**Intervalli di confidenza per la media**

1. Si desidera stimare il punteggio medio SAT (Scholastic Assessment Test) per tutti gli studenti che hanno frequentato il corso di preparazione al test SAT di Ethan-Davies negli ultimi 2 anni. Si sceglie un campione casuale semplice di 100 studenti da un elenco completo di tutti gli studenti di Ethan-Davies che hanno frequentato il corso negli ultimi due anni, rilevando che il punteggio medio SAT per il campione è pari a 1940 punti. Supponiamo di conoscere che la deviazione standard della popolazione è di 83 punti.

a. Produrre un intervallo di confidenza del 95% del punteggio medio al test SAT per la popolazione degli studenti Ethan-Davis.

b. Interpretare l'intervallo.

2. In uno studio incentrato sulla carenza di farmacisti negli Stati Uniti, è stato selezionato un campione casuale semplice di 250 farmacisti di recente assunzione a cui è stato somministrato un questionario. Una delle domande rivolte agli intervistati ha riguardato l’entità del "bonus di firma" (un incentivo economico) che, come neoassunto, ha ricevuto per unirsi al suo attuale datore di lavoro. Il bonus medio campionario è risultato pari a $4789 (fonte: ASHP Staffing Survey, ashp.org). Svolgere i seguenti punti.

a. Stimare il bonus di firma medio per la popolazione di farmacisti di recente assunzione negli Stati Uniti mediante un intervallo di confidenza al 95%, assumendo che la deviazione standard del bonus di firma nella popolazione sia pari a $3000.

b. Interpretare l'intervallo di confidenza trovato al punto a.

3. Si desidera stimare gli anni di anzianità media dei dipendenti che lavorano per Kaneko Ltd. Viene selezionato casualmente un campione di 49 lavoratori dell’azienda. L'anzianità media campionaria è di 13.6 anni. Supponiamo di sapere che la deviazione standard della popolazione sia di 5.2 anni.

a. Costruire e interpretare un intervallo di confidenza al 95% dell'anzianità media per la popolazione dei dipendenti dell’azienda.

b. Costruire un intervallo di confidenza al 80%.

4. Certification Magazine ha condotto uno studio sulle retribuzioni dei professionisti IT (Information Tecnology), contattando un campione casuale semplice di 7130 specialisti IT certificati. Lo stipendio medio del campione era $96677 (fonte: CertMag.com). Svolgere i seguenti punti.

a. Stimare lo stipendio medio per la popolazione di professionisti IT rappresentati dal campione mediante un intervallo di confidenza. Utilizzare un livello di confidenza del 99% e assumere che la deviazione standard per la popolazione degli stipendi IT sia di $8200.

b. Interpretare l'intervallo ottenuto al punto a.

5. Fare riferimento all'esercizio 3. Supponiamo che la dimensione del campione sia 20 anziché 49. Supponiamo anche che la deviazione standard della popolazione sia sconosciuta e la deviazione standard campionaria sia pari a 5.2 anni.

a. Determinare gli intervalli di confidenza al 95%, 99% e 80% per l’anzianità media della popolazione dei dipendenti dell’azienda e commentare il risultato.

6. L’ufficio Controllo Qualità di Hershey Mechanical vuole stimare la vita media delle nuove guarnizioni O-ring che l'azienda ha appena introdotto nel mercato. Su un campione casuale semplice di 25 O-ring si osserva una vita media di 1560 ore. La deviazione standard campionaria è di 68 ore.

a. Costruire e interpreta l'intervallo di confidenza al 95% della vita media della popolazione complessiva degli O-ring prodotti dall'azienda.

7. Sora-Tobu-Jutan Tecnologies vuole stimare il tempo medio necessario per completare l'assemblaggio del suo nuovo robot aspirapolvere. Dopo aver fatto trascorrere un po’ di tempo per permettere ai lavoratori di apprendere bene i nuovi procedimenti di assemblaggio, i supervisori scelgono 15 assemblaggi da analizzare. Per questi 15 assemblaggi, il tempo di assemblaggio medio è di 38.4 minuti. La deviazione standard campionaria dei tempi di assemblaggio è pari a 5.6 minuti.

a. Indicare la popolazione oggetto di studio.

b. Costruire e interpretare una stima intervallare del tempo di assemblaggio medio necessario a completare il prodotto da parte dell’azienda. Scegliere a questo scopo un livello di confidenza del 95%. Assumere che la distribuzione della popolazione dei tempi di montaggio sia normale.

8. Fence-U-In sta testando la forza di rottura della catena che l'azienda produce per i suoi prodotti di recinzione. Supponiamo che venga selezionato un campione di cinque catene e che vengano rilevati i seguenti risultati:

|  |  |
| --- | --- |
| **Collegamento** | **Forza di rottura (in libbre)** |
| 1 | 840 |
| 2 | 820 |
| 3 | 790 |
| 4 | 850 |
| 5 | 700 |

Calcolare l'intervallo di confidenza al 90% per la forza di rottura media, assumendo che la forza di rottura dell’intera popolazione di catene sia approssimativamente normale.

9. Nell'ambito di uno studio di marketing, Sleep ‘n’ Dream vuole stimare, per i clienti che hanno recentemente acquistato un materasso a molle, il numero medio di negozi visitati prima di effettuare il loro acquisto. Su un campione di sei clienti selezionato in modo casuale si osservano i seguenti risultati:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Visite del negozio** | 6 | 10 | 4 | 5 | 7 | 4 |

Utilizzare i dati dell’esempio per costruire una stima intervallare del numero medio di visite per la popolazione di clienti che ha recentemente acquistato un materasso a molle, con livello di confidenza del 90%. Supponiamo che i valori siano distribuiti nella popolazione in modo approssimativamente normale.

