



Su 5459 donne in gravidanza seguite dall'Aarhus University Hospital in Danimarca, sono stati raccolti dati, in un periodo di due anni, sulla durata della gestazione. E' stata osservata una media pari a 281.9 giorni, e una deviazione standard pari a 11.4 giorni. Un bambino viene classificato prematuro se il tempo gestazione è pari 258 giorni o inferiore. Se i tempi gestazionali sono distribuiti normalmente, quale proporzione di bambini dovrebbe nascere prematuramente?



Riporta lo z -score utilizzato nella ricerca di un intervallo di confidenza per una proporzione con livello di fiducia 0.9973



Il *Mental Development Index* (MDI) della Scala di Bailey dello sviluppo infantile, è una misura standardizzata utilizzata in studi sui bambini ad alto rischio. Questa variabile ha una distribuzione approssimativamente normale con media pari a 100 e una deviazione standard pari a 16. Trova il punteggio di MDI pari al secondo quartile.



Una variabile psicologica misurata nella popolazione ha una distribuzione *gaussiana*, con $\mu = 0.10$ e $\sigma = 316.23$. Supponi di ripetere la misurazione su un campione casuale indipendente di 1000 soggetti, e indica con \bar{y} la media dei valori rilevati. Trova la media e l'errore standard della distribuzione campionaria di \bar{y} .



Uno studio ha riportato che 486 donne avevano una media di 8.3 amici intimi ($s = 15.6$) e 354 uomini avevano una media di 8.9 amici intimi ($s = 15.5$). Calcola un intervallo di confidenza al 95% per la differenza tra le medie di popolazione dei maschi e delle femmine, assumendo l'omogeneità della varianza nelle due popolazioni di riferimento.



Si consideri il seguente esempio basato sui dati contenuti nel *data frame* **Ginzberg**.

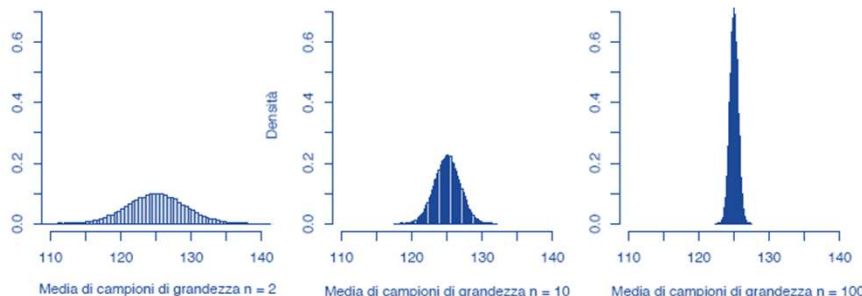
- La variabile dipendente (**adjdep**) è il punteggio sulla *Beck self-report depression scale*.
- I soggetti sono stati divisi in due gruppi a secondo del punteggio ottenuto su una scala (**adjsimp**) che misura il *bisogno di vedere il mondo "in bianco e nero"*
- I soggetti con punteggi minori o uguali alla media del campione su questa variabile sono stati assegnati al gruppo "*semplicità = bassa*"; gli altri sono stati assegnati al gruppo "*semplicità = alta*".
- Per ciascun gruppo sono state calcolate la media e la deviazione standard dei punteggi sulla *Beck self-report depression scale*.
- I risultati sono i seguenti:

Depressione	n	\bar{y}	s
semplicità bassa	46	0.8308	0.4433
semplicità alta	36	1.2162	0.4901

Calcolare la statistica test t-Student per campioni indipendenti, assumendo l'omogeneità della varianza dei due gruppi.



Osservando queste distribuzioni di medie campionarie, si evince che:



- la media campionaria è uno stimatore corretto (unbiased). la media campionaria è uno stimatore preciso. la media campionaria è uno stimatore consistente. la media campionaria è uno stimatore potente.



