ESERCIZIO 5.19

Riporta il valore del t-score necessario per:

a) Calcolare un intervallo di confidenza al 95% con 5 osservazioni.

b) Calcolare un intervallo di confidenza al 95% con 15 osservazioni.

c) Calcolare un intervallo di confidenza al 95% con gdl = 25.

d) Prendere una decisione in una verifica di ipotesi unilaterale con $\alpha = 0.05$. N = 18.

e) Prendere una decisione in una verifica di ipotesi bilaterale con $\alpha = 0.05$. Gdl = 30.

RISOLUZIONE:

Per tutti i punti uso la tavola T a pag. 287 del libro di Luccio. I gradi di libertà (gdl) corrispondono a (n – 1).

a) Devo trovare quel t-score che per 4 gdl lascia a destra il 0.025 (2.5%) dell'area. Infatti, mi interessa quel valore che lascia all'esterno dell'intervallo il 5% dell'area.

Incrociando quindi 4 gdl con 0.025 trovo che **t = 2.78**.

b) Ragionamento uguale al punto precedente, ma in questo caso gdl = 14.

Incrociando quindi 14 gdl con 0.025 trovo che **t = 2.14**.

c) Ragionamento uguale al punto **a)**, ma in questo caso gdl = 25.

Incrociando quindi 25 gdl con 0.025 trovo che **t = 2.06**.

d) Per prendere una decisione in un'ipotesi unilaterale ho bisogno di quel t-score che, per 17 gdl, lascia a destra il 0.05 (5%) dell'area (questo t-score "critico" lo confronterei poi con il t-score associato al valore campionario).

Incrociando quindi 17 gdl con 0.05 trovo che **t = 1.74**.

e) Per prendere una decisione in un'ipotesi bilaterale ho bisogno di quel t-score che, per 30 gdl, lascia sia a destra che a sinistra il 0.025 (2.5%) dell'area. Infatti, mi interessa quel valore che lascia all'esterno, nel complesso, un'area del 5%. (questo t-score "critico" lo confronterei poi con il t-score associato al valore campionario).

Incrociando quindi 30 gdl con 0.025 trovo che **t = 2.04**.

ESERCIZIO 5.20

Trova e interpreta l'intervallo di confidenza al 95% per μ, se $\overline{x}$ = 70 e s = 10, basato su una dimensione campionaria pari a:

a) 5

b) 20

RISOLUZIONE:

Per trovare l'intervallo di confidenza di una media μ, la formula è **$\overline{x}$ ± t*se**. Dove: **t** è il **t-score** associato a dei certi gradi di libertà (n – 1) e a un certo livello di confidenza e dove: **se** = $\frac{s}{\sqrt{n}}$ **.**

a)

Trovo il **t-score** che per 4 gdl lascia alla destra della distribuzione il 0.025 (2.5%) dell'area (così nel complesso, sia a destra che a sinistra, lascia il 5% dell'area).

Nella tavola T a pag. 287 del libro di Luccio, incrociando 4 gdl con 0.025 trovo che **t = 2.78**.

Calcolo l'**se: se** = $\frac{10}{\sqrt{5}}$ **= 4.47**.

Ho quindi i dati per calcolare l'intervallo:

I.C. = 70 ± 2.78*4.47 = 70 ± 12.43 = (57.57, 82.43).

L'interpretazione dell'intervallo è: Ho una fiducia del 95% nel ritenere che l'intervallo da me trovato include il valore medio μ della popolazione. Ovvero, immaginando di ripetere il processo di campionamento innumerevoli volte, ho il 95% di fiducia nel ritenere che l'intervallo da me trovato sia uno di quei intervalli che contiene il valore μ della popolazione.

b)

Quello che cambia dal punto **a)** è la dimensione campionaria, quindi variano i gdl e l'**se**.

Trovo il **t-score** che per 19 gdl lascia alla destra della distribuzione il 0.025 (2.5% dell'area) (così nel complesso, sia a destra che a sinistra, lascia il 5% dell'area).

