ESERCIZIO 6.16

In un sondaggio del Pew Research Center (14 maggio 2003) condotto su 1201 adulti è stato chiesto "Ritieni che i programmi attuati per incrementare il numero di studenti neri e delle minoranze nei campus dei college siano una buona o una cattiva cosa?". Il 60% ha detto *buona* il 30% ha detto *cattiva* e il 10% non ha saputo rispondere. Sia π la proporzione della popolazione che ritiene siano una buona cosa. Fai una verifica di ipotesi con H_0: $\pi = 0.50$ contro H_1: $\pi > 0.50$. Prendi una decisione per $\alpha = 0.05$ e fornisci un'interpretazione del P-valore.

RISOLUZIONE:

Il problema mi chiede di fare una verifica di ipotesi per la proporzione di adulti che ritiene che i programmi attuati siano una buona cosa. Quindi la proporzione campionaria che considero è quella corrispondente al 60% (= 0.60). Quindi $\hat{\pi} = 0.60$.

Le ipotesi sono: H_0: $\pi \leq 0.50$ vs H_1: $\pi > 0.50$

Come primo passo calcolo la statistica test z associata alla proporzione campionaria.

La formula è: $\boldsymbol{z = \frac{\hat{\pi}-\pi_0}{se_0}}$; dove $\boldsymbol{\pi_0}$ è la proporzione assunta sotto l'ipotesi nulla e dove $\boldsymbol{se_0}$ è l'errore standard della distribuzione assunta sotto l'ipotesi nulla. $\boldsymbol{se_0 = \sqrt{\frac{\pi_0 * (1-\pi_0)}{n}}}$.

Quindi: $\boldsymbol{se_0 = \sqrt{\frac{0.50 * (1-0.50)}{1201}}}$ **= 0.014.**

Quindi: $\boldsymbol{z = \frac{0.60-0.50}{0.014}}$ **= 7.14.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di 7.14. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 3.99 la probabilità è inferiore a 0.0000, quindi per 7.14 è ancora inferiore.

Quindi: **p-valore = 0.0000…**

L'interpretazione del **p-valore** è: Sotto l'assunzione della veridicità dell'ipotesi nulla, la probabilità di ottenere per caso una proporzione pari a 0.60 o superiore è uguale a 0.0000….

Per prendere una decisione con un livello di significatività $\alpha = 0.05$ confronto il **p-valore** che ho ottenuto e vedo se è inferiore ad α.

0.0000… < 0.05.

Essendo il p-valore inferiore rifiuto l'ipotesi nulla.

ESERCIZIO 6.7

Secondo un accordo sindacale, il reddito medio per tutti i lavoratori di profilo senior della linea di assemblaggio di una grande azienda deve essere pari a 500$ alla settimana. Una rappresentante di un gruppo di donne decide di analizzare se il reddito medio μ delle lavoratrici è conforme a questo accordo. Per un campione casuale di cento donne occupate sono stati ottenuti i valori, $\bar{x}$ = 475$ e s = 60$.

a) Verifica se il reddito medio delle donne occupate è inferiore a 500$ alla settimana. Esplicita le ipotesi, calcola il test e il P-valore. Prendi una decisione per α = 0.05 e interpreta il risultato.

b) Riporta e interpreta il P-valore per H_1: μ > 500.

RISOLUZIONE:

a)

Il problema mi chiede di verificare se il reddito delle donne occupate è **inferiore** a 500$ alla settimana.

Le ipotesi quindi sono: H_0: μ ≥ 500 vs H_1: μ < 500

Come secondo passo calcolo la statistica test z associata alla media campionaria.

La formula è: $\boldsymbol{z} = \frac{\bar{x} - \mu_0}{se_0}$; dove $\boldsymbol{\mu_0}$ è la media assunta sotto l'ipotesi nulla e dove $\mathbf{se_0}$ è l'errore standard della distribuzione assunta sotto l'ipotesi nulla. $\boldsymbol{se_0} = \frac{s}{\sqrt{n}}$.

Quindi: $\boldsymbol{se_0} = \frac{60}{\sqrt{100}}$ **= 6.**

Quindi: $\boldsymbol{z} = \frac{475-500}{6}$ **= - 4.167.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di -4.167. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 3.99 la probabilità è inferiore a 0.0000, quindi per -4.167 è ancora inferiore.

