischio e a basso rischio. Il 20% dei clienti e' ad alto rischio, il 30% a medio rischio e il 50% a basso rischio. La probabilita' che un cliente abbia almeno un incidente nel corso di quest'anno e' 0.25 per la classe ad alto rischio, 0.16 per la classe a medio rischio e 0.10 per la classe a basso rischio.
   (a) Trovare la probabilita' che un cliente scelto a caso abbia almeno un incidente nel corso di quest'anno.
   (b) Trovare la probabilita' che un cliente sia ad alto rischio, sapendo che tale cliente ha avuto almeno un incidente quest'anno.

2. La durata X (in centinaia di ore) di un transistor e' una variabile aleatoria con funzione di ripartizione

$$F(x) = \begin{cases} 0 & y < 0 \\ 1 - e^{-x^2}, & x \geq 0 \end{cases}$$

   (a) Mostrare che $F(x)$ ha le proprieta' di una funzione di ripartizione.
   (b) Trovare la funzione di densita'.
   (c) Trovare la probabiita' che il transistor funzioni per almeno 200 ore.

3. (a) Definire la covarianza tra due variabili aleatorie discrete X e Y. (b) Che cosa misura tale indice? (c) Che valori puo' assumere? Come si interpreta? (d) Qual e' l'interpretazione nelle situazioni estreme? (e) Se X e' un altezza in cm e Y e' un peso in kg e la covarianza tra i due e' 7.5, calcolare la covarianza tra l'altezza espressa in metri $X^* = X/100$ e il peso Y espresso in kg.

Soluzioni

1. (a) Lo spazio campionario e' partizionato in tre eventi A = cliente ad alto rischio, M = cliente a medio rischio e B = cliente a basso rischio. Quindi si considera l'evento I = il cliente scelto ha almeno un incidente nell'anno. Pertanto per la formula delle probabilita' totali

$$\begin{aligned} P(I) &= P(I \cap A) + P(I \cap M) + P(I \cap B) \\ &= P(I \mid A)P(A) + P(I \mid M)P(M) + P(I \mid B)P(B) \\ &= 0.25 \times 0.2 + 0.16 \times 0.3 + 0.1 \times 0.5 = 0.148 \end{aligned}$$

(b) Si puo' usare la formula di Bayes:

$$P(A \mid I) = \frac{P(I \cap A)}{P(I)} = \frac{0.05}{0.148} = 0.3378$$

2. (a) $F(-\infty) = 0$. Il limite per $x \to \infty$ e' inoltre uguale a 1. Inoltre la funzione e' non decrescente perche' se $a \leq b$

$$0 \leq e^{-b^2} \leq e^{-a^2} \leq 1.$$

(b) La densita' e' 0 per $x < 0$ e

$$f(x) = \frac{d}{dx}\{1 - e^{-x^2}\} = 2xe^{-x^2}$$

per $x \geq 0$. (c) La probabilita' cercata e'

$$P(X > 2) = 1 - F(2) = 1 - (1 - e^{-4}) = 1/e^4 = 1.8\%.$$

3. (a) La covarianza tra due variabili discrete Xe Y aventi funzione di probabilita' congiunta $p(x, y)$ e valori attesi $\mu_x = EX$ e $\mu_y = EY$ e'

$$\text{cov}(X, Y) = \sum_x \sum_y (x - \mu_x)(y - \mu_y)p(x, y) = E\{(Y - EX)(Y - EY)\}.$$

(b) L'indice misura il grado di associazione lineare tra X e Y. (c) Puo' assumere valori compresi nell'intervallo

$$-\sigma_x\sigma_y \leq \text{cov}(X, Y) \leq \sigma_x\sigma_y$$

dove σ_x e σ_y sono gli scarti quadratici medi (le radici quadrate delle varianze) di X e Y rispettivamente. Se e' positiva, c'e' concordanza tra X e Y (a scarti di un segno dalla media di X si accompagnano scarti dello stesso segno dalla media di Y). Se e' negativa, c'e' discordanza (a scarti di un segno dalla media di X si accompagnano scarti di segno opposto dalla media di Y). (d) La covarianza e' uguale al prodotto $\sigma_x\sigma_y$ se e solo se $Y = a + bX$ con $b > 0$. La covarianza e' uguale a $-\sigma_x\sigma_y$ se e solo se $Y = a + bX$ con $b < 0$.

(e) Si tratta di calcolare

$$\text{cov}(X/100, Y) = \sum_x \sum_y (x/100 - \mu_x/100)(y - \mu_y)p(x, y) = 1/100 \times \text{cov}(X, Y) = .075$$