Nella tavola T a pag. 287 del libro di Luccio, incrociando 19 gdl con 0.025 trovo che **t = 2.09**.

Calcolo l'**se: se** = $\frac{10}{\sqrt{20}}$ **= 2.236**.

Ho quindi i dati per calcolare l'intervallo:

I.C. = 70 ± 2.06*2.236 = 70 ± 4.67 = (65.33, 74.67).

L'interpretazione dell'intervallo è: Ho una fiducia del 95% nel ritenere che l'intervallo da me trovato include il valore medio μ della popolazione. Ovvero, immaginando di ripetere il processo di campionamento innumerevoli volte, ho il 95% di fiducia nel ritenere che l'intervallo da me trovato sia uno di quei intervalli che contiene il valore μ della popolazione.

ESERCIZIO C1

Riporta i valori del χ^2 necessari per:

a) Trovare il 10% dell'area sottesa dalla coda destra della distribuzione. Gdl = 7.

b) Trovare il 2.5% dell'area sottesa dalla coda sinistra della distribuzione. Gdl = 9.

c) Trovare il 5% dell'area equamente sottesa da entrambe le code. Gdl = 10.

d) Trovare il 10% dell'area equamente sottesa da entrambe le code. Gdl = 20.

e) Prendere una decisione in una verifica di ipotesi unilaterale H1: $\sigma > x$ con $\alpha = 0.01$. Gdl = 30.

RISOLUZIONE:

Per tutti i punti utilizzo la tavola χ^2 a pag. 294-295 del libro di Luccio. I gradi di libertà (gdl) corrispondono a (n – 1).

a) Devo trovare quel χ^2-score che per 7 gdl lascia a destra il 0.100 (10%) dell'area.

Incrociando quindi 7 gdl con 0.100 trovo che $\boldsymbol{\chi^2}$**= 12.0170**.

b) Devo trovare quel χ^2-score che per 9 gdl lascia a sinistra il 0.025 (2.5%) dell'area. Visto che la tavola del libro riporta le aree lasciate nella coda destra cerco quel valore che lascia il (1 – 0.025) dell'area, ovvero quel valore che lascia il 0.975 (97.5%) dell'area.

Incrociando quindi 9 gdl con 0.975 trovo che $\boldsymbol{\chi^2}$ **= 2.70039**.

c) Devo trovare quel χ^2-score che per 10 gdl lascia sia a sinistra che a destra il 0.025 (2.5%) dell'area. Così, nel complesso viene lasciata un'area del 5%. Quindi, per trovare i due valori del χ^2 corrispondenti, cerco prima il χ^2 che lascia a destra il 0.975 (97.5%) dell'area (ragionamento analogo al punto **b)**) e poi cerco il χ^2 che lascia a destra il 0.025 (2.5%) dell'area.

Incrociando quindi 10 gdl con 0.975 trovo che $\boldsymbol{\chi^2}$ **= 3.24697.**

Incrociando quindi 10 gdl con 0.025 trovo che $\boldsymbol{\chi^2}$ **= 20.4832.**

d) Ragionamenti analoghi al punto **c)**. In questo caso però i gdl sono 20 mentre l'area lasciata sia a sinistra che a destra è il 0.05 (5%). Così nel complesso viene lasciata un'area del 10%. Quindi, per trovare i due valori del χ^2 corrispondenti, cerco prima il χ^2 che lascia a destra il 0.950 (95%) dell'area (ragionamento analogo al punto **b)**) e poi cerco il χ^2 che lascia a destra il 0.050 (5%) dell'area.

Incrociando quindi 20 gdl con 0.950 trovo che $\boldsymbol{\chi^2}$ **= 10.8508.**

Incrociando quindi 20 gdl con 0.050 trovo che $\boldsymbol{\chi^2}$ **= 31.4104.**

e) Per prendere una decisione in un'ipotesi unilaterale destra, con $\alpha = 0.01$, ho bisogno di quel χ^2-score che, per 30 gdl, lascia a destra il 0.01 (1%) dell'area. In questo modo ho il χ^2-score "critico" che confronterei poi con il χ^2-score associato al valore campionario.

