Metodi Matematici e Statistici

Prova scritta – 17/2/2009

Le domande valgono 3 punti ciascuna.

1. Data $X \sim N(-6, 4)$, trovare un numero λ tale che $P(X < \lambda) = 0.9$.

2. Data $N \sim \mathcal{P}(\lambda)$, trovare per quali valori di λ vale $E[N^2 - N] > 5$.

3. Una staffetta di 4 corridori impiega complessivamente un tempo T a coprire il percorso. Il tempo T è la somma di quattro tempi esponenziali indipendenti T_1, T_2, T_3, T_4, di media 12 secondi. Calcolare media e varianza di T.

4. Sia X una v.a. discreta tale che $P(X = 0) = 0.1$, $P(X = 1) = 0.1$, $P(X = 2) = 0.8$. Calcolare $E\left[\frac{X}{1+X^2}\right]$.

5. Se X, Y sono gaussiane canoniche indipendenti, calcolare $E\left[X^2 e^{2Y}\right]$.

6. Date due binomiali indipendenti $N_1, N_2 \sim B\left(2, \frac{1}{3}\right)$, calcolare $P(N_1 = N_2)$.

7. Una cartoleria vende ogni giorno un numero aleatorio N di penne, $N \sim \mathcal{P}(4)$. Il rifornimento arriva solo una volta ogni 25 giorni lavorativi. La cartoleria pensa che un rifornimento di 120 penne sia sufficiente. Con che probabilità lo è?

8. In una zona del mondo le persone di età inferiore ai 50 anni sono il doppio di quelle di età superiore. Una certa malattia legata all'età colpisce il 2% della popolazione sotto i 50 anni, mentre il 7% di quella superiore ai 50 anni. In quella zona del mondo, che percentuale di persone soffre di quella malattia?

9. Vorremmo stimare la media di una v.a. di Poisson. Sappiamo che tale media è compresa tra 2 e 6. Vorremmo poter determinare quanti esperimenti sono necessari per avere una precisione pari ad 1, a livello 95%. Discutere questo problema non banale.

10. I passeggeri che viaggiano intorno alle ore 24 tra Firenze e Pisa in certi giorni dell'anno vengono contati, per 10 giorni. I valori sono

$$17, 13, 9, 23, 14, 12, 16, 11, 14, 21.$$

Un certo servizio ferroviario viene cancellato se il numero medio di passeggeri è inferiore a 20. A livello di significatività 95%, c'è ragione di cancellare quel servizio serale?

1 Soluzioni

1.
$$\lambda = -6 + 2q_{0.9} = -6 + 2 \cdot 1.28 = -3.44.$$

2.
$$\begin{aligned} E\left[N^2 - N\right] &= Var\left[N\right] + E\left[N\right]^2 - E\left[N\right] \\ &= \lambda + \lambda^2 - \lambda = \lambda^2 > 5 \end{aligned}$$

se e solo se $\lambda > \sqrt{5}$.

3. $T = T_1 + T_2 + T_3 + T_4$,
$$E\left[T\right] = 4E\left[T_1\right] = 48$$
$$Var\left[T\right] \overset{\text{per l'indipendenza}}{=} 4Var\left[T_1\right] = 576.$$

4. $\frac{X}{1+X^2} = 0$ con prob. 0.1, $\frac{1}{2}$ con prob. 0.1, $\frac{2}{5}$ con prob. 0.8, quindi
$$E\left[\frac{X}{1+X^2}\right] = \frac{1}{2} \cdot 0.1 + \frac{2}{5} \cdot 0.8 = 0.37$$

5.
$$\begin{aligned} E\left[X^2 e^{2Y}\right] &= E\left[X^2\right] E\left[e^{2Y}\right] = \left(Var\left[X\right] + E\left[X\right]^2\right) \phi_Y\left(2\right) \\ &= e^2 = 7.389. \end{aligned}$$