10. La Koufos Importazioni vuole stimare il valore medio dei suoi crediti correnti. Nel complesso, ci sono 2000 conti. Viene selezionato un campione casuale semplice di 40 conti. L'importo medio dovuto all'impresa nel campione è di $6500. La deviazione standard campionaria è di $1600. Svolgere i seguenti punti.

a. Determinare l'intervallo di confidenza al 95% del credito medio.

11. Dei ricercatori intendono stimare la produzione giornaliera media per le 2300 aziende dell'industria metallurgica. Decidono di selezionare in modo casuale 60 aziende la cui produzione giornaliera media è di 1560 unità, con una deviazione standard di 320 unità.

a. Costruire un intervallo di confidenza al 99% della produzione media giornaliera per azienda, riferita alla popolazione di aziende del settore metallurgico.

b. Stimare la produzione giornaliera totale dell'industria metallurgica nel suo complesso, con un intervallo di confidenza al 99%.

12. Si segnala che i contribuenti americani spendono 6.6 miliardi di ore a compilare annualmente i moduli fiscali (fonte: all-headlinenews.com). Supponiamo di selezionare un campione casuale di 1500 contribuenti e di scoprire che il tempo medio campionario di compilazione dei moduli fiscali è di 28.4 ore, con una deviazione standard di campione di 5.7 ore. Svolgere i seguenti punti.

a. Calcolare un intervallo di confidenza al 95% per il tempo medio necessario a compilare i moduli fiscali per la popolazione di tutti i contribuenti americani.

13. Per testare l’efficacia pulente di due diversi detersivi per bucato, ADG Labs seleziona casualmente cinque diverse pezze di tessuto e sporca ciascuna di esse in modo uniforme. Ogni pezza viene tagliata a metà. Per ogni pezza sporcata e tagliata, una metà selezionata a caso viene lavata con il detergente A e l'altro viene lavata con il detergente B. Viene quindi utilizzato uno strumento particolare che misura la pulizia di ciascun campione lavato. La tabella seguente mostra i risultati dell’esperimento. Supponendo che tutte le condizioni necessarie riguardanti la popolazione siano soddisfatte, utilizzare l'approccio per dati appaiati per costruire un intervallo di confidenza al 95% della differenza media dell’efficacia pulente dei due detersivi.

Livelli di pulizia delle pezze lavate

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Pezza** | **1** | **2** | **3** | **4** | **5** |
| **Detersivo 1** | 91 | 86 | 88 | 92 | 88 |
| **Detersivo 2** | 83 | 78 | 79 | 86 | 84 |

14. Per valutare l'efficacia di una nuova terapia per la riduzione dell'appetito, vengono selezionati cinque soggetti. Prima che cominci il programma terapeutico, viene misurato l'apporto calorico giornaliero dei cinque partecipanti. Dopo il trattamento, l'apporto calorico dei cinque partecipanti viene nuovamente misurato. La tabella seguente mostra i risultati osservati prima e dopo il trattamento. Per questo esperimento con campioni dipendenti, costruire un intervallo di confidenza al 90% della differenza media nell'apporto calorico giornaliero prima e dopo il trattamento.

Assunzione di calorie

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Soggetto** | **1** | **2** | **3** | **4** | **5** |
| **Prima** | 2340 | 2190 | 1960 | 3450 | 2960 |
| **Dopo** | 1950 | 2230 | 1670 | 2920 | 2475 |

ESERCIZI VERIFICA D’IPOTESI

1. Le ipotesi in contrapposizione di un test sono le seguenti:

Un campione casuale di dimensione 64 viene prelevato dalla popolazione di interesse. La media campionaria è pari a 1125. Si assuma che la deviazione standard della popolazione sia 80. In base ai dati campionari, al livello di significatività , l’ipotesi nulla è da rifiutare?

1. Le ipotesi in contrapposizione di un test sono le seguenti:

Un campione casuale di dimensione 100 viene prelevato dalla popolazione di interesse. La media campionaria è 1955. Si assuma che la deviazione standard della popolazione sia 150. In base ai dati campionari, al livello di significatività , l’ipotesi nulla è da rifiutare?

1. Le ipotesi in contrapposizione di un test sono le seguenti:

Un campione casuale di dimensione 36 viene prelevato dalla popolazione di interesse. La media campionaria è pari a 2525. Si assuma che la deviazione standard della popolazione sia di 120. In base ai dati campionari, al livello di significatività , l’ipotesi nulla è da rifiutare?

4. Secondo i ricercatori della Carnegie Mellon University, la vita media dell’unità hard drive (disco rigido) di un computer è di 600000 ore, ovvero 3.1 anni (fonte: science.newsfactor.com). Si supponga che un’azienda affermi che la durata dei suoi dischi rigidi è più lunga, offrendo come prova uno studio basato su un campione di 64 dischi rigidi di propria produzione. Se la durata media campionaria è di 3.22 anni, possiamo affermare che c’è evidenza empirica a supporto dell’affermazione dell’azienda? Si utilizzi un livello di significatività e si supponga che la deviazione standard della durata nella popolazione di dischi rigidi prodotti dall’azienda sia pari a anni. Si formalizzi il problema come una verifica delle seguenti ipotesi:

5. La Lundberg Survey, che riporta i prezzi del carburante presso le stazioni di rifornimento a livello nazionale, ha recentemente riferito che il prezzo medio della benzina negli Stati Uniti è di $0.90 al litro (fonte: moneynews.com). Tuttavia, gran parte dei cittadini di Los Angeles crede che il prezzo medio attuale per la popolazione delle stazioni di rifornimento di questa città sia più alto. Si decide di selezionare un campione casuale di 49 stazioni a Los Angeles e si rileva che il prezzo medio della benzina per le stazioni campionate è pari a $0.91 al litro. Possiamo affermare che l’evidenza empirica supporta l’ipotesi alternativa e che dovremmo rifiutare l'ipotesi nulla che il prezzo medio della benzina delle stazioni di servizio di Los Angeles sia non superiore a $0.90, ad un livello di significatività ? Si assume che la deviazione standard dei prezzi della benzina per la popolazione costituita da tutti le stazioni di servizio di Los Angeles sia di $0.05 al litro.

