Statistica I. Ingegneria Gestionale. II° Scritto del 21/12/2011

Cerchiare, su questo foglio, le risposte corrette e risolvere per esteso gli esercizi sui fogli assegnati.

Esercizio 1. Un ingegnere gestionale di un'azienda che produce veicoli e macchine per scavi, esamina le vendite delle ultime settimane. Le vendite di certe macchine da scavo è sporadica e decide di descriverla con una v.a. discreta. Invece le vendite di veicoli più leggeri da trasporto è più frequente e quindi la descrive con una gaussiana.

i) (4 pt) Le macchine da scavo vendute nelle ultime 10 settimane sono 2, 3, 1, 2, 4, 3, 1, 2, 1, 3. Se si escludono non più di 20 settimane su 100, al massimo quante macchine si aspetta di venderne in futuro, settimanalmente? Scegliere una v.a. discreta ragionevole per questo problema, tra Poisson e binomiale, motivando la scelta.

$$, \quad , \quad , \quad , \quad 3.$$

ii) (3 pt) Supponiamo ora di sapere che la vendita di tali macchine avviene ad aziende ben note a priori. Ci sono 50 aziende di quel tipo. Supponendole indipendenti e con la stessa probabilità p di comprare un macchina da scavo in una settimana, quanto stimate che valga p? Descrivere in dettaglio il modello matematico utilizzato.

$$, \quad , \quad , \quad , \quad 0.044.$$

iii) (4 pt) Riguardo ai veicoli da trasporto, i dati delle ultime 9 settimane forniscono media 14.2 e deviazione 5.7. Se si esclude una settimana su 20, in che intervallo simmetrico rispetto alla media cadono le vendite settimanali? Spiegare il calcolo anche tramite un disegno.

$$, \quad , \quad [3.028, 25.372] .$$

iv) (3 pt) Inoltre, con un livello di fiducia pari al 95%, in che intervallo cade il valore vero della media settimanale delle vendite di veicoli?

$$, \quad , \quad [10.476, 17.924] .$$

v) (4 pt) Tenendo conto di questa incertezza, accettando un rischio del 10%, volendo avere settimanalmente un magazzino in grado di servire tutte le richieste, qual'è il numero di veicoli che si deve tenere? Oltre a dare il risultato, spiegare bene il calcolo eseguito ed il significato del risultato.

$$, \quad , \quad , \quad , \quad 25.22.$$

vi) (4 pt) L'anno dopo sembra che il mercato dei veicoli da trasporto sia più vivace. Non sappiamo se le vendite medie siano rimaste le stesse dell'anno precedente o siano aumentate (diminuite non sembra realistico). Per capire se sono davvero aumentate, vengono registrati i dati delle vendite relativi a 12 settimane, che forniscono una media pari a 18.4 e viene calcolata la probabilità che uscisse un valore così estremo per puro caso. Che valore si trova? Partire scrivendo con precisione le v.a. esaminate e scrivendo letteralmente la probabilità richiesta.

, , , , 0.006.

Esercizio 2. i) (2 pt) Calcolare la funzione generatrice nel punto $t = \log_e 2$ della v.a. X^2, dove $X \sim B(1, 0.3)$.

, , , , 1. 3.

ii) (3 pt) Calcolare $P(X_1 + X_2 \leq 1)$ dove X_1 e X_2 sono copie indipendenti di X.

, , , , 0.91.

iii) (3 pt) Se $Y \sim N(0, 9)$ è indipendente da X, calcolare $P(XY < 1)$.

, , , , 0.889.

Quantili utili per questo compito:

$x =$	-1.654	$1/3$	0.842	1.28	1.96	2.552	2.9	3.297	1.96
$\Phi(x) =$	0.049	0.6305	0.8	0.9	0.975	0.994	0.9981	0.9995	0.975

Formule e teoremi potenzialmente utili. 1) Se X è $B(n, p)$, allora $P(N = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, $E[X] = np$, $Var[X] = np(1 - p)$; se X è $\mathcal{P}(\lambda)$, allora $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $E[X] = Var[X] = \lambda$; se X è esponenziale di parametro λ, $F(x) = 1 - e^{-\lambda x}$, $E[X] = \sigma[X] = \frac{1}{\lambda}$. 2) Se $X_1, ..., X_n$ sono $N(\mu, \sigma^2)$, allora $P\left(\left|\overline{X} - \mu\right| \leq \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1 - \alpha$. 3) Se $X_1, ..., X_n$ sono $N(\mu, \sigma^2)$, allora $Z = \frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}$ è una $N(d, 1)$ con $d = \frac{\mu - \mu_0}{\sigma} \sqrt{n}$.