Incrociando quindi 30 gdl con 0.01 trovo che $\boldsymbol{\chi^2}$ **= 50.8922**.

ESERCIZIO DI ESEMPIO

Si supponga di aver scelto a caso 16 scolaresche dalla popolazione italiana, omogenee come numero, sesso ed età dei componenti. Per queste scolaresche si rileva il tempo (ore) dedicato ad attività ricreative. Le statistiche riassuntive del campione danno una media $\overline{x} = 3.338$, $s_1^2 = 2.897$ e $s_1 = 1.702$.

a) Calcola un intervallo di confidenza al 95% per σ^2.

b) Verifica se vi sono evidenze a favore del fatto che la varianza del tempo dedicato ad attività ricreative delle scolaresche sia maggiore di 1, prendi una decisione con $\alpha = 0.05$.

c) Verifica se la varianza ottenuta da queste scolaresche sia significativamente maggiore della varianza ottenuta da altre scolaresche ($n = 13$, $s_2^2 = 2.134$) italiane. $\alpha = 0.05$.

RISOLUZIONE:

a)

Per calcolare l'intervallo di confidenza al 95% per la varianza σ^2 la formula è: $\frac{(n-1)*s^2}{\chi^2_{1-(\alpha/2)}} \geq \sigma^2 \geq \frac{(n-1)*s^2}{\chi^2_{(\alpha/2)}}$ Dove **n** è il numero di osservazioni, s^2 è la varianza campionaria,

α è 0.05 (corrisponde a 1 – livello di fiducia dove il livello di fiducia è 95% in questo caso),

$\chi^2_{1-(\alpha/2)}$ e $\chi^2_{(\alpha/2)}$ sono le statistiche del χ^2 che per (n -1) gdl sono associate all'area espressa dal pedice (in questo caso 1 – 0.025 = **0.975** e **0.025**).

Di questi mi mancano le statistiche del χ^2. Le cerco quindi nella tavola χ^2 a pag. 294-295 del libro di Luccio.

Incrociando 15 gdl con 0.975 trovo che $\chi^2_{0.975} = 6.26214$.

Incrociando 15 gdl con 0.025 trovo che $\chi^2_{0.025} = 27.4884$.

A questo punto sostituisco nella formula e trovo l'intervallo:

$\frac{(16-1)*2.897}{6.26214} \geq \sigma^2 \geq \frac{(16-1)*2.897}{27.4884} = $ **$6.94 \geq \sigma^2 \geq 1.58$.** Quindi l'intervallo è **(6.94, 1.58).**

b)

Per verificare la congettura faccio una verifica di ipotesi. Visto che mi chiede di verificare se σ^2 è maggiore di 1, le ipotesi che pongo sono queste:

H_0: $\sigma^2 \leq 1$ e H_1: $\sigma^2 > 1$.

Come prossimo passo devo calcolare la statistica test. Per la varianza utilizzo la statistica test del χ^2: $\chi^2 = \frac{(n-1)*s^2}{\sigma_0^2}$. Dove:

n è il numero di osservazioni, s^2 è la varianza campionaria, σ^2_0 è il valore della varianza assunto sotto l'ipotesi nulla. Ho tutti i valori quindi procedo con il calcolo: $\chi^2 = \frac{(16-1)*2.897}{1} = 43.455$.

Ho trovato la statistica test, adesso dovrei trovare il p-valore ad essa associato e verificare se è inferiore ad α (0.05). Però la tavola χ^2 a pag. 294-295 del libro di Luccio non riporta i p-valori ma riporta delle statistiche test χ^2 che per determinati gdl lasciano alla loro destra una certa percentuale d'area della distribuzione.