6. Negli anni passati, il ​​numero medio di errori presenti nelle relazioni emesse dalla sezione contabile della Sterling Thompson, Inc. era pari a 5.2. L'azienda ora adotta un nuovo sistema. In un campione casuale di 36 recenti rapporti emessi utilizzando il nuovo sistema, il numero medio di errori è stato pari a 4.6. Alla luce di questo dato, possiamo sostenere l'affermazione che il tasso di errore medio per tutti i rapporti emessi con il nuovo sistema sia inferiore alla precedente media di 5.2? Si utilizzi per tale verifica un livello di significatività pari a 0.01. Si ponga come ipotesi nulla e come ipotesi alternativa. Si assuma inoltre che la deviazione standard di errori nella popolazione di tutte le nuove relazioni emesse con il nuovo sistema sia pari a 1.8.

7. Mediabistro.com riporta che, in media, le aziende negli Stati Uniti hanno 14709 follower su Twitter. Si seleziona un campione casuale di 36 aziende e si osserva un numero medio campionario di follower su Twitter pari a 12950. Questo dato può rappresentare evidenza empirica a favore dell’ipotesi che il numero medio di follower su Twitter per la popolazione delle imprese statunitensi sia inferiore alla media nazionale? Si utilizzi un livello di significatività pari a . Si supponga che la deviazione standard della popolazione sia pari a 4200 follower. Le ipotesi da verificare sono:

8. Uno studio recente ha riportato che Comcast Cable era il fornitore di banda larga più veloce negli Stati Uniti, con una velocità media di download pari a 17.2 megabit al secondo (fonte: gigaom.com). In un campione casuale di 49 osservazioni delle prestazioni di Comcast nella grande area di Minneapolis, si è osservata una velocità media di download pari a 16.7 megabit al secondo. Possiamo affermare che questo risultato rappresenta evidenza empirica a supporto dell’ipotesi che la velocità media di download della Comcast nell’area di Minneapolis sia più lenta di 17.2 megabit per secondo? Utilizzare un livello di significatività uguale a e si assuma che la deviazione standard della popolazione sia di 1.4 megabit al secondo. Le ipotesi da verificare sono:

9. Le ipotesi contrapposte per un particolare test di ipotesi sono di seguito indicate:

Si assuma che la deviazione standard della popolazione sia nota e pari a 80 km e che si selezioni un campione casuale di dimensione 64. Usando un livello di significatività pari a , si svolga quanto segue.

a. Esprimere la regola decisionale espressa in km (rispetto alla media campionaria).

b. Se la media campionaria è pari a km, l’ipotesi nulla dovrebbe essere rifiutata? Motivare la risposta.

10. Le ipotesi contrapposte per un particolare test di ipotesi sono di seguito indicate:

Si assuma che la deviazione standard della popolazione sia nota e pari a 140 secondi e che venga selezionato un campione casuale di dimensione 49. Usando un livello di significatività pari a , svolgere i seguenti punti.

a. Se la media campionaria fosse di secondi, si dovrebbe rifiutare l'ipotesi nulla? Motivare la risposta.

11. Le ipotesi contrapposte per un particolare test di ipotesi sono di seguito indicate:

Si assuma che la deviazione standard della popolazione sia nota e pari a $60 e che venga selezionato un campione casuale di dimensione 36. Usando un livello di significatività pari a , si svolgano i seguenti punti.

a. Se la media campionaria è pari a $266, si dovrebbe rifiutare ipotesi nulla? Motivare la risposta.

12. Secondo il National Center for Health Statistics, l'altezza media di una donna adulta negli Stati Uniti Stati è pari a 161.8 cm (fonte: pediatrics.about.com). Si sospetta che una scarsa nutrizione infantile possa influire negativamente sull'altezza delle donne adulte provenienti da famiglie a basso reddito. Si decide di selezionare un campione casuale di 100 donne adulte cresciute in famiglie a basso reddito e si scopre che l'altezza media campionaria è di 160.5 cm. È possibile utilizzare questo risultato per sostenere l'ipotesi che l'altezza media delle donne provenienti da famiglie a basso reddito è inferiore alla media nazionale pari a 161.8 cm? Si definisca la regola decisionale per il test rispetto alla media campionaria espressa in cm. Si utilizzi un livello di significatività pari al e una deviazione standard di popolazione pari a cm. Si consideri come ipotesi nulla e come ipotesi alternativa .

13. La media settimanale delle ore di lavoro dagli operai impiegati nel settore manifatturiero degli Stati Uniti è ufficialmente pari a (fonte: bls.gov). Si sospetta che la media in una determinata zona di interesse sia più alta. Si decide di selezionare un campione casuale di addetti alla produzione delle aziende manifatturiere locali e si rileva che la media settimanale campionaria delle ore di lavoro è pari a ore. Questo dato può essere utilizzato per sostenere che la settimana lavorativa media nella produzione locale è più lunga di quella nazionale dichiarata di ore? Esprimere la regola decisionale appropriata rispetto alla media campionaria in ore. Si utilizzi un livello di significatività di e si assuma una deviazione standard di popolazione pari a ore. Si consideri come ipotesi nulla e come ipotesi alternativa ).

