ESERCIZIO 1

Attraverso un'indagine effettuata su un campione di 83 pellicole cinematografiche, si verifichi, a livello 0.01, se vi sono differenze, in media, tra i budget dei film di produzione americana ed i budget dei film di produzione non americana degli ultimi 20 anni. Il campione di cui si dispone, costituito da 68 film americani e 15 non americani, fornisce le seguenti sintesi: $\bar{x} = 67.12, \bar{y} = 39.32, s_x = 43.17, s_y = 51.03$, essendo X il budget dei film americani ed Y il budget di quelli non americani. Si assumano distribuzioni normali con varianze uguali per i budget.

Soluzione

Si tratta di un test sulla differenza di medie per popolazioni normali indipendenti con varianze non note e uguali.
Si pongano $\mu_X = \mathbf{E}(X)$, $\mu_Y = \mathbf{E}(Y)$ e $\sigma^2 = \mathbf{V}\mathrm{ar}(X) = \mathbf{V}\mathrm{ar}(X)$.

Il problema è dunque

$$\begin{cases} H_0 : & \mu_X = \mu_Y \quad (\text{cioè } \mu_X - \mu_Y = 0) \\ H_1 : & \mu_X \neq \mu_Y \quad (\text{cioè } \mu_X - \mu_Y \neq 0) \end{cases}$$

La regione di rifiuto del test di livello 0.01 è:

$$R_{0.01} : |Z| = \left| \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_p^2}{n_X} + \frac{S_p^2}{n_Y}}} \right| \geq z_{\alpha/2} = 2.576;$$

è stato usato il quantile della distribuzione normale standard anzichè quello della distribuzione t di Student in quanto i gradi di libertà, 81, sono elevati per cui i quantili corrispondenti sono simili. Dal momento che

$$s_p^2 = \frac{43.17^2 \cdot 67 + 51.03^2 \cdot 14}{81} = 1991.62$$

e

$$z_{oss} = \left| \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_p^2}{n_X} + \frac{s_p^2}{n_Y}}} \right| = 2.183 < 2.576$$

non si rifiuta H_0 al livello $\alpha = 0.01$.

Si ottiene la stessa risposta calcolando il p-value=$2\mathbf{P}(Z \geq 2.183) = 0.029 > 0.01$, dove la probabilità è stata calcolata supponendo vera H_0 (ovvero, supponendo nulla la differenza tra le medie). Dunque, al livello 0.01, non si rileva evidenza di una differenza fra i budget USA e non USA.

ESERCIZIO 2

In una cassa per prelievo automatico di contante, riempita settimanalmente, si cerca il quantitativo ottimale di denaro da immettere per evitare, da una parte, che la cassa resti vuota e, dall'altra parte, che ci sia una quantità eccessiva di denaro. I dati in possesso della banca indicano che, in passato, mediamente venivano prelevate 400 (migliaia) di euro per settimana. Si stabilisca, a livello 0.05, se é cambiata la quantità media di contante prelevato per settimana, attraverso un test basato su un campione di 8 settimane in ciascuna delle quali viene rilevata la quantità di contante prelevata; la media osservata sul campione è pari a 380, mentre la varianza campionaria risulta pari a 342.8571. Si scrivano le ipotesi nulla ed alternativa e si precisi la regione di rifiuto del test. Cambia il risultato del test se il livello adottato è 0.01? E se è 0.1? Si assuma una distribuzione normale per il contante prelevato.

Soluzione

Si tratta di un test sulla media di una popolazione normale con varianza non nota. Le ipotesi sono:

$$\begin{cases} H_0 : & \mu = 400 \\ H_1 : & \mu \neq 400 \end{cases}$$

La regione di rifiuto è:

$$R_{0.05} : |T| = \left| \frac{\bar{X} - 400}{\sqrt{S^2/n}} \right| \geq t_{\alpha/2}^{n-1} = t_{0.025}^{7} = 2.365.$$

Dal momento che $t_{oss} = \left| \frac{\bar{x}-400}{\sqrt{s^2/n}} \right| = \left| \frac{380-400}{\sqrt{342.8571/8}} \right| = 3.055 > 2.365$, si rifiuta H_0 al livello $\alpha = 0.05$. Si ottiene la stessa risposta osservando che il p-value è compreso tra 0.01 e 0.02, quindi <0.05 (si guardino le tavole della distribuzione t di Student, in corrispondenza a 7 gradi di libertà).
Se invece $\alpha = 0.01$ non si rifiuta (p-value> α) mentre se $\alpha = 0.1$ si rifiuta.

ESERCIZIO 3

Un'indagine sui possessori di carta di credito, effettuata su un campione di ampiezza 20, ha fornito una spesa mensile media pari a 820 euro, con uno scarto quadratico medio pari a 156 euro. Si verifichi, a livello 0.05, se la spesa media mensile nell'ambito di tutti i possessori di carta di credito è aumentata, rispetto al passato in cui era pari a 800 euro, assumendo un modello gaussiano per la distribuzione del carattere.

Soluzione

Si ponga X = spesa mensile con carta di credito; si assume $X \sim \mathcal{N}(\mu, \sigma^2)$.
Le ipotesi sono:

$$\begin{cases} H_0 : & \mu = 800 \quad (\text{o } \mu \leq 800) \\ H_1 : & \mu > 800 \end{cases}$$

e la regione di rifiuto del test di livello 0.05 è

$$R_{0.05} : T = \frac{\bar{X} - 800}{\sqrt{S^2/n}} \geq t_{\alpha}^{n-1} = t_{0.05}^{19} = 1.729.$$

Dal momento che $t_{oss} = \frac{820-800}{156/\sqrt{20}} = 0.573 < 1.729$, non si rifiuta H_0 al livello $\alpha = 0.05$. Non si rileva dunque evidenza empirica che la spesa sia aumentata.

ESERCIZIO 4

Un processo produttivo è tarato per inscatolare un certo prodotto alimentare in confezioni da 1000 grammi; il processo è ritenuto sotto controllo se il peso medio delle confezioni prodotte è quindi pari a 1000 grammi. Per controllare se il processo è fuori controllo si estrae un campione di 80 confezioni, il cui peso medio è pari a 997 grammi e la deviazione standard pari a 16 grammi.
a) Si scrivano le ipotesi nulla ed alternativa.
b) Si calcoli il p-value del test; se ne descriva il significato.
c) Si utilizzi il p-value per stabilire a livello 0.05 se il processo è fuori controllo.
d) Si dica, giustificando la risposta, se sono necessarie ipotesi sul modello per eseguire il test.

Soluzione

a) Le ipotesi sono:

$$\begin{cases} H_0: & \mu = 1000 \\ H_1: & \mu \neq 1000 \end{cases}$$

b) Il valore osservato per la statistica test è $z_{\mathrm{OSS}} = \frac{997-1000}{16/\sqrt{80}} = -1.677$. Il p-value è:

$$\text{p-value} = 2\mathbf{P}(Z \geq |z_{\mathrm{OSS}}|) = 2\mathbf{P}(Z \geq 1.677) = 0.093;$$

si tratta della probabilità, valutata supponendo vera H_0 (supponendo cioè $\mu = 1000$) che la statistica test assuma valori "uguali o più estremi" di quello effettivamente osservato nel campione.

c) Poichè il p-value è maggiore di 0.05, non si rifiuta l'ipotesi nulla a questo livello; dunque non si rileva evidenza empirica che il processo sia fuori controllo.

d) Non è necessaria alcuna ipotesi sul modello in quanto il campione è sufficientemente ampio (n=80).