Quindi posso trovare qual è quel χ^2 "critico" che lascia alla sua destra il 0.05 per verificare se il mio χ^2 è maggiore o inferiore a esso.

Se il mio χ^2 è superiore a quello "critico" allora il suo p-valore è inferiore a 0.05. E rifiuto l'ipotesi nulla.

Se il mio χ^2 è inferiore a quello "critico" allora il suo p-valore è superiore a 0.05. E non rifiuto l'ipotesi nulla.

Quindi, nella tavola, incrociando 15 gdl con 0.05 trovo che: $\boldsymbol{\chi^2_{0.050} = 27.9958}$**.**

Vedo che $\boldsymbol{\chi^2 > \chi^2_{0.050}}$ **(43.455 > 27.9958).**

Quindi concludo che: sotto l'assunzione della veridicità dell'ipotesi nulla sarebbe improbabile trovare una varianza di 2.897. Quindi reputo questa ipotesi implausibile (rifiuto quindi l'ipotesi nulla) e considero più plausibile che il valore di σ^2 sia superiore a 1.

c)

Questo punto mi sta chiedendo di fare un confronto fra due varianze. Per verificare questa congettura faccio una verifica di ipotesi. Visto che mi chiede di verificare se la varianza delle 16 scolaresche di partenze è maggiore alla varianza delle 13 nuove scolaresche le ipotesi che pongo sono:

H_0: $\sigma^2_1 \leq \sigma^2_2$ e H_1: $\sigma^2_1 > \sigma^2_2$.

Come passo successivo devo calcolare la statistica test. Quando si fa un confronto fra varianze il test appropriato è il test che utilizza la distribuzione F. Quindi calcolo il valore della statistica test $\boldsymbol{F = \frac{s_1^2}{s_2^2}}$.

Ho a disposizione le varianze campionarie delle due scolaresche quindi ricavo la statistica test.

$\boldsymbol{F = \frac{2.897}{2.134} = 1.358}$.

Ho trovato la statistica test, adesso dovrei trovare il p-valore ad essa associato e verificare se è inferiore ad α (0.05). Per procedere utilizzo la Tavola F a pag. 288 e 289 del libro di Luccio.

La tavola F non riporta i p-valori ma riporta delle statistiche test F che per determinati gdl lasciano alla loro destra il 0.05 (5%) dell'area della distribuzione (scelgo questa con 0.05 appunto perché il punto **c)** pone α = 0.05). Quindi posso trovare qual è quel F "critico" che lascia alla sua destra il 0.05 per verificare se il mio F è maggiore o inferiore a esso.

Se il mio F è superiore a quello "critico" allora il suo p-valore è inferiore a 0.05. E rifiuto l'ipotesi nulla.

Se il mio F è inferiore a quello "critico" allora il suo p-valore è superiore a 0.05. E non rifiuto l'ipotesi nulla.

La distribuzione F non ha un solo gdl, ma ne ha due. Uno dei gdl è rappresentato sulle righe e l'altro dei gdl è rappresentato sulle colonne.

I gdl rappresentati sulle colonne sono i gdl della varianza (n – 1) del numeratore della frazione $\boldsymbol{F = \frac{s_1^2}{s_2^2}}$.

I gdl rappresentati sulle righe sono i gdl della varianza (n – 1) del denominatore della frazione $\boldsymbol{F = \frac{s_1^2}{s_2^2}}$.

Quindi in questo caso i gdl del numeratore sono 15, mentre quelli del denominatore 12. Quindi, per α = 0.05, all'incrocio fra questi due gdl trovo che $\boldsymbol{F_{15,12} = 2.6169}$**.**

Vedo che $\boldsymbol{F < F_{15,12}}$ **(1.358 < 2.6169).**

Quindi concludo che: sotto l'assunzione della veridicità dell'ipotesi nulla sarebbe plausibile trovare una varianza di 2.897. Quindi reputo questa ipotesi plausibile (non rifiuto quindi l'ipotesi nulla) e ritengo plausibile che le due varianze siano uguali.