14. Lo stato della Florida ha riferito che il sussidio di disoccupazione medio per le persone diventate disoccupate per le conseguenze dei danni creati dall’uragano Charley è stato di $224 (fonte: floridajobs.org). Si supponga di selezionare un campione casuale di 200 lavoratori disoccupati a causa dell'uragano Charley e di osservare un sussidio di disoccupazione medio campionario di $207. A fronte di questo dato, è possibile smentire l’affermazione dello stato della Florida e concludere che il sussidio di disoccupazione medio è in realtà inferiore a $224? Indicare l'appropriata regola decisionale espressa rispetto alla media campionaria in dollari. Si utilizzi un livello di significatività di e una deviazione standard della popolazione di $206. Si consideri come ipotesi nulla e come ipotesi alternativa .

15. Negli ultimi cinque anni, il tempo medio per risolvere i reclami dei clienti presso la Allen Retailing è stato di 3.4 giorni. La Allen ha recentemente istituito una nuova procedura di elaborazione dei reclami e la ha testata su un campione casuale di 75 reclami. Il tempo medio per risolvere i reclami nel campione è risultato pari a 2.9 giorni. A fronte di questo risultato, si può sostenere che il tempo medio per la risoluzione dei reclami utilizzando la nuova procedura sia inferiore a 3.4 giorni? Si indichi la regola decisionale appropriata rispetto alla media in giorni. Si utilizzi un livello di significatività di 0.05 e una deviazione standard della popolazione pari a 1.2 giorni.

16. Le ipotesi contrapposte per un certo test sono di seguito definite:

Si supponga che la deviazione standard della popolazione sia nota e pari a 80. Su un campione casuale di dimensione 64 si osserva una media campionaria di 1020.

a. Cacolare il p-value del test.

b. Decidere in merito al rifiuto dell’ipotesi nulla con un livello di significatività pari a 0.05. Decidere in merito al rifiuto dell’ipotesi nulla con un livello di significatività pari a 0.01. Motivare le risposte.

17. Le ipotesi contrapposte di un test sono di seguito definite:

Si supponga che la deviazione standard della popolazione sia nota e pari a 36. Un campione casuale di dimensione 81 ha una media campionaria di 488.

a. Calcolare il p-value del test.

b. Decidere in merito al rifiuto dell’ipotesi nulla con un livello di significatività pari a 0.10. Decidere in merito al rifiuto dell’ipotesi nulla con un livello di significatività pari a 0.05. Motivare le risposte.

18. Le ipotesi contrapposte di un test sono di seguito definite:

Si supponga che la deviazione standard della popolazione sia nota e pari a 200. Un campione casuale di dimensione 100 ha una media campionaria di 2560.

a. Calcolare il p-value del test.

b. Decidere in merito al rifiuto dell’ipotesi nulla con un livello di significatività pari a 0.01. Decidere in merito al rifiuto dell’ipotesi nulla con un livello di significatività pari a 0.10. Motivare le risposte.

c. Qual è il livello di significatività minimo per il quale verrebbe rifiutata l’ipotesi nulla?

20. Apple riferisce che mediamente un download di una app dal suo App Store porta allo sviluppatore un guadagno pari a 17.5 centesimi (fonte: Forbes.com). Un gruppo di sviluppatori di app speciali afferma che la media per le loro app è inferiore. In un campione di 100 delle loro app, il gruppo rileva un guadagno medio per download pari a 16.2 centesimi.

a. Considerando 17.5 centesimi come ipotesi nulla, calcolare il p-value del test da applicare per verificare se il gruppo di sviluppatori ha ragione. Si assuma una deviazione standard della popolazione di 5.4 centesimi.

b. Il gruppo di sviluppatori di app può utilizzare questo risultato campionario per sostenere che il guadagno medio ottenuto dalle loro apps è inferiore a 17.5 centesimi, se il livello di significatività è fissato a 0.05? E se è pari a 0.01? Motivare le risposte.

21. Gethuman.com sostiene che il tempo medio di attesa per l'opzione del servizio clienti Live Chat di Amazon è di 5 minuti (fonte: gethuman.com). In un campione casuale di 100 tentativi di usare la Live Chat, si scopre che il tempo medio di attesa è di 5 minuti e 30 secondi.

a. Considerando minuti come ipotesi nulla, calcolare il p-value del test sulla media del tempo di attesa per il servizio Live Chat. Si assuma una deviazione standard della popolazione di 3.4 minuti.

b. Con un livello di significatività pari a 0.05, l’ipotesi nulla secondo la quale il tempo medio di attesa sia non superiore a 5 minuti è da rifiutare? E con un livello di significatività di 0.01?

c. Qual è il minimo livello di significatività per il quale avremmo un risultato statisticamente significativo?

22. Secondo l'ultima edizione dell’indagine Zagat, il costo medio per una cena nei migliori ristoranti statunitensi è $35.37. In un sondaggio condotto su 36 ristoranti locali di una certa zona selezionati casualmente, si rileva un costo medio campionario per una cena di $32.80. Tale evidenza empirica consente di sostenere che la media locale è inferiore alla media nazionale di $35.37 dichiarata da Zagat? Si utilizzi un livello di significatività di 0.05 e si assuma una deviazione standard dei prezzi per la popolazione di ristoranti locali pari a $6.60. Calcolare il p-value del test e spiegarne il significato.

