Esercizi riassuntivi di Inferenza Statistica

Esercizio 1
Su un campione casuale di 20 negozi di musica di una regione italiana la media settimanale di dischi di musica Jazz venduti è risultata pari a 160 con deviazione standard 46. In una diversa regione italiana su un campione casuale di 15 negozi la corrispondente media è risultata pari a 90 con deviazione standard 35. Sotto l'ipotesi di normalità del numero di dischi Jazz venduti settimanalmente da ciascun negozio,

1. si costruisca un intervallo di confidenza di livello 90% per la vera media settimanale di dischi Jazz venduti da un negozio nella prima regione.

2. si verifichi al livello 5% l'ipotesi che non vi sia differenza significativa nelle vendite medie settimanali di musica Jazz di un negozio tra le due regioni, contro l'ipotesi che il Jazz sia più popolare nella prima regione.

Esercizio 2
Da un'indagine pre-elettorale, su 150 persone intervistate 50 dichiarano di sostenere un certo partito politico. Si trovi un intervallo di confidenza di livello 0,9 per la percentuale reale di persone che sostengono tale partito. Se su 450 persone intervistate 150 risultano sostenere il partito, come si modifica l'intervallo di confidenza rispetto a quello costruito su 150 individui?

Esercizio 3
Una banca assume che i prelievi effettuati con il Bancomat si distribuiscono secondo una variabile casuale normale. Osservati 30 prelievi, con un prelievo medio pari a 125 euro e scarto quadratico medio parti a 32 euro,

1. costruire un intervallo di confidenza di livello 0,99 per lo scarto quadratico medio reale dei prelievi

2. Si verifichi, al livello 10% l'ipotesi che il livello medio dei prelievi sia pari a 150 euro contro l'ipotesi che sia inferiore a 150 euro.

Soluzioni degli esercizi riassuntivi di Inferenza Statistica

Esercizio 1

Si indichi con $\bar{x}$=160 il numero medio di dischi Jazz venduti in una settimana negli $n_1 = 20$ negozi della prima regione e con $s_1 = 46$ la deviazione standard. Per la seconda regione, $\bar{y}$=90 è il numero medio di dischi Jazz venduti negli $n_2 = 15$ negozi esaminati e $s_2 = 35$ la deviazione standard.

1. Sotto ipotesi di normalità del numero di dischi venduti settimanalmente da un negozio, un intervallo di confidenza di livello 90% per la media settimanale μ_1 di dischi Jazz venduti in un negozio della prima regione è

$$\left(\bar{x} \pm t_{(n_1-1);0,95}\sqrt{\frac{s_1^2 \cdot n_1}{(n_1-1)\cdot n_1}}\right) = \left(160 \pm 1,7291\sqrt{\frac{46^2 \cdot 20}{19 \cdot 20}}\right) = (142, 178)$$

2. Se indichiamo con μ_2 la media dei dischi Jazz venduti settimanalmente da un negozio nella seconda regione, il sistema di ipotesi che vogliamo sottoporre a verifica è

$$\begin{cases} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 > \mu_2 \end{cases}$$

Se assumiamo che le varianze nelle due popolazioni di negozi di dischi siano uguali, una statistica test adatta per questo problema di verifica di ipotesi è

$$\frac{\bar{x} - \bar{y}}{\sqrt{s_P^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{160 - 90}{\sqrt{1839,24\left(\frac{1}{20} + \frac{1}{15}\right)}} = 4,78$$

dove s_P^2 indica la stima combinata della varianza comune delle due popolazioni, ossia

$$s_P^2 = \frac{s_1^2 \cdot n_1 + s_2^2 \cdot n_2}{n_1 + n_2 - 2} = \frac{46^2 \cdot 20 + 35^2 \cdot 15}{20 + 15 - 2} = 1839,24$$

Se vogliamo verificare H_0 contro H_1 al livello α =5%, rifiutiamo H_0 in favore di H_1 se la statistica test (4,78) risulta maggiore di $t_{20+15-2;0.95} = 1,69$. La condizione è verificata quindi rigettiamo H_0 e accettiamo H_1 al livello di significatività specificato.

