Statistica I. Ingegneria Gestionale. Compitino del 28/05/2010

Esercizio. All'ufficio che riceve e tiene nota degli ordini di un certo materiale, viene chiesto di fare delle analisi statistiche sulla quantità Q, di quel materiale, ordinata settimanalmente. L'ufficio utilizza la registrazione degli ordini relativamente a 52 settimane. Dette $q_1, ..., q_{52}$ le quantità settimanali ordinate in tali 52 settimane, si sappia che $q_1 + ... + q_{52} = 287.8$, $q_1^2 + ... + q_{52}^2 = 1782.5$.

1) Stimare la quantità *media* settimanale, dichiarando l'errore assoluto al 95% (usare per semplicità i quantili gaussiani).

$$5.5346 \pm 0.524, \quad 5.5346 \pm 0.837, \quad 5.9246 \pm 0.524, \quad 5.9246 \pm 0.837$$

2) Prevedere la *probabilità* di ricevere, la prossima settimana, ordini per una quantità maggiore di 8.

0.102, 0.203, 0.304, 0.405

3) Volendo avere già in magazzino l'occorrente per accontentare tutte le ordinazioni con probabilità 0.85, *quanto* si deve tenere?

7.54, 5.74, 3.45, 9.82

4) Lo studente avrà fino ad ora tacitamente assunto che Q sia gaussiana. Se invece decidesse che Q è esponenziale, come risponderebbe alla domanda 2?

0.2356, 0.1653, 0.3287, 0.4728

5*) Abbiamo così sviluppato due modelli per Q, uno gaussiano ed uno esponenziale. L'anno dopo, osservando gli ordini per 20 settimane, si riscontra per 6 volte un valore della quantità ordinata maggiore di 8. Al 95% ritenete che il modello gaussiano sia ancora valido? E quello esponenziale? Quale vi sembra il migliore?

Gaussiano valido: sì, no. Esponenziale valido: sì, no.

Eventuali osservazioni:

6) Se in una settimana la quantità ordinata è ≤ 2, quella settimana viene segnata come critica. Su 10 settimane, che probabilità c'è di avere almeno una

settimana critica? Usare la Q gaussiana.

0.2895, 0.1496, 0.3225, 0.4852

7) Su 100 settimane, che probabilità c'è di averne più di 7 critiche? Usare la Q gaussiana.

0.0217, 0.1023, 0.0823, 0.0012

8*) Supponiamo che si accetti la gaussianità di Q ma che, dopo due anni, si abbia il dubbio che a causa della crisi economica ci sia stato un calo sistematico di ordinazioni. Se esse avessero una media settimanale pari o inferiore a 4, questo obbligherebbe ad una riconversione della produzione. Si decide di esaminare Q per due mesi, cioè 9 settimane. Useremo significatività 95%. Che potenza avrà il test che eseguiremo, pensando alla sua capacità di metterci in allarme relativamente al pericolo di riconversione? Quante settimane servirebbero per avere una potenza del 90%?

Descrivere concisamente il test che si vuole eseguire, ed i ragionamenti e calcoli per trovare la potenza:

H1) Si registrano poi i valori $q_1, ..., q_9$. Essi forniscono $q_1 + ... + q_9 = 39.5$. Che valore p-dei-dati troviamo?

0.0375, 0.0057, 0.0182, 0.0728

H2) Metà degli ordini vengono dall'estero, l'altra metà dall'Italia. Gli ordini dall'estero portano ad un guadagno maggiore di K il 60% delle volte, mentre quelli dall'Italia solo il 20% delle volte. Con che percentuale, globalmente parlando, il guadagno è maggiore di K?

0.4, 0.1, 0.2, 0.3

1 Soluzioni

1. Stima della media: $\frac{287.8}{52} = 5.534$. Stima della deviazione:

$$S^2 = \frac{1}{51}\left(1782.5 - 52 \cdot 5.5346^2\right) = 3.7186, \quad S = 1.9284.$$

Errore assoluto al 95%, usando per semplicità l'approssimazione gaussiana:

$$\frac{1.9284 \cdot 1.96}{\sqrt{52}} = 0.524.$$

2.

$$P(Q \geq 8) = 1 - \Phi\left(\frac{8 - 5.5346}{1.9284}\right) = 1 - \Phi(1.2785) = 1 - 0.898 = 0.102.$$

3)

$$q = 5.5346 + 1.9284 \cdot 1.04 = 7.5401.$$

4) Intanto stimiamo λ con $\frac{1}{5.5346}$, poi calcoliamo

$$P(Q \geq 8) = e^{-\lambda \cdot 8} = \exp\left(-\frac{1}{5.5346} 8\right) = 0.23564.$$