Quindi: **p-valore = 0.0000…**

L'interpretazione del **p-valore** è: Sotto l'assunzione della veridicità dell'ipotesi nulla, la probabilità di ottenere per caso una media pari a 475$ o inferiore è uguale a 0.0000….

Per prendere una decisione con un livello di significatività α = 0.05 confronto il **p-valore** che ho ottenuto e vedo se è inferiore ad α.

0.0000… < 0.05. Essendo il p-valore inferiore rifiuto l'ipotesi nulla.

Quindi il risultato ottenuto mi dice che ho forti evidenze per ritenere che il reddito medio delle lavoratrici non sia pari a 500$ ma che sia invece inferiore.

b)

Per ricavare il **p-valore** relativo all'ipotesi H_1: μ > 500 mi basta calcolare 1 - il p-valore ottenuto al punto **a)**. Infatti, rispondere a questa ipotesi alternativa corrisponde a guardare la probabilità sottesa alla distribuzione a partire dallo z-score -4.167 in poi invece che la probabilità sottesa prima di -4.167.

Quindi **p-valore$_2$ = 1 – 0.0000… ≈ 1.**

Ovvero, sotto l'assunzione di veridicità dell'ipotesi nulla, la probabilità di ottenere per caso una media pari a 475$ o superiore è circa uguale a 1.

ESERCIZIO 6.17

Un test scientifico di astrologia ha selezionato 116 adulti. Per ogni adulto, è stato mostrato a un astrologo l'oroscopo corrispondente a tre profili di personalità, uno appartenente allo stesso adulto (ottenuto tramite un questionario) e altri due riferiti ad altri due adulti scelti a caso dal gruppo sperimentale; all'astrologo è stato chiesto di indicare quale fra i tre profili di personalità era quello proprio dell'adulto. L'astrologo ha formulato previsioni corrette 40 volte su 116.

a) Esplicita le ipotesi per verificare se l'astrologo riesce a fare previsioni più accurate di quelle dovute al caso. Trova il P-valore, prendi una decisione con $\alpha = 0.05$ e interpreta i risultati.

RISOLUZIONE:

a)

Per verificare se l'astrologo riesce a fare previsione più accurate di quelle dovute al caso mi devo prima chiedere quale proporzione di risposte corrette mi attendo se l'astrologo tirasse a indovinare.

Ci sono 116 adulti. Essendoci 3 profili di personalità per ogni adulto ed essendoci solo un profilo di personalità corretto per ogni adulto, la probabilità di indovinare casualmente è pari a 1/3. Quindi $\pi = 0.33$.

Quindi le ipotesi che formulo sono: $H_0: \pi \leq 0.33$ vs $H_1: \pi > 0.33$.

La proporzione di risposte giuste fatte dall'astrologo è pari a: $\boldsymbol{\hat{\pi} = 40/116 = 0.345}$. Quindi la proporzione campionaria è $\boldsymbol{\hat{\pi} = 0.345}$.

Come primo passo per verificare l'ipotesi, calcolo la statistica test z associata alla proporzione campionaria.

La formula è: $\boldsymbol{z} = \frac{\hat{\pi}-\pi_0}{se_0}$; dove $\boldsymbol{\pi_0}$ è la proporzione assunta sotto l'ipotesi nulla e dove $\boldsymbol{se_0}$ è l'errore standard della distribuzione assunta sotto l'ipotesi nulla. $\boldsymbol{se_0} = \sqrt{\frac{\pi_0*(1-\pi_0)}{n}}$.

Quindi: $\boldsymbol{se_0} = \sqrt{\frac{0.33*(1-0.33)}{116}}$ **= 0.044.**

Quindi: $\boldsymbol{z} = \frac{0.345-0.33}{0.044}$ **= 0.34.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di 0.34. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 0.34 la probabilità è pari a 0.3669.

Quindi: **p-valore = 0.3669**

L'interpretazione del **p-valore** è: Sotto l'assunzione della veridicità dell'ipotesi nulla, la probabilità di ottenere per caso una proporzione pari a 0.345 o superiore è uguale a 0.3669 (36.69%).

