Statistica I. Ingegneria Gestionale. Scritto dell'8/06/2010

Esercizio 1. Un negozio, a consuntivo di un anno di lavoro, calcola che la vendita media settimanale di un certo prodotto è pari a 21.5 con deviazione standard 3.5. Non troppo soddisfatto dell'andamento, effettua della pubblicità specifica per quel prodotto, per 20 settimane, registrando le vendite settimanali: registra i valori $v_1, ..., v_{20}$, tali che $v_1 + ... + v_{20} = 465$, $v_1^2 + ... + v_{20}^2 = 11064$. Ad occhio il valore è sopra la media.

i) Al massimo, con quale significatività può ritenere che la pubblicità abbia fatto aumentare le vendite, piuttosto che ritenere che l'aumento sia puramente causale?

$$0.012\,9, \qquad ...$$

Spiegare che ragionamenti sono stati svolti.

ii) Supponiamo che il vero valore della vendita media sia passato da 21.5 a 24. Con che probabilità si accorgerebbe che c'è stato un aumento, seguendo la metodologia del punto precedente, relativamente ad $\alpha = 0.05$?

$$0.9382, \qquad ...$$

Svolgere per esteso, spiegando i ragionamenti principali.

iii) Decidiamo che, a causa della pubblicità, siamo in un nuovo regime di vendite, di cui il campione $v_1, ..., v_{20}$ è rappresentativo per stimare media e deviazione standard. Sulla base di questi dati, che probabilità c'è di vendere meno di 20 alla settimana?

$$0.186\,7, \qquad ...$$

iv) Abbiamo il dubbio che solo 20 rilevazioni siano poche per stimare bene la nuova vendita media settimanale, al 95%. Vorremmo un errore relativo del 10% al massimo. Quante rilevazioni servono?

$$10, \qquad ...$$

v) Un dipendente ha un'intuizione: gli sembra che le vendite del prodotto fino ad ora esaminato vadano di pari passo con le vendite di un secondo prodotto, di cui disponiamo delle vendite $w_1, ..., w_{20}$ relative alle stesse 20 settimane dell'altro. Sapendo che $w_1 + ... + w_{20} = 149$, $w_1^2 + ... + w_{20}^2 = 1420$, $v_1 w_1 + ... + v_{20} w_{20} = 3702$, che correlazione trova tra i due?

$$0.849\,44, \qquad ...$$

Esercizio 2. i) Se $X \sim N(-2, 4)$, trovare λ tale che $P(X < \lambda) = 0.9$.

$$0.56, \qquad ...$$

ii) Se inoltre $Y \sim B(1, 0.2)$ ed X, Y sono indipendenti, calcolare $P(XY > 0.5)$.

$$0.021\,12, \qquad ...$$

iii) Se infine $Z \sim N(5, 9)$ e sono indipendenti, calcolare $E\left[Y^2 e^{-Z}\right]$.

$$0.121\,31, \qquad ...$$

iv) Calcolare poi $Var\left[\log_e(1 + Y \cdot (e-1))\right]$.

$$0.16, \qquad ...$$

v) Se $Y_1, ..., Y_{30}$ sono distribuite come Y e sono indipendenti, calcolare $P(Y_1 + ... + Y_{30} > 8)$.

$$0.181\,4, \qquad ...$$

1 Soluzioni

Esercizio 1. $v_1 + ... + v_{20} = 465$, $v_1^2 + ... + v_{20}^2 = 11064$

$$\mu_0 = 21.5, \qquad \sigma = 3.5$$

$$\overline{v} = \frac{465}{20} = 23.25, \qquad S^2 = \frac{1}{19}\left(11064 - 20 \cdot (23.25)^2\right) = 13.303$$

quindi $S = 3.6473$.

i) Si presuppone che la pubblicità aumenti le vendite quindi si effettua un test unilaterale, del tipo

$$\frac{\overline{v} - \mu_0}{\sigma}\sqrt{n} > q_{1-\alpha}$$

La domanda chiede il p-value, quindi la probabilità di avere valori più estremi, quindi, essendo

$$\frac{\overline{v} - \mu_0}{\sigma}\sqrt{n} = \frac{23.25 - 21.5}{3.5}\sqrt{20} = 2.2361$$