5. Dobbiamo fare un test per la proporzione p di volte in cui si supera 8. Nell'ipotesi gaussiana vale $p = 0.102$, mentre $\widehat{p} = \frac{6}{20} = 0.3$. Possiamo eseguire un test gaussiano, grazie all'approssimazione fornita dal TLC. Vale $z = \frac{0.3-0.102}{\sqrt{0.102(1-0.102)}}\sqrt{20} = 2.9258$, da confortarsi con $q_{0.975} = 1.96$. Il test è significativo, rifiutiamo il modello gaussiano. Nel caso esponenziale vale $p = 0.235$, quindi $z = \frac{0.3-0.2356}{\sqrt{0.2356(1-0.2356)}}\sqrt{20} = 0.67866$ che non supera 1.96. Il test non è significativo, non rifiutiamo il modello esponenziale, che quindi è migliore dell'altro. Potevamo effettuare il test chi quadro, ottenendo lo stesso risultato.

6. Ad ogni settimana associamo una Bernoulli che vale 1 se quella settimana è critica. Il parametro p vale

$$p = P(Q \leq 2) = \Phi\left(\frac{2 - 5.5346}{1.9284}\right) = \Phi(-1.8329) = 1 - \Phi(1.8329) = 1 - 0.9664 = 0.0336.$$

Il numero N di settimane critiche è la somma di queste Bernoulli, che è una $B(10, p)$. Quindi

$$P(N \geq 1) = 1 - P(N = 0) = 1 - (1-p)^{10} = 1 - 0.9664^{10} = 0.289549.$$

7)

$$\begin{aligned} P(X_1 + ... + X_{100} > 7) &\approx 1 - \Phi\left(\frac{7 - 100 \cdot p}{\sqrt{100}\sqrt{p(1-p)}}\right) \\ &= 1 - \Phi\left(\frac{7 - 100 \cdot 0.0336}{\sqrt{100}\sqrt{0.0336(1-0.0336)}}\right) \\ &= 1 - \Phi(2.02) = 1 - 0.9783 = 0.0217. \end{aligned}$$

8) Vogliamo un test potente nell'evidenziare un calo, quindi usiamo il test unilaterale sinistro, confrontando cioè $\overline{x}$ con $\mu_0 - \frac{\sigma q_{0.95}}{\sqrt{n}}$ (nel senso che il test è siglificativo se $\overline{x} < \mu_0 - \frac{\sigma q_{0.95}}{\sqrt{n}}$) oppure $z = \frac{\overline{x}-\mu_0}{\sigma}\sqrt{n}$ con $q_{0.95}$. La potenza relativa al valore critico $\mu = 4$ è $1 - \beta(4)$, dove

$$\beta(4) = P_{\mu=4}\left(\overline{X} > \mu_0 - \frac{\sigma q_{0.95}}{\sqrt{n}}\right) = 1 - \Phi\left(\frac{\mu_0 - \frac{\sigma q_{0.95}}{\sqrt{n}} - 4}{\frac{\sigma}{\sqrt{n}}}\right)$$
$$= 1 - \Phi\left(\frac{5.5346 - 4}{1.9284}\sqrt{9} - 1.65\right) = 1 - \Phi(0.737\,37) = 1 - 0.7673 = 0.232\,7.$$

Quindi la potenza è 0.7673.

Per trovare la numerosità che fornisce potenza 0.9 dobbiamo risolvere l'equazione

$$\Phi\left(\frac{5.5346 - 4}{1.9284}\sqrt{n} - 1.65\right) = 0.9$$

(o meglio trovare il minimo n per cui vale il $\geq$). Quindi

$$\frac{5.5346 - 4}{1.9284}\sqrt{n} - 1.65 = 1.29$$
$$\frac{5.5346 - 4}{1.9284}\sqrt{n} = 1.29 + 1.65 = 2.\,94$$
$$\sqrt{n} = \frac{2.\,94 \cdot 1.9284}{5.5346 - 4} = 3.\,694\,4$$

da cui $n = 14$.

9) Dobbiamo calcolare $P(Z < z)$ dove

$$z = \frac{\frac{39.5}{9} - 5.\,534}{1.9284}\sqrt{9} = -1.\,781\,4.$$

Quindi il p-value è

$$\Phi(-1.\,781\,4) = 1 - \Phi(1.\,781\,4) = 1 - 0.9625 = 0.0375.$$

10)

$$P(>K) = P(>K|estero)\,P(estero) + P(>K|Italia)\,P(Italia)$$
$$= 0.6 \cdot \frac{1}{2} + 0.2 \cdot \frac{1}{2} = 0.4.$$