Per prendere una decisione con un livello di significatività $\alpha = 0.05$ confronto il **p-valore** che ho ottenuto e vedo se è inferiore ad α.

$0.3669 < 0.05$.

Essendo il p-valore superiore NON rifiuto l'ipotesi nulla.

ESERCIZIO 6.19

In un'elezione a sindaco a Madison, Wisconsin, vi sono due candidati: il candidato A e il candidato B. La metà esatta dei residenti attualmente è a favore di ciascuno di essi.

a) Dato un campione casuale di 400 votanti, 230 hanno votato per il candidato B. Saresti disposto/a a prevedere che i residenti preferiranno il candidato B? Perché?

b) Dato un campione casuale di 40 votanti, 23 hanno votato per il candidato B. Saresti disposto/a a prevedere che i residenti preferiranno il candidato B? Perché?

RISOLUZIONE:

a)

Per verificare se sono in grado di fare una previsione faccio una verifica di ipotesi. Mi chiedo se, dati i risultati ottenuti in questo campione (400 partecipanti), ho buoni motivi di ritenere che la popolazione preferisca il candidato B. Una preferenza in questo caso è esemplificata da una proporzione di voti superiore a 0.50.

Quindi le ipotesi che formulo sono: H_0: $\pi \leq 0.50$ vs H_1: $\pi > 0.50$.

La proporzione di voti a favore per il candidato B è pari a: **$\hat{\pi}$ = 230/400 = 0.575**. Quindi la proporzione campionaria è **$\hat{\pi}$ = 0.575**.

Come primo passo per verificare l'ipotesi, calcolo la statistica test z associata alla proporzione campionaria.

La formula è: $z = \frac{\hat{\pi}-\pi_0}{se_0}$; dove π_0 è la proporzione assunta sotto l'ipotesi nulla e dove se_0 è l'errore standard della distribuzione assunta sotto l'ipotesi nulla. $se_0 = \sqrt{\frac{\pi_0 * (1-\pi_0)}{n}}$.

Quindi: $se_0 = \sqrt{\frac{0.50 * (1-0.50)}{400}}$ **= 0.025.**

Quindi: $z = \frac{0.575-0.50}{0.025}$ **= 3.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di 3. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 3 la probabilità è pari a 0.0013.

Quindi: **p-valore = 0.0013.**

L'interpretazione del **p-valore** è: Sotto l'assunzione della veridicità dell'ipotesi nulla, la probabilità di ottenere per caso una proporzione pari a 0.575 o superiore è pari a 0.0013 (0.13%).

Questo risultato mi sta a indicare che: se nella popolazione non ci fosse una preferenza per il candidato B, aver ottenuto casualmente una proporzione di voti pari a 0.575 (57.5%) sarebbe molto improbabile. Infatti, la probabilità di ottenere un risultato di tale entità o superiore per motivi di variabilità campionaria è appena del 0.13%.

Ovvero, ottenere 230 voti su 400 per il candidato B è abbastanza improbabile se immaginassimo che la popolazione non esprimesse preferenze (= 200 voti su 400).

Se decidessi di prendere una decisione statistica fissando $\alpha = 0.05$ rifiuterei l'ipotesi nulla. ($0.0013 < 0.05$).

Quindi sarei disposto a prevedere che i residenti preferiscano il candidato B.

b)

La risoluzione di questo punto è identica a quella del punto precedente, ciò che cambia sono i valori. Infatti, in questo caso n = 40 e i voti per il candidato B sono 23.

Quindi le ipotesi che formulo sono: $H_0: \pi \leq 0.50$ vs $H_1: \pi > 0.50$.

La proporzione di voti a favore per il candidato B è pari a: **$\hat{\pi}$ = 23/40 = 0.575**. Quindi la proporzione campionaria è **$\hat{\pi}$ = 0.575**. Noto che è la stessa del punto **a)**.

$\boldsymbol{se_0} = \sqrt{\frac{\mathbf{0.50*(1-0.50)}}{\mathbf{40}}}$ **= 0.079.** Noto che invece l'errore standard è maggiore.

