A

TEST DI SIGNIFICATIVITÀ PER UNA MEDIA

Cinque parti:

1. assunzioni
2. ipotesi
3. test
4. P-valore
5. conclusioni

A

TEST DI SIGNIFICATIVITÀ PER UNA MEDIA

1. assunzioni

a) variabile quantitativa

b) casualità → dati ottenuti utilizzando la randomizzazione, come nel caso di un campione casuale semplice

c) la variabile ha una distribuzione normale nella popolazione

A

TEST DI SIGNIFICATIVITÀ PER UNA MEDIA

2. ipotesi

H_o: $\mu = \mu_o$

H_a: $\mu \neq \mu_o$ (H_a: $\mu < \mu_o$ oppure H_a: $\mu > \mu_o$)

A

TEST DI SIGNIFICATIVITÀ PER UNA MEDIA

3. test → t-score

descrive quanto dista (in termini di numero di errori standard) la stima puntuale dal valore ipotizzato in H_o

$t = (\bar{y} - \mu_o)/se$

dove l'errore standard stimato $se = s/\sqrt{n}$

A

TEST DI SIGNIFICATIVITÀ PER UNA MEDIA

4. P-valore

il P-valore è il peso dell'evidenza contro H_o

→ il P-valore è la probabilità, se vera H_o, che la statistica test sia pari al valore osservato o a uno compreso nell'insieme dei valori più estremi che forniscono evidenze ancora più forti contro H_o

P = probabilità a due code per H_a: $\mu \neq \mu_o$

P = probabilità a destra del t-valore osservato per H_a: $\mu > \mu_o$

P = probabilità a sinistra del t-valore osservato per H_a: $\mu < \mu_o$

A

TEST DI SIGNIFICATIVITÀ PER UNA MEDIA

5. conclusioni

interpretazione del P-valore → piccoli valori di P forniscono forti evidenze contro H_o e a favore di H_a

Si rifiuta H_o se $P \leq \alpha$ (livello di significatività)

B

TEST DI SIGNIFICATIVITÀ PER UNA PROPORZIONE

Cinque parti:

1. assunzioni
2. ipotesi
3. test
4. P-valore
5. conclusioni

B

TEST DI SIGNIFICATIVITÀ PER UNA PROPORZIONE

1. assunzioni

a) casualità → dati ottenuti utilizzando la randomizzazione, come nel caso di un campione casuale semplice

b) dimensione campionaria sufficientemente ampia in modo che la distribuzione della proporzione campionaria sia approssimativamente normale

B

TEST DI SIGNIFICATIVITÀ PER UNA PROPORZIONE

2. ipotesi

H_o: $\Pi = \Pi_o$

H_a: $\Pi \neq \Pi_o$ (H_a: $\Pi < \Pi_o$ oppure H_a: $\Pi > \Pi_o$)

B

TEST DI SIGNIFICATIVITÀ PER UNA PROPORZIONE

3. test

$z = (\hat{\Pi} - \Pi_o)/se_o$

dove l'errore standard sotto l'ipotesi nulla, $se_o = \sqrt{\Pi_o\,(1 - \Pi_o)/n}$

B

TEST DI SIGNIFICATIVITÀ PER UNA PROPORZIONE

4. P-valore

il P-valore è il peso dell'evidenza contro H_o

→ il P-valore è la probabilità, se vera H_o, che la statistica test sia pari al valore osservato o a uno compreso nell'insieme dei valori più estremi che forniscono evidenze ancora più forti contro H_o

P = probabilità a due code per H_a: $\Pi \neq \Pi_o$

P = probabilità a destra del valore osservato z per H_a: $\Pi > \Pi_o$

P = probabilità a sinistra del valore osservato z per H_a: $\Pi < \Pi_o$

B

TEST DI SIGNIFICATIVITÀ PER UNA PROPORZIONE

5. conclusioni

interpretazione del P-valore → piccoli valori di P forniscono forti evidenze contro H_o e a favore di H_a

Si rifiuta H_o se $P \leq \alpha$ (livello di significatività)

ESERCITAZIONE 6

PROBLEMA 1

Per una verifica di H_o: $\mu = 0$ contro H_a: $\mu \neq 0$ con $n = 1000$, la statistica test t è pari a 1.04.

a) trova il p-valore e interpretalo
b) supponi che t = -2.50 piuttosto che 1.04. Trova il p-valore. Questo valore fornisce un'evidenza contro l'ipotesi nulla più forte o più debole? Fornisci una spiegazione.
c) quando t = 1.04, trova il p-valore per: (i) H_a: $\mu > 0$, (ii) H_a: $\mu < 0$

Soluzione

a) gdl=n-1=999 quindi è possibile usare la normale standardizzata per approssimare la distribuzione t; p-valore=2 x 0.1492=0.30 → con α=0.05 non rifiuto H_o (è plausibile che μ sia uguale a 0)
b) p-valore=2 x 0.0062=0.012 → con α=0.05 rifiuto H_o; questo p-valore fornisce un'evidenza più forte contro H_o; in generale, p-valori più piccoli danno evidenze più forti
c) (i) p-valore=0.1492; (ii) p-valore= 1-0.1492=0.851