23. Si supponga di testare le seguenti ipotesi sulla media:

La dimensione campionaria è 81. La media campionaria è 975. La deviazione standard della popolazione è pari a 90. Il livello di significatività sia 0.05.

a. Calcolare il valore osservato della statistica test .

b. Calcolare il p-value appropriato.

c. Decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

25. Si supponga di testare le seguenti ipotesi:

La dimensione del campione è 64. La media campionaria è 107. La deviazione standard della popolazione è pari a 36. Il livello di significatività prefissato è 0.01.

a. Calcolare il valore osservato della statistica test .

b. Calcolare il p-value appropriato.

c. Decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

26. Si supponga di voler verificare le seguenti ipotesi:

La dimensione campionaria è 36. La media campionaria è 940. La deviazione standard della popolazione è pari a 120. Il livello di significatività è 0.05.

a. Quali sono i valori critici del test? Indicare la regola decisionale.

b. Calcolare il valore osservato della statistica test .

c. Decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

27. Si supponga di voler verificare le seguenti ipotesi:

La dimensione del campione è pari a 100. La media campionaria è uguale a 635. La deviazione standard della popolazione è pari a 140. Il livello di significatività prefissato è 0.10.

a. Calcolare il valore osservato della statistica test .

b. Calcolare il p-value appropriato.

c. Decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

28. Si supponga di testare le seguenti ipotesi:

La dimensione campionaria è uguale a 144. La media campionaria è pari a 7040. La deviazione standard della popolazione è 240. Si utilizzi un livello di significatività pari a 0.01.

a. Calcolare il p-value appropriato.

b. Decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

29. Si supponga di testare le seguenti ipotesi:

La dimensione del campione è 49. La media campionaria è 68. La deviazione standard della popolazione è pari a 21. Si utilizzi un livello di significatività pari a 0.05.

a. Calcolare il p-value appropriato.

b. Decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

30. Il Centro nazionale per le statistiche educative (NCES) riporta che il punteggio medio degli studenti della 12a classe sottoposti al test di lettura nazionale è stato 295 (su una scala da 1 a 500) (fonte: nces.ed.gov). Un campione casuale di 100 alunni della 12a classe nelle scuole della contea di Blanchet, in Alabama, ha recentemente sostenuto il test. Gli studenti del campione hanno ottenuto un punteggio medio di 301. Date le seguenti ipotesi:

(la media della popolazione di studenti di Blanchet è uguale a quella nazionale)

(la media della popolazione di studenti di Blanchet è diversa da quella nazionale).

a. Calcolare il p-value per questo test a due code. Si assuma una deviazione standard della popolazione di 38 punti.

b. Decidere in merito al rifiuto o meno dell’ipotesi nulla ponendo . Motivare la risposta.

31. Si stima che il tempo di inattività medio per i computer presso le aziende di tutto il paese è pari a 175 ore all'anno (fonte: networkworld.com). In un campione di 50 aziende situate al di fuori delle principali aree metropolitane, il tempo medio di inattività è stato di 192 ore. Date le seguenti ipotesi:

(la media per le imprese al di fuori delle aree metropolitane è uguale a quella nazionale)

(la media per le imprese al di fuori delle aree metropolitane non è uguale a quella nazionale)

a. Calcolare il p-value per questo test a due code. Si assuma una deviazione standard della popolazione di 43 ore.

b. Usando un livello di significatività di 0.05, decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

32. La C-Net.com riporta che la durata media della carica di una batteria attiva (da una ricarica alla successiva) per il telefono modello 1661 di Nokia è di 9.3 ore (fonte: reviews.cnet.com). In un test basato su 64 ricariche, si rileva una durata media campionaria della carica della batteria per questo modello di 8.9 ore. Date le seguenti ipotesi:

(la media della durata della carica di una batteria per il Nokia 1661 è di 9.3 ore)

(la media della durata della carica di una batteria per il Nokia 1661 non è di 9.3 ore)

a. Calcolare il p-value per questo test a due code. Si assuma una deviazione standard della popolazione di 2.2 ore.

b. Usando un livello di significatività di 0.05, decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

33. L'Internal Revenue Service (IRS) riferisce che il rimborso medio dall’imposizione fiscale nell’anno passato è stato di $2236. L'IRS ritiene che il rimborso medio di quest'anno avrà la stessa media dell'anno precedente. La media campionaria del rimborso in un campione casuale di 500 casi di quest’anno è $2350. Date le seguenti ipotesi:

(la media del rimborso quest’anno è uguale a quella dell’anno precedente)

(la media del rimborso quest’anno non è uguale a quella dell’anno precedente)

a. Calcolare il p-value del test. Si assuma una deviazione standard della popolazione di $724.

b. Usando un livello di significatività di 0.05, decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

34. La Parkrose Bakery cuoce grandi quantità di prodotto in enormi forni. I fornai richiedono che a metà del ciclo di cottura, la temperatura media per i prodotti sia di 325 gradi F. Un lotto con temperatura media inferiore o superiore richiederà una regolazione diversa nelle impostazioni del forno. Per verificare questa condizione, un campione casuale di 36 prodotti viene selezionato casualmente a metà di ogni ciclo e viene calcolata la temperatura media campionaria dei dati osservati.

a. Supponendo che il campione abbia una temperatura media di 323.4 gradi, calcolare il p-value del test. Si assuma una deviazione standard della popolazione di 4.5 gradi.

b. Scegliendo un livello di significatività di 0.01, decidere in merito al rifiuto o meno dell’ipotesi nulla. Motivare la risposta.