$\boldsymbol{z} = \frac{\mathbf{0.575-0.50}}{\mathbf{0.079}}$ **= 0.95.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di 0.95. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 0.95 la probabilità è pari a 0.1711.

Quindi: **p-valore = 0.1711.**

L'interpretazione del **p-valore** è: Sotto l'assunzione della veridicità dell'ipotesi nulla, la probabilità di ottenere per caso una proporzione pari a 0.575 o superiore è uguale a 0.1711 (17.11%).

In questo caso i risultati mi dicono che: se nella popolazione non ci fosse una preferenza per il candidato B, aver ottenuto casualmente una proporzione di voti pari a 0.575 (57.5%) sarebbe relativamente probabile. Infatti, la probabilità di ottenere un risultato di tale entità o superiore per motivi di variabilità campionaria è del 17.11%.

Ovvero, ottenere 23 voti su 40 per il candidato B non è così improbabile se immaginassimo che la popolazione non esprimesse preferenze (= 20 voti su 40).

Se decidessi di prendere una decisione statistica fissando $\alpha = 0.05$ NON rifiuterei l'ipotesi nulla. ($0.1711 < 0.05$).

Quindi in questo caso NON sarei disposto a prevedere che i residenti preferiscano il candidato B.

ESERCIZIO 6.23

Jones e Smith conducono separatamente uno studio per verificare H_0: $\mu \leq 500$ contro H_1: $\mu > 500$, ciascuno con $n = 1000$. Jones ottiene $\bar{x} = 516.4$, con $s = 316.23$. Smith ottiene $\bar{x} = 516.5$, con $s = 316.23$.

a) Mostra che z = 1.64 e P-valore = 0.0505 per Jones e mostra che z = 1.65 e P-valore = 0.0495 per Smith.

b) Scegliendo $\alpha = 0.050$, indica per ciascun studio se il risultato è "statisticamente significativo".

RISOLUZIONE:

a)

Le ipotesi sono: H_0: $\mu \leq 500$ vs H_1: $\mu > 500$

<u>Jones:</u>

La formula per calcolare lo z-score è: $\boldsymbol{z} = \frac{\bar{x} - \mu_0}{se_0}$; dove $\boldsymbol{\mu_0}$ è la media assunta sotto l'ipotesi nulla e dove $\mathbf{se_0}$ è l'errore standard della distribuzione assunta sotto l'ipotesi nulla. $\boldsymbol{se_0} = \frac{s}{\sqrt{n}}$.

Quindi: $\boldsymbol{se_0} = \frac{316.23}{\sqrt{1000}}$ **= 10.**

Quindi: $\boldsymbol{z} = \frac{516.4 - 500}{10}$ **= 1.64.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di 1.64. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 1.64 la probabilità è pari a 0.0505.

Quindi: **p-valore = 0.0505.**

<u>Smith:</u>

La formula per calcolare lo z-score è: $\boldsymbol{z} = \frac{\bar{x} - \mu_0}{se_0}$; dove $\boldsymbol{\mu_0}$ è la media assunta sotto l'ipotesi nulla e dove $\mathbf{se_0}$ è l'errore standard della distribuzione assunta sotto l'ipotesi nulla. $\boldsymbol{se_0} = \frac{s}{\sqrt{n}}$.

Quindi: $\boldsymbol{se_0} = \frac{316.23}{\sqrt{1000}}$ **= 10.**

Quindi: $\boldsymbol{z} = \frac{516.5 - 500}{10}$ **= 1.65.**

Adesso che ho trovato la statistica test cerco, sulla tavola z a pag. 286 del libro di Luccio, la probabilità associata a un punto z di 1.65. Questa probabilità corrisponde al suo **p-valore**. Vedo che per uno z-score di 1.65 la probabilità è pari a 0.0495.

Quindi: **p-valore = 0.0495.**

b)

Un risultato è statisticamente significativo se il suo **p-valore** è inferiore al livello di significatività α prefissato. In questo caso $\alpha = 0.05$. Vedo che:

Per <u>Jones</u>: $0.0505 > 0.05$ Quindi NON è statisticamente significativo.

Per <u>Smith</u>: $0.0495 < 0.05$ Quindi è statisticamente significativo.