PROBLEMA 2

Trova e interpreta il p-valore per la verifica di H_o: $\mu = 100$ contro H_a: $\mu \neq 100$, se un campione ha:

a) $n = 400$, $\bar{y} = 103$, $s = 40$
b) $n = 1600$, $\bar{y} = 103$, $s = 40$
c) commenta l'effetto di n sui risultati di un test di significatività

Soluzione

i due casi hanno la stessa media e deviazione standard, ma la grandezza del campione è diversa

a) con n=400
   se=2
   t=(103-100)/2=1.5
   p-valore=2 x 0.0668=0.134
   se α=0.05 non rifiuto H_o
b) con n=1600
   se=1
   t=(103-100)/1=3
   p-valore=2 x 0.00135=0.003
   se α=0.05 rifiuto H_o
c) all'aumentare di n anche una piccola differenza (100 vs 103) può produrre un p-valore piccolo

PROBLEMA 3

Secondo un accordo sindacale, il reddito medio dei lavoratori senior nell'azienda XY deve essere pari a \$ 500 a settimana. Si decide di analizzare se il reddito medio μ delle lavoratrici donne è conforme all'accordo. Per un campione casuale di nove donne occupate sono stati ottenuti i valori, $\bar{y}$ = \$ 410, s = \$ 90.

a) verifica se il reddito medio delle donne lavoratrici è differente da \$ 500 alla settimana. Esplicita le assunzioni, le ipotesi, il test e il p-valore. Interpreta il risultato
b) riporta e interpreta il p-valore per H_a: $\mu < 500$
c) riporta e interpreta il p-valore per H_a: $\mu > 500$

Soluzione

a) assunzioni

variabile quantitativa; casualità; normalità

ipotesi

H_o: $\mu = 500$

H_a: $\mu \neq 500$

test

se=30

t=(400-500)/30=-3

p-valore

dgl=8

(0.005 x 2) < p-valore < (0.010 x 2) = 0.01 < p-valore < 0.02

conclusioni

con α=0.05 rifiuto H_o

b) H_a: $\mu < 500$ → p-valore=0.01 (evidenze forti che la media sia inferiore a 500)

c) H_a: $\mu > 500$ → p-valore=0.99

PROBLEMA 4

Per una verifica di H_o: $\pi = 0.50$, la proporzione campionaria è 0.35 secondo un campione di dimensione 100.

a) mostra che la statistica test è z= -3.0
b) trova e interpreta il p-valore per H_a: $\pi < 0.50$
c) per un livello di significatività pari a 0.05, quale decisione puoi prendere?
d) se la decisione presa in c) fosse errata, di quale tipo di errore si tratterrebbe?

Soluzione

a) $se_o = \sqrt{(0.50 \times 0.50)/100} = 0.05$

$z = (0.35 - 0.50)/0.05 = -3$

b) p-valore=0.00135; forte evidenza in favore di H_a

c) con α=0.05 rifiuto H_o

d) errore di I tipo

PROBLEMA 5

Uno studio considera se il punteggio medio μ ad un esame di ammissione a psicologia nel 2016 sia in qualche modo differente dalla media 500 registrata nel 2006. Verifica H_o: $\mu = 500$ contro H_a: $\mu \neq 500$, se per un campione casuale a livello nazionale di 10000 studenti che hanno sostenuto l'esame nel 2016, $\bar{y} = 497$, s = 100. Commenta il risultato.

Soluzione

se = 1

t = (497 - 500)/1 = -3

p-valore = 2 x 0.00135 = 0.003

con α=0.05 rifiuto H_o

il risultato è statisticamente altamente significativo ma poco significativo dal punto di vista pratico (la differenza è minima, solo 3 punti, ma essendo n molto grande questo è sufficiente per far emergere una differenza)

PROBLEMA 6

In una elezione a sindaco vi sono due candidati.

a) dato un campione casuale di 400 votanti, 230 hanno votato per un certo candidato. Sei disposto a prevedere il vincitore? perchè?

b) dato un campione casuale di 40 votanti, 23 hanno votato per un certo candidato. Sei disposto a prevedere il vincitore? perchè?

Soluzione

la proporzione campionaria è la stessa (0.575) nei due casi; sono le dimensioni campionarie che differiscono

a) $se_o = 0.025$
$z = (0.575 - 0.50)/0.025 = 3$
p-valore = 2 x 0.00135 = 0.003
con α=0.05 rifiuto H_o e quindi posso prevedere il vincitore

b) $se_o = 0.079$
$z = (0.575 - 0.50)/0.079 = 0.95$
p-valore = 2 x 0.1711 = 0.342
con α=0.05 non rifiuto H_o e quindi non posso prevedere il vincitore