35. Si supponga di testare le seguenti ipotesi:

e

La dimensione campionaria è 25. La media campionaria è 1545 e la deviazione standard campionaria è uguale a 75. Si usi un livello di significatività di 0.05 e si assuma che la distribuzione della popolazione sia normale

a. Calcolare il valore osservato della statistica test .

b. Decidere in merito al rifiuto o meno dell’ipotesi nulla utilizzando l’approccio del p-value mediante uso delle tavole della distribuzione t di Student.

36. Si considerino le seguenti ipotesi da testare:

e

La dimensione campionaria è 16. La media campionaria è pari a 3920 e la deviazione standard campionaria è uguale a 200. Si usi un livello di significatività di 0.01 e si assuma che la distribuzione della popolazione sia normale.

a. Calcolare il valore osservato della statistica test .

b. Decidere in merito al rifiuto o meno dell’ipotesi nulla mediante l’uso delle tavole della distribuzione t di Student.

37. Si consideri il problema in cui si vuole verificare le seguenti ipotesi:

e

La dimensione campionaria è uguale a 9. La media campionaria è pari a 3260 e la deviazione standard campionaria a 120. Si utilizzi un livello di significatività di 0.10 e si assuma che la distribuzione della popolazione sia normale.

a. Calcolare il valore osservato della statistica test .

b. Decidere in merito al rifiuto o meno dell’ipotesi nulla mediante uso delle tavole della distribuzione t di Student.

38. Secondo i Centri per il Controllo e la Prevenzione delle Malattie, l’aspettativa di vita media negli Stati Uniti è salita a 77.6 anni (fonte: cnn.com). Dal rapporto emerge che i residenti nelle zone rurali del Sud-ovest hanno una vita più breve rispetto alla popolazione generale. In un campione casuale di 250 residenti del Sud-ovest rurale recentemente defunti, la media di vita era di 75.9 anni con una deviazione standard campionaria di 9.8 anni. Tale risultato campionario è sufficiente per rifiutare l'ipotesi nulla sotto definita? Si utilizzi un livello di significatività di 0.05.

(l’aspettativa di vita media per i residenti dell’area rurale del Sud-ovest è almeno quanto quella della popolazione generale)

(l’aspettativa di vita media per i residenti dell’area rurale del Sud-ovest è inferiore a quella della popolazione generale)

39. Secondo la ePaynews.com, l’importo medio delle transazioni online è $187 (fonte: epaynews.com/.) Si supponga di osservare un campione casuale di 50 transazioni online realizzate tramite il sito Web di una nota azienda e di scoprire che l'importo medio campionario per transazione è pari a $195. La deviazione standard nel campione è uguale a $36. Alla luce dei dati, con un livello di significatività di 0.05, possiamo dire che vi sia evidenza empirica per rifiutare l'ipotesi nulla a favore dell’ipotesi alternativa?

(L'importo medio della transazione effettuate sul sito web dell’azienda nota è uguale all’importo medio delle transazioni online.)

(L'importo medio della transazione effettuate sul sito web dell’azienda nota non è uguale all’importo medio delle transazioni online.)

40. Uno studio sugli Annali della Medicina d'Urgenza ha definito che il tempo medio di attesa per i pazienti del pronto soccorso negli ospedali della California (cioè il tempo trascorso dall'arrivo al pronto soccorso fino alla visita di un medico) è di 56 minuti (fonte: acep.org). L’Ospedale di San Luca a Los Angeles, volendo condurre un proprio studio, tiene traccia di un campione casuale di 200 pazienti. Il tempo medio di attesa per i pazienti del campione è di 60.4 minuti. La deviazione standard campionaria è uguale a 28.6 minuti. Alla luce dei dati, al livello di significatività , l'ipotesi nulla secondo la quale il tempo medio di attesa per i pazienti del San Luca è uguale al tempo medio della California va rifiutata?

Verifica ipotesi DUE CAMPIONI

1. Si intende ricorrere a due campioni casuali di dimensione 50 selezionati dalla popolazione 1 e dalla popolazione 2, per verificare le seguenti ipotesi sulla differenza delle medie delle due popolazioni:

1. Si supponga che le medie campionarie siano rispettivamente uguali a e . Calcolare il valore osservato della statistica test, assumendo che le deviazioni standard delle popolazioni siano pari a e .
2. Applicare la regola decisionale basata sulla regione critica per la statistica Z, per decidere in merito al rifiuto dell’ipotesi nulla al livello di significatività .

2. Fare riferimento all’esercizio 1.

a. Calcolare il p-value e spiegarne il significato.

b. Utilizzare il p-value per decidere se l’ipotesi nulla è da rifiutare o meno. Motivare la risposta.

3. Si ha un campione casuale di dimensione 100 per ciascuna di due popolazioni. Si intende usare i dati campionari per verificare le seguenti ipotesi:

1. Si supponga che le medie campionarie siano rispettivamente e . Calcolare il valore osservato della statistica test, assumendo che le deviazioni standard delle popolazioni siano pari a e .
2. Applicare la regola decisionale basata sulla regione critica per la statistica Z, per decidere in merito al rifiuto dell’ipotesi nulla al livello di significatività .

4. Fare riferimento all’esercizio 3.

a. Calcolare il p-value e spiegarne il significato.

b. Utilizzare il p-value per decidere se l’ipotesi nulla è da rifiutare o meno. Motivare la risposta.

