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

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

Esercizio 1. Nel 2006, un giornalaio vendeva in media 76.3 quotidiani al giorno, con una deviazione standard pari a 15.6. Nel 2007 entra in gioco un concorrente nella stessa zona.

i) (4 pt) Per capire se questo ha modificato la media delle vendite, il giornalaio registrerà i valori delle vendite per n giorni ed eseguirà: A = un test bilaterale; B = un test unilaterale destro; C = un test unilaterale sinistro:

$$..., \quad ..., \quad C$$

Motivare la scelta tra A, B e C, descrivere il test e dire su che base potrebbe scegliere il valore di n.

ii) (4 pt) Scelto $n = 20$, registrati i valori per 20 giorni, riscontra una media aritmetica giornaliera pari a $\overline{x} = 64.8$. Dire com'è definito il p-value di questo test e, partendo dalla definizione data, calcolare quanto vale. Commentare il risultato dal punto di vista applicativo:

, , , , 0.0005.

iii) (4 pt) Se il valore medio vero delle vendite fosse sceso a 60, eseguendo un test al 90%, con che probabilità ce ne accorgeremmo? Scrivere matematicamente la probabilità richiesta e, partendo da lì, ricavare il risultato, usando solamente le formule riportate sotto.

, , , , 0.9996.

iv) (2 pt) Siamo ormai nel 2007, quindi tralasciamo il problema del confronto col 2006. Quanto è precisa la stima della media di vendite giornaliere, ottenuta dai dati raccolti sopra, al 95%?

, , , , 6.837.

v) (4 pt) Con questa stima, quanto può essere sicuro di vendere giornalmente, al 90% (la preoccupazione riguarda il valore minimo delle vendite)? Spiegare il calcolo tramite un disegno.

, , , , 44.832.

vi) (4 pt) Riguardo invece alle vendite di una rivista sulle motociclette, le vendite giornaliere sono poche e meglio descrivibili con una v.a. di Poisson di parametro 3. Per quale percentuale di giorni lavorativi vende almeno 2 di queste riviste? Oltre a trovare la soluzione numerica dell'esercizio, spiegare su che base si è tradotto il concetto di "percentuale di giorni lavorativi" nella quantità calcolata.

, , , , 0.801.

Esercizio 2. Nel seguito, X è gaussiana $N(4,25)$, Y è Bernoulli $B(1,0.1)$. Si suppongano X e Y indipendenti.

i) (2 pt) Calcolare $P(0<X<8)$

, , 0.576, ..., ...

ii) (3 pt) Calcolare $P(X<4Y)$

, , 0.241, ..., ...

iii) (3 pt) Calcolare $E\left[YX^2e^Y\right]$.

, , 11.145, ..., ...

Quantili utili per questo compito:

$x=$	-2.582	-1.654	0.8	1.28	1.85	1.96	2.537	3.297	3.392
$\Phi(x)=$	0.0049	0.049	0.788	0.9	0.967	0.975	0.994	0.9995	0.9996

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 $P\left(\left|\overline{X}-\mu\right|\le\frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)=1-\alpha$. 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) C, in quanto immagina che la variazione della media sia in negativo, se c'è stata, ed in tal caso è più facile che il test risulti significativo (potenza maggiore). L'ipotesi nulla è che la media sia 76.3. Quando avrà i dati $x_1, ..., x_n$, scelta la significatività α, dovrà calcolare $\overline{x}$, $z = \frac{\overline{x}-76.3}{15.6}\sqrt{n}$, $q_{1-\alpha}$ e vedere se $z < -q_{1-\alpha}$ (in tal caso rifiuterà l'ipotesi nulla, cioè affermerà che la concorrenza ha diminuito la sua vendita media). Specificando una certa potenza relativamente ad un valore della media μ che gli sembri critico, da scoprire se valido, può trovare n che permette di avere tale potenza.