1 Soluzioni

Esercizio 1. i) Usiamo una Poisson, non avendo dati che permettano l'uso delle binomiali. La media del campione è 2.2, che quindi viene presa come stima del parametro λ della Poisson (che ne è la media teorica). Cerchiamo allora il numero "massimo" m di macchine vendute settimanalmente all'80% (20 settimane su 100), cioè detto N la v.a. che conta il numero di macchine vendute settimanalmente, cerchiamo m tale che $P(N \leq m) \geq 0.8$. Troviamolo per tentativi, ricordando che N è una $\mathcal{P}(2.2)$, quindi $P(N = k) = e^{-2.2}\frac{2.2^k}{k!}$. Vale

$$P(N \leq 2) = e^{-2.2}\left(\frac{2.2^0}{0!} + \frac{2.2^1}{1!} + \frac{2.2^2}{2!}\right) = 0.622$$
$$P(N \leq 3) = e^{-2.2}\left(\frac{2.2^0}{0!} + \frac{2.2^1}{1!} + \frac{2.2^2}{2!} + \frac{2.2^3}{3!}\right) = 0.819$$

quindi il numero "massimo" è 3.

ii) Indichiamo con $X_1, ..., X_{50}$, delle Bernoulli di parametro p che indicano se la rispettiva azienda compra una macchina. Il numero N di macchine vendute è la somma $X_1 + ... + X_{50}$ che, per un noto teorema, è una binomiale $B(50, p)$. La media è $50p$. Abbiamo, dai dati, una media stimata pari a 2.2 per cui risulta $p = \frac{2.2}{50} = 0.044$.

iii) Le spiegazioni sono come nei compiti precedenti. Una settimana su 20 significa che va esclusa una probabilità $\alpha = \frac{1}{20} = 0.05$, che va distribuita metà da una parte e metà dall'altra, quindi, detti $a < b$ gli estremi di tale intervallo, vale

$$a = \mu - \sigma q_{0.975} = 14.2 - 5.7 \cdot 1.96 = 3.028$$
$$a = \mu + \sigma q_{0.975} = 14.2 + 5.7 \cdot 1.96 = 25.372$$

iv) Questo invece è l'intervallo di confidenza della media, quindi

$$\mu = 14.2 \pm \frac{5.7 \cdot 1.96}{\sqrt{9}} = 14.2 \pm 3.724.$$

v) La media vera non è 14.2 ma, al 95% (cioè con livello di fiducia 95% che ciò che stiamo per dire sia corretto) è compresa tra 10.476 e 17.924. Nell'ottica della scorta di magazzino, la situazione peggiore è che sia $\mu = 17.924$. Se questo accade, la quatità da tenere, con rischio 10%, è

$$\lambda = 17.924 + 5.7 \cdot q_{0.9} = 17.924 + 5.7 \cdot 1.28 = 25.22.$$

Abbiamo una confidenza pari al 95% che, a meno di un rischio del 10%, questa quantità sia sufficiente.

vi) Dobbiamo supporre che non sia cambiata la media, cioè sia ancora 14.2, e la deviazione sia 5.7, però un campionamento di numerosità 12 fornisca un valore medio superiore a 18.4. Indichiamo con $X_1, ..., X_{12}$ i valori che escono da questo campionamento (che, ripetiamo, stiamo assumendo essere di tipo $N\left(14.2, 5.7^2\right)$), con $\overline{X}$ la loro media aritmentica, con Z la stessa quantità standardizzata, cioè $Z = \frac{\overline{X}-14.2}{5.7}\sqrt{12}$, che è una $N(0,1)$, con z il valore sperimentale di questa variabile, cioè $z = \frac{18.4-14.2}{5.7}\sqrt{12} = 2.552$. Dobbiamo calcolare

$$P(Z > z) = 1 - P(Z \leq z) = 1 - \Phi(2.552) = 1 - 0.994 = 0.006.$$

Esercizio 2. i) La v.a. X^2 è ancora $B(1, 0.3)$ quindi $\varphi(t) = 0.3 \cdot e^t + 0.7$. Pertanto $\varphi(\log_e 2) = 0.3 \cdot 2 + 0.7 = 1.3$.

ii) Per un noto teorema, $X_1 + X_2$ è una $B(2, 0.3)$, quindi

$$\begin{aligned} P(X_1 + X_2 \leq 1) &= P(X_1 + X_2 = 0) + P(X_1 + X_2 = 1) = 1 - P(X_1 + X_2 = 2) \\ &= 1 - \binom{2}{2} 0.3^2 (1 - 0.3)^0 = 0.91. \end{aligned}$$

iii)

$$\begin{aligned} P(XY < 1) &= P(XY < 1 | X = 0) P(X = 0) + P(XY < 1 | X = 1) P(X = 1) \\ &= 0.7 \cdot P(0 \cdot Y < 1) + 0.3 \cdot P(1 \cdot Y < 1) \\ &= 0.7 \cdot 1 + 0.3 \cdot \Phi\left(\frac{1}{3}\right) = 0.7 \cdot 1 + 0.3 \cdot 0.6305 = 0.889. \end{aligned}$$