Esercizio 2

1. Viene richiesto un intervallo di confidenza di livello $1-\alpha$=0,9 per la percentuale di persone che sostengono un certo partito politico. La frequenza relativa dei sostenitori del partito sul campione considerato è $\hat{p} = 50/150 = 0,33$. Poiché $n\hat{p}$=150 $\cdot$ 0.33 e $n\hat{q}$=150 $\cdot$ 0.77 sono entrambi maggiori di 5, possiamo usare l'intervallo di confidenza approssimato per una percentuale (basato sull'approssimazione normale della distribuzione di $\hat{p}$),

$$\left(\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

   Nel nostro caso $1-\alpha = 0,9$ e quindi $1-\alpha/2$=0,95 e $z_{0.95} = 1,645$, da cui l'intervallo desiderato è

$$\left(0,33 \pm 1,645\sqrt{\frac{0,33 \cdot 0,77}{150}}\right) = (0,27; 0,397) = (27\%, 39,7\%)$$

2. Con i nuovi dati $\hat{p}$ diventa $\hat{p} = 150/450 = 0,33$. Pertanto la stima di p non cambia rispetto al caso precedente. Tuttavia il nuovo intervallo di confidenza al livello 0,9 è

$$\left(0,33 \pm 1,645\sqrt{\frac{0,33 \cdot 0,77}{450}}\right) = (0,297; 0,369) = (29,7\%, 36,9\%)$$

   che risulta più corto rispetto all'intervallo precedente. Questo è l'effetto della maggiore numerosità campionaria che, riducendo il termine $\hat{p}\hat{q}/n$, conduce ad un intervallo più piccolo, e quindi più preciso.

Esercizio 3

1. Per quanto affermato dal testo, i prelievi al Bancomat possono considerarsi distribuiti normalmente con media μ e varianza σ^2 ignote. L'ipotesi di normalità consente di costruire un intervallo di confidenza per la varianza σ^2 usando l'espressione

$$\left(\frac{s^2 \cdot n}{\chi^2_{n-1;1-\alpha/2}}, \frac{s^2 \cdot n}{\chi^2_{n-1;\alpha/2}}\right)$$

   Pertanto, per lo scarto quadratico medio σ possiamo usare

$$\left(\sqrt{\frac{s^2 \cdot n}{\chi^2_{n-1;1-\alpha/2}}}, \sqrt{\frac{s^2 \cdot n}{\chi^2_{n-1;\alpha/2}}}\right)$$

Nel caso in esame $s^2 = 32^2$, $1-\alpha = 0,99$ e quindi $\alpha = 0,01$, $\alpha/2$=0,005 e $1-\alpha/2$=0,995 e $n = 30$. Dalle tavole $\chi^2_{29,0,005} = 13,12$ e $\chi^2_{29,0,995} = 52,34$. Pertanto, un intervallo di livello 0,99 per σ è

$$\left(\sqrt{\frac{32^2 \cdot 30}{52,34}}, \sqrt{\frac{32^2 \cdot 30}{13,12}}\right) = (24,23; 48,39) \text{ euro}$$

2. Il sistema di ipotesi che si vuole sottoporre a verifica al livello $\alpha = 10\%$ è

$$\begin{cases} H_0 : & \mu = 150 \\ H_1 : & \mu < 150 \end{cases}$$

Essendo la varianza σ^2 ignota, una condizione di rifiuto di H_0 al livello 0,1 è

$$\frac{\bar{x} - 150}{\sqrt{\frac{s'^2}{30}}} = \frac{\bar{x} - 150}{\sqrt{\frac{s^2 \cdot 30}{29 \cdot 30}}} < -t_{29;0,9}$$

Essendo $\bar{x}$=125 e $s = 32$, la statistica test risulta pari a -4,20 che, confrontata con $-t_{29;0,9} = -1,3$, porta al rifiuto di H_0 e ad accettare H_1 al livello $\alpha = 0,1$.