6. $P\left(N_1 = 0\right) = \left(\frac{2}{3}\right)^2$, $P\left(N_1 = 1\right) = 2\frac{1}{3}\frac{2}{3}$, $P\left(N_1 = 2\right) = \left(\frac{1}{3}\right)^2$,
$$\begin{aligned} P\left(N_1 = N_2\right) &= P\left(N_1 = N_2 | N_2 = 0\right) P\left(N_2 = 0\right) \\ &\quad + P\left(N_1 = N_2 | N_2 = 1\right) P\left(N_2 = 1\right) \\ &\quad + P\left(N_1 = N_2 | N_2 = 2\right) P\left(N_2 = 2\right) \end{aligned}$$

$$\begin{aligned} &= P\left(N_1 = 0\right) P\left(N_2 = 0\right) + P\left(N_1 = 1\right) P\left(N_2 = 1\right) \\ &+ P\left(N_1 = 2\right) P\left(N_2 = 2\right) \\ &= P\left(N_1 = 0\right)^2 + P\left(N_1 = 1\right)^2 + P\left(N_1 = 2\right)^2 \\ &= \left(\frac{2}{3}\right)^4 + \left(2\frac{1}{3}\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^4 = 0.407. \end{aligned}$$

7. Indichiamo le vendite di 25 giorni con $N_1, ..., N_{25}$. Dobbiamo calcolare

$$
\begin{aligned}
& P\left(N_1+...+N_{25} \leq 120\right) \\
& = P\left(\frac{N_1+...+N_{25}-25E\left[N_1\right]}{\sqrt{Var\left[N_1\right]}\sqrt{25}} \leq \frac{120-25E\left[N_1\right]}{\sqrt{Var\left[N_1\right]}\sqrt{25}}\right) \\
& \sim P\left(Z \leq \frac{120-25E\left[N_1\right]}{\sqrt{Var\left[N_1\right]}\sqrt{25}}\right) = \Phi\left(\frac{120-25E\left[N_1\right]}{\sqrt{Var\left[N_1\right]}\sqrt{25}}\right) \\
& = \Phi\left(\frac{120-25\cdot 4}{2\cdot\sqrt{25}}\right) = \Phi\left(2\right) = 0.977
\end{aligned}
$$

8. A = inferiore a 50, B = superiore a 50, M = malattia:

$$
\begin{aligned}
P\left(M\right) &= P\left(M|A\right)P\left(A\right)+P\left(M|B\right)P\left(B\right) \\
&= 0.02\cdot\frac{2}{3}+0.07\cdot\frac{1}{3} = 0.036.
\end{aligned}
$$

9. Se si trattasse di una gaussiana, la formula sarebbe (partendo da $\frac{\sigma q_{0.975}}{\sqrt{n}} \leq 1$)

$$n \geq \left(\sigma q_{0.975}\right)^2.$$

Per il teorema limite centrale, le formule sono approssimativamente valide anche senza ipotesi di gaussianità, se la numerosità è abbastanza elevata (cosa che qui è deducibile solo a posteriori). Circa σ, sappiamo che la media della Poisson è compresa tra 2 e 6. Quindi la numerosità richiesta deve essere (approssimativamente)

$$n \geq 6\cdot q_{0.975} = 6\cdot 1.96 = 11.76$$

quindi servono 12 esperimenti. Il risultato è molto approssimato in quanto questo numero è basso e quindi l'uso stesso dell'approssiamazione gaussiana non è ottimale.

10. La media campionaria è 15, la deviazione 4.372 circa, la numerosità 10. Ipotizziamo che il numero sia ≥ 20 ed eseguiamo un test unilaterale al 95%. Calcoliamo

$$z = \frac{15-20}{4.372}\sqrt{10} = -3.6165$$

e confrontiamolo con

$$-q_{0.95} = 1.64.$$

Il valore di z supera il quantile, quindi il test è significativo. Possiamo rifiutare l'ipotesi, cioè ritenere che il servizio debba essere cancellato.