CORSO DI STATISTICA (parte 2) - ESERCITAZIONE 5

Dott.ssa Antonella Costanzo

a.costanzo@unicas.it

Esercizio 1. Test ipotesi proporzione (grandi campioni)

E' necessario valutare l'opera svolta da un particolare ufficio di assistenza sociale e quindi si estraggono a caso dall'archivio 125 fascicoli relativi ad altrettanti casi. La percentuale di successi è risultata del 55% mentre la percentuale di successi richiesti ai singoli uffici è, di norma, del 60%. L'ufficio lavora secondo la norma? Si controlli scegliendo un livello di significatività del 5%.

Sol.

1.Definizione del sistema di ipotesi:

$$H_0 : p = 0.60$$
$$H_1 : p < 0.60$$

2. Livello di significatività

$\alpha = 0.05$

3. Costruzione della statistica test

Sfrutto il TLC, infatti n è sufficientemente grande per cui la statistica test è:

$$Z_{stat} = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \rightarrow N(0,1)$$

Il livello di significatività è $\alpha = 0.05$, il test è unidirezionale quindi il quantile della distribuzione Normale standard da individuare è:

$z_{0.95} = -1.645$

4. Definizione della regola di decisione (regione di rifiuto)

Se $Z_{stat} < -1.645$ oppure $|Z_{stat}| > 1.645$ allora rifiutiamo l'ipotesi nulla

5. A partire dal campione calcolo il valore della statistica sotto l'ipotesi nulla

$$Z_{stat} = \frac{0.55 - 0.60}{\sqrt{\frac{0.6(1-0.6)}{125}}} = -1.14$$

6. Decisione

Poichè $|Z_{stat}| = 1.111 < 1.645$ non si rifiuta l'ipotesi nulla. Ciò vuol dire che la differenza riscontrata fra la percentuale di successi osservata e quella attesa è spiegabile in termini di errore di campionamento.

6.bis il p-value

Definizione:
Il p-value o livello di significatività osservato è la probabilità di osservare valori della statistica test meno favorevoli ad H_0del valore effettivamente ottenuto. In altre parole, il p-value è il più piccolo livello di significatività tale da indurre al rifiuto dell'ipotesi nulla sulla base dei dati osservati. Esso rappresenta la probabilità che i dati non compatibili con l'ipotesi nulla siano stati osservati quando in realtà H_0 era vera. Di conseguenza un p-value molto piccolo è un forte indicatore del fatto che H_0 non è vera. Se il p-value<α, rifiuto quindi l'ipotesi nulla.

Nel nostro caso, essendo il test unidirezionale:

$$P\left(Z_{stat} < Z_{stat|H_o}\right) = P(Z_{stat} < -1.14) = 1 - P(Z_{stat} < 1.14) = 1 - 0.8729 = 0.1271$$

Siccome il p-value > α=0.05, allora non rifiuto l'ipotesi nulla.

Esercizio 2. Test ipotesi media, varianza ignota (grandi campioni)

Una ditta produttrice di alimenti per l'infanzia dichiara che il contenuto medio di calcio per 100 g di un dato alimento è di 200 mg. In un campione di 121 scatole da 100 g si è rilevato un contenuto medio di 191 mg con uno scarto quadratico medio corretto di 15 mg. Sapendo che il contenuto in calcio si distribuisce normalmente, si può ritenere che la ditta ha dichiarato un valore di calcio troppo elevato? Si scelga un livello di significatività dell'1%

Sol.

1.*Definizione del sistema di ipotesi:*

$$H_0 : \mu = 200$$
$$H_1 : \mu < 200$$

2. Livello di significatività

$\alpha = 0.01$

3. Costruzione della statistica test

Siccome la varianza della popolazione è ignota ma il campione è sufficientemente grande, allora:

$$Z_{stat} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \sim N(0,1)$$

Il livello di significatività è $\alpha = 0.01$, il test è a una coda, quindi il quantile della distribuzione normale da individuare è:

$$z_{0.99} = 2.325$$

4. *Definizione della regola di decisione (regione di rifiuto)*

Se $Z_{stat} < -2.325$ oppure, equivalentemente $|Z_{stat}| \geq 2.325$ allora rifiutiamo l'ipotesi nulla

5. A partire dal campione calcolo il valore della statistica sotto l'ipotesi nulla

$$Zstat = \frac{191-200}{15/\sqrt{121}} = -6.6$$

6. Decisione

Poichè $Z_{stat} = -6.6 < -2.325$ allora rifiutiamo l'ipotesi nulla.

Esercizio 3. Determinazione della numerosità campionaria

In un processo produttivo che funziona a regime, il peso dei pezzi prodotti si distribuisce normalmente con media 50 g e varianza 25. Se in un campione di pezzi, in cui la media è 52 g, l'ipotesi nulla di un divario casuale fra la media attesa del processo e la media del campione è stata rifiutata al 5% (concludendo così che il processo non funziona a regime), quante unità conteneva almeno il campione?

Sol.

L'ipotesi nulla da sottoporre a verifica è: $H_0: \mu = 50$ contro l'ipotesi alternativa $H_0: \mu \neq 50$.

La varianza della popolazione è $\sigma^2 = 25$

Le informazioni relative al campione sono $\bar{X} = 52$. La statistica test per controllare l'ipotesi nulla è:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{52 - 50}{5/\sqrt{n}}$$

Poiché l'ipotesi nulla è stata rifiutata per $\alpha = 0.05$, e al rifiuto di H_0 possono portare sia valori molto elevati, sia valori molto bassi, deve essere: $|Z| \geq z_{0.975} = 1.96$ cioè

$$\frac{2}{5/\sqrt{n}} \geq 1.96$$

$\sqrt{n} \geq 4.9$, $n \geq 24.01$. Poiché n deve essere intero, sarà $n \geq 24$