5. In uno studio condotto presso gli Ospedali dell’Università di Chicago, sono state confrontate le prestazioni di medici specializzati nella cura di pazienti ospedalizzati con quelle di internisti generici che dedicano solo una piccola parte del loro tempo alla cura dei pazienti ospedalizzati. Una delle variabili esaminate era il tempo trascorso in ospedale dai pazienti sotto la cura di queste due categorie di medici. Su un campione di pazienti curati da internisti è stato rilevato un tempo di permanenza medio di 4.59 giorni, a fronte di un tempo medio pari a 4.10 giorni per i pazienti curati da medici ospedalieri (fonte: uchospitals.edu). Effettuare un test per verificare l’ipotesi che non via sia differenza tra i tempi medi di permanenza in ospedale dei due gruppi di pazienti, usando un livello di significatività di 0.01. Si supponga che le dimensioni campionarie siano 488 (internisti) e 162 (ospedalieri) e le deviazioni standard delle due popolazioni rispettivamente 1.5 giorni e 1.4 giorni.

6. I vigili del fuoco di Boulder hanno condotto uno studio sui tempi medi di risposta alle emergenze delle due stazioni. Su un campione di 125 emergenze per la North Boulder Station, il tempo medio di risposta è stato di 16.2 minuti. Su un campione di 160 emergenze gestite dalla South Boulder Station, il tempo medio di risposta è stato di 14.3 minuti. Sulla base di questi dati possiamo rifiutare l’ipotesi nulla che non vi sia differenza in media nei tempi di risposta al livello di significatività 0.05? Assumere che le deviazioni standard dei tempi di risposta per le due popolazioni di emergenze siano note e pari a 4.5 minuti, per North Boulder, e 5.2 minuti, per South Boulder.

7. In uno studio agricolo effettuato per esaminare l'impatto di una nuova miscela di pesticidi sul raccolto di colza canadese, un campione di 459 aziende agricole che utilizzano il nuovo prodotto

e un campione di 295 aziende agricole che utilizzano quello convenzionale sono stati selezionati casualmente. La media delle entrate lorde per acro, per il campione di aziende che ricorrono al nuovo mix di pesticidi, è stata di $181.90. La media per acro per il campione di fattorie che usano il prodotto convenzionale è stata di $152.10 (fonte: canola-council.org). Effettuare un test di ipotesi ad una coda per determinare se questi dati portano a rifiutare l’ipotesi nulla che la media delle entrate lorde della popolazione di aziende del primo gruppo non sia superiore a quella delle aziende del secondo gruppo. Utilizzare un livello di significatività di 0.05 e assumere che le deviazioni standard delle due popolazioni siano note e pari a $83.40 e $75.20 rispettivamente.

8. Si supponga di osservare un campione casuale di dimensione 10 da ciascuna di due popolazioni normali, e di voler testare le seguenti ipotesi riguardanti le medie delle due popolazioni:

1. Usando un livello di significatività di 0.05, identificare i valori critici per la statistica test T e la regola decisionale.
2. Supponiamo che le medie campionarie siano e . Calcolare il valore osservato della statistica test e decidere in merito al rifiuto o meno dell’ipotesi nulla. Le deviazioni standard campionarie sono e . Assumere che le deviaizoni standard delle due popolazioni siano uguali.
3. Risolvere il problema seguendo l’approccio del p-value.

9. In riferimento all’esercizio 8, assumere che le deviazioni standard campionarie siano entrambe pari a 100 e mostrare come questo influisce sui risultati finali.

10. Si supponga di osservare un campione casuale di dimensione 10 dalla popolazione 1 e un campione casuale di dimensione 13 dalla popolazione 2 e di voler testare le seguenti ipotesi riguardanti le medie delle due popolazioni:

Assumere che le distribuzioni delle due popolazioni siano normali con uguali deviazioni standard.

1. Usare un livello di significatività di 0.01, identificare i valori critici per la statistica test T e la regola decisionale.
2. Supponiamo che le medie campionarie siano e . Calcolare il valore osservato della statistica test e decidere in merito al rifiuto o meno dell’ipotesi nulla. Le deviazioni standard campionarie sono e . Assumere che le deviazioni standard delle due popolazioni siano uguali.
3. Risolvere il problema seguendo l’approccio del p-value.

11. Viene condotto uno studio per confrontare il tempio medio di allenamento di due gruppi di addetti alla sicurezza degli aeroporti. In un campione casuale di addetti alla sicurezza dipendenti pubblici, il tempo medio di allenamento è risultato pari a 68.2 ore, con una deviazione standard campionaria di 10.4 ore. In un campione casuale di addetti alla sicurezza dipendenti privati, il tempo medio di allenamento è stato di 65.4 ore, con una deviazione standard di 12.3 ore. Alla luce di questi dati, è da rigettare l’ipotesi che la differenza tra i tempi medi di allenamento sia nulla al livello di significatività 0.05?

a. Per rispondere alla domanda precedente si supponga che i campioni siano composti da 12 dipendenti pubblici e 16 dipendenti di aziende private e si assuma che le distribuzioni delle popolazioni siano normali e abbiano uguali deviazioni standard

b. Per rispondere alla domanda precedente si supponga che i campioni siano composti da 120 dipendenti pubblici e 160 dipendenti di aziende private.

12. Fare riferimento all’esercizio 11, riguardante la degenza media di pazienti seguiti da due diverse categorie di medici (ospedalieri e internisti generici). Si supponga ora che le due deviazioni standard siano ignote e che vengano rilevati i seguenti dati: i pazienti sotto le cure di internisti generici hanno fatto registrare una permanenza media in ospedale di 4.59 giorni e una deviazione standard di 1.8 giorni, mentre pazienti curati da medici ospedalieri hanno fatto registrare una permanenza media di 4.1 giorni e una deviazione standard di 1.3 giorni. Volendo verificare se vi è o meno differenza non nulla dei tempi medi di permanenza in ospedale, definire le ipotesi statistiche del problema ed eseguire il test al livello di significatività nelle tre distinte condizioni seguenti, che riguardano le dimensioni campionarie:

a. i campioni sono costituiti da 10 pazienti curati da internisti e 15 pazienti curati da ospedalieri. Si assuma che le distribuzioni delle popolazioni siano normali con uguale varianza.

b. I campioni sono costituiti da 100 pazienti curati da internisti e 150 pazienti curati da ospedalieri.