ii) Il p-value è la probabilità che l'indicatore $Z = \frac{\overline{X}-76.3}{15.6}\sqrt{20}$, sotto l'ipotesi nulla (secondo tale ipotesi Z è $N(0,1)$), assuma valori più estremi del valore sperimentale z,

$$z = \frac{64.8 - 76.3}{15.6}\sqrt{20} = -3.297$$

quindi

$$p := P(Z \leq -3.297) = \Phi(-3.297) = 1 - \Phi(3.297) = 0.0005.$$

E' molto piccolo, quindi il giornalaio potrà affermare con elevata significatività che l'entrata in gioco del concorrente ha avuto effetto.

iii) Dobbiamo calcolare $P_{\mu=60}(Z \leq -q_{0.9})$ dove $Z = \frac{\overline{X}-76.3}{15.6}\sqrt{20}$. Risulta

$$Z \sim N(d, 1), \qquad d = \frac{60 - 76.3}{15.6}\sqrt{20} = -4.673$$

in quanto la media vera ora è 60, $\overline{X}$ proviene da una $N(60, 15.6)$. Vale

$$P(Z \leq -q_{0.9}) = \Phi(-q_{0.9} - d) = \Phi(-1.28 + 4.673) = \Phi(3.393) = 0.9996.$$

iv) Per la teoria dell'intervallo di confidenza, la precisione δ è data da

$$\delta = \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

che approssimiamo a ($\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $1 - \frac{\alpha}{2} = 0.975$)

$$\delta = \frac{15.6 \cdot 1.96}{\sqrt{20}} = 6.837.$$

v) Si deve tracciare una gaussiana X di media 64.8 e deviazione 15.6, e cercare il valore λ per cui $P(X > \lambda) = 0.9$: affermeremo che siamo sicuri di vendere più di λ. Il punto λ è qundi alla sinistra di 64.8. Tracciando anche la gaussiana canonica per confronto, in cui l'analogo di λ sarebbe il quantile $q_{0.1} = -q_{0,9} = -1.28$, si riconosce che

$$\lambda = 64.8 - 15.6 \cdot 1.28 = 44.832.$$

vi) Dobbiamo calcolare la probabilità che sia $X \geq 2$ dove X è $\mathcal{P}(3)$, cioè $P(X = k) = e^{-3}\frac{3^k}{k!}$, quindi

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - e^{-3}\frac{3^0}{0!} - e^{-3}\frac{3^1}{1!} = 0.801.$$

Il numero $P(X \geq 2)$ è cira uguale alla percentuale di giorni in cui accade $X \geq 2$ a causa della legge dei grandi numeri. Infatti, introducendo delle v.a. $Y_1, Y_2, ...$ di Bernoulli che valgono 1 se $X \geq 2$ e zero altrimenti, che corrispondano ai diversi giorni, la LGN dice che $\frac{Y_1+...+Y_n}{n} \to E[Y_1] = P(X \geq 2)$ e $\frac{Y_1+...+Y_n}{n}$ rappresent la percentuale di giorni in cui avviene $X \geq 2$.

Esercizio 2. i)

$$P(0 < X < 8) = \Phi\left(\frac{8-4}{5}\right) - \Phi\left(\frac{0-4}{5}\right) = \Phi(0.8) - \Phi(-0.8) = 0.788 - 1 + 0.788 = 0.576.$$

ii)

$$\begin{aligned} P(X < 4Y) &= P(X < 4Y|Y = 0)\,P(Y = 0) + P(X < 4Y|Y = 1)\,P(Y = 1) \\ &= 0.9 \cdot P(X < 0) + 0.1 \cdot P(X < 4) \\ &= 0.9 \cdot \Phi\left(\frac{0-4}{5}\right) + 0.1 \cdot \Phi\left(\frac{4-4}{5}\right) \\ &= 0.9 \cdot \Phi(-0.8) + 0.1 \cdot \Phi(0) = 0.9 \cdot (1 - 0.788) + 0.1 \cdot 0.5 = 0.241. \end{aligned}$$

iii)

$$\begin{aligned} E\left[YX^2e^Y\right] &= E\left[X^2\right]E\left[Ye^Y\right] = \left(Var[X] + \mu_X^2\right)\left(0 \cdot e^0 P(Y = 0) + 1 \cdot e^1 P(Y = 1)\right) \\ &= (25 + 16)(e \cdot 0.1) = 11.145. \end{aligned}$$