calcoliamo

$$P(Z > 2.2361) = 1 - \Phi(2.2361) = 1 - 0.9871 = 0.0129.$$

ii) Si tratta di calcolare la potenza. Sulla base del solito disegno (qui non riportato) la potenza è la probabilità, secondo la gaussiana $\overline{X} \sim N\left(24, \frac{3.5^2}{20}\right)$, di cadere nell'intervallo che produce il rifiuto dell'ipotesi $\mu_0 = 21.5$. Tale intervallo è, per quanto appena visto, $[21.5 + \frac{3.5 \cdot q_{0.95}}{\sqrt{20}}, \infty)$. Quindi, usando $q_{0.95} = 1.65$,

$$\begin{aligned} P\left(\overline{X} \geq 21.5 + \frac{3.5 \cdot q_{0.95}}{\sqrt{20}}\right) &= 1 - \Phi\left(\frac{21.5 - 24}{3.5}\sqrt{20} + 1.65\right) \\ &= 1 - \Phi(-1.5444) = \Phi(1.5444) = 0.9382. \end{aligned}$$

iii)

$$\begin{aligned} P(V < 20) &= \Phi\left(\frac{20 - \overline{v}}{S}\right) = \Phi\left(\frac{20 - 23.25}{3.6473}\right) = \Phi(-0.89107) \\ &= 1 - \Phi(0.89107) = 1 - 0.8133 = 0.1867. \end{aligned}$$

iv) L'errore assoluto è

$$\delta = \frac{3.6473 \cdot 1.96}{\sqrt{n}}$$

quindi quello relativo, usando come valore della media 23.25,

$$\frac{\delta}{23.25} = \frac{3.6473 \cdot 1.96}{23.25\sqrt{n}}$$

che deve essere al massimo 0.1, quindi

$$\sqrt{n} \geq \frac{3.6473 \cdot 1.96}{23.25 \cdot 0.1} = 3.0747$$

ovvero $n \geq 9.4538$, quindi $n = 10$.

v) $w_1 + ... + w_{20} = 149$, $w_1^2 + ... + w_{20}^2 = 1420$, $v_1 w_1 + ... + v_{20} w_{20} = 3702$,

$$\overline{w} = \frac{149}{20} = 7.45, \qquad S_W^2 = \frac{1}{19}\left(1420 - 20 \cdot (7.45)^2\right) = 16.313$$

$$S_W = \sqrt{16.313} = 4.0389$$

$$r^2 = \frac{\frac{1}{n-1}\sum (v_i - \overline{v})(w_i - \overline{w})}{S_V S_W} = \frac{\frac{1}{19}(3702 - 20 \cdot 23.25 \cdot 7.45)}{3.6473 \cdot 4.0389} = 0.84944.$$

Esercizio 2. i)

$$\lambda = -2 + 2 \cdot q_{0.9} = -2 + 2 \cdot 1.28 = 0.56$$

ii)

$$\begin{aligned} P(XY > 0.5) &= P(Y = 1, X > 0.5) = P(Y = 1) P(X > 0.5) \\ &= 0.2 \cdot \left(1 - \Phi\left(\frac{0.5 - (-2)}{2}\right)\right) = 0.2 \cdot (1 - \Phi(1.25)) \\ &= 0.2 \cdot (1 - 0.8944) = 0.02112. \end{aligned}$$

iii)

$$E\left[Y^2 e^{-Z}\right] = E\left[Y^2\right] E\left[e^{-Z}\right] = \left(Var[Y] + \mu_Y^2\right) \phi_Z(-1)$$

$$\left(0.2 \cdot 0.8 + (0.2)^2\right) e^{-5+\frac{9 \cdot 1}{2}} = 0.12131.$$

iv)

$$\log_e(1 + X \cdot (e-1)) = 0 \text{ se } X = 0, \ = 1 \text{ se } X = 1.$$

Quindi $\log_e(1 + X \cdot (e-1))$ è una $B(1, 0.2)$, la cui varianza è $0.2 \cdot 0.8 = 0.16$.

v)

$$\begin{aligned} P(Y_1 + ... + Y_{30} > 8) &= P\left(\frac{Y_1 + ... + Y_{30} - 30 \cdot 0.2}{\sqrt{30 \cdot 0.16}} > \frac{8-6}{\sqrt{30 \cdot 0.16}}\right) \\ &= 1 - \Phi\left(\frac{8-6}{\sqrt{30 \cdot 0.16}}\right) = 1 - \Phi(0.91287) = 1 - 0.8186 = 0.1814. \end{aligned}$$