13. Helene Parker Ltd. vuole verificare se due promozioni diverse presso i punti vendita per la sua nuova linea di prodotti per la cura della pelle produrranno diversi livelli di vendita. L’azienda ha abbinato i negozi, creando coppie di punti vendita simili per dimensione e posizione, e ha selezionato casualmente 5 coppie di negozi. Per ogni coppia selezionata, a un negozio scelto a caso è stata assegnata la promozione 1 e all’altro la promozione 2. Alla fine del mese di promozione vengono rilevate le vendite per ciascun negozio. Usando i dati della tabella e assumendo che le condizioni sulla popolazione necessarie per l’applicazione del test siano sodisfatte, condurre un test per verificare l’ipotesi che la differenza delle vendite medie delle due popolazioni di negozi (individuate in base al tipo di promozione) sia pari a zero. Usare un livello di significatività di 0.01.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Punto vendita** | | | | |
| **Tipo di promozione** | **1** | **2** | **3** | **4** | **5** |
| Promozione 1 | 5.6 | 7.6 | 8.4 | 9.2 | 6.0 |
| Promozione 2 | 5.8 | 6.0 | 7.2 | 8.9 | 5.7 |
| Differenza | -0.2 | +1.6 | +1.2 | +0.3 | +0.3 |

14. Montgomery Sports ha selezionato sei giocatori di tennis dell’Università della Georgia per verificare la durabilità delle sue due nuova corde per racchetta in poliestere. La tabella seguente riporta la durata in ore delle corde testate prima che si rompessero. Trattando i sei giocatori come un campione di dati appaiati, e assumendo che tutte le condizioni necessarie per la popolazione siano soddisfatte, condurre un test per verificare l’ipotesi nulla che non vi sia differenza nella durata media dei due diversi tipi di corda. Usare un livello di significatività di 0.10.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Giocatore** | | | | | |
| **Tipo di corda** | **1** | **2** | **3** | **4** | **5** | **6** |
| Corda tipo 1 | 15.2 | 17.6 | 10.4 | 9.2 | 13.4 | 12.6 |
| Corda tipo 2 | 14.3 | 16.0 | 8.2 | 8.9 | 15.7 | 9.3 |

15. Per verificare l’efficacia di un nuovo programma di sensibilizzazione sulla sicurezza, cinque impiegati della XYZ Inc. vengono selezionati casualmente. Prima che partecipino ad un corso sulla sicurezza, ai cinque dipendenti viene chiesto di compilare un questionario che consiste in 100 domande su temi riguardanti la sicurezza. Dopo aver partecipato al corso, gli stessi cinque lavoratori svolgono di nuovo una prova di verifica sulle conoscenze. La tabella seguente illustra i risultati prima e dopo la prova. Assumendo che le necessarie condizioni sulla popolazione siano soddisfatte, eseguire un test per verificare l’ipotesi che i punteggi medi prima e dopo il corso non cambino. Usare un livello di significatività di 0.05

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Lavoratore** | | | | |
| **Tipo di promozione** | **1** | **2** | **3** | **4** | **5** |
| Prima del corso | 45 | 36 | 52 | 58 | 63 |
| Dopo il corso | 68 | 59 | 77 | 85 | 75 |

16. Kellen Auto Products ha testato su strada il nuovo additivo per benzina destinato a ridurre i consumi nelle auto per passeggeri. Nelle prove sono state usate cinque Chevrolet Orion e cinque diversi guidatori. Ciascun guidatore, con il mezzo assegnatogli, ha effettuato 500 miglia di guida ordinaria giornaliera utilizzando l’additivo e 500 miglia senza usare l’additivo. Per ogni guidatore, se usare prima la benzina con additivo o quella senza additivo è stato stabilito casualmente. I risultati sono di seguito riportati.

Risultati de test in miglia per gallone (mpg)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Auto** | | | | |
| **Tipo di benzina** | **1** | **2** | **3** | **4** | **5** |
| Con additivo | 35.7 | 31.3 | 36.1 | 33.0 | 38.9 |
| Senza additivo | 32.2 | 29.3 | 34.5 | 32.8 | 36.2 |

Assumendo che le necessarie condizioni sulla popolazione siano soddisfatte, eseguire un test per verificare l’ipotesi che la differenza tra i consumi medi con i due tipi di benzina sia pari a zero. Usare un livello di significatività di 0.05

17. L’Università di Alta Vista sta cercando di valutare l’efficacia del nuovo corso di scrittura per matricole. Ciascuno studente frequentante è stato sottoposto ad una prova di scrittura pre-corso e ad una prova post-corso, entrambe valutate da un professore di inglese che ignorava quale delle prove fosse pre-corso e quale fosse post-corso. I risultati sono riportati sotto.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Studente** | | | | | |
|  | **1** | **2** | **3** | **4** | **5** | **6** |
| Pre-corso | 78 | 85 | 63 | 91 | 83 | 68 |
| Post-corso | 73 | 87 | 56 | 89 | 73 | 66 |

Assumendo che le necessarie condizioni sulla popolazione siano soddisfatte, eseguire un test per verificare l’ipotesi che la differenza tra i punteggi medi prima e dopo il corso sia pari a zero. Usare un livello di significatività di 0.10.