L'interpretazione dell'intervallo di fiducia al 95%:

- La probabilità che il parametro stia nell'intervallo è 0.95. La probabilità che il parametro stia nell'intervallo è $1 - 0.95 = 0.05$. Su un numero molto grande di repliche indipendenti della stima intervallare, ci si attende che il 5% delle stime non conterranno il parametro. Nell'intervallo troviamo il 95% dei dati della popolazione di riferimento.
-



Un test consiste di 100 domande VERO - FALSO. Joe non ha studiato, così, in ciascuna domanda, tenta di indovinare la risposta corretta. Si trovi la probabilità che Joe ottenga un punteggio di almeno 70 punti, passando quindi l'esame.



Per una variabile distribuita normalmente, determina la probabilità che un'osservazione ricada tra $\mu - 0.85\sigma$ e $\mu + 0.35\sigma$.



L'interpretazione dell'intervallo di fiducia al 95%:

- La probabilità che il parametro stia nell'intervallo è 0.95. La probabilità che il parametro stia nell'intervallo è $1 - 0.95 = 0.05$. Su un numero molto grande di repliche indipendenti della stima intervallare, ci si attende che il 5% delle stime non conterranno il parametro. Nell'intervallo troviamo il 95% dei dati della popolazione di riferimento.
-



Il Teorema del Limite Centrale implica che (scegli la risposta esatta):

- la distribuzione campionaria assomiglia sempre più alla distribuzione della popolazione al crescere dell'ampiezza campionaria. Le distribuzioni della popolazione sono normali se la popolazione è piuttosto numerosa. Tutte le variabili hanno distribuzioni dei dati campionari a forma campanulare se un campione casuale è costituito da almeno 30 osservazioni. Per grandi campioni casuali, la distribuzione campionaria di \bar{y} è approssimativamente normale, indipendentemente dalla forma della distribuzione della popolazione.



In risposta alla domanda sul numero di ore giornaliere trascorse davanti alla TV, le risposte fornite da otto soggetti che si sono identificati come Hindu sono state 2,3,2,1,0,1,4,3. Determina una stima puntuale della media della popolazione per gli Hindu e il margine di errore per questa stima.



Nel 2006, il Florida Poll realizzato per la Florida International University ha domandato se le attuali normative ambientali sono troppo restrittive o meno. Dei 1200 rispondenti, 229 hanno detto che lo sono. Trova un intervallo di confidenza al 95% e al 99% per il valore del parametro.



1. Un operaio deve entrare in una cisterna. Si ha probabilità di .54 di avere cefalea, se si respira aria satura di CO₂, e di 0.07, se si respira aria pura. Le cisterne sono 3 con aria satura di CO₂ e 7 con aria pura. Quale sarà allora la probabilità di essere entrato in una cisterna con aria pura, se l'operaio ha cefalea?



3. Si abbia il seguente campione: 2, 3, 3, 6, 4, 2, 4, 5. Calcola gli intervalli di fiducia della varianza dell'universo di appartenenza per $\alpha = 0.10$.



La media di un campione è 144 ± 32 . Calcola gli intervalli di confidenza della varianza, con $\alpha = 0,025$.

Quant'è la somma di 10 punti z di una distribuzione elevati al quadrato?

n

χ_{10}^2

χ_9^2



Per una variabile aleatoria *normale*, quale proporzione di casi è racchiusa tra $\mu - 0.67\sigma$ e $\mu + 0.67\sigma$?



Per una variabile aleatoria *normale*, quale proporzione di casi è racchiusa tra $\mu - 0.49\sigma$ e $\mu + 0.49\sigma$?



Trovare il valore z corrispondente al 43esimo percentile.



4.14 Se z è un numero positivo tale che l'intervallo tra $\mu - z\sigma$ e $\mu + z\sigma$ contiene la metà dell'area di una distribuzione normale, allora quale è il percentile di $\mu + z\sigma$?

25 esimo

35 esimo

65 esimo

75 esimo



5.33 Uno studio ha stimato il reddito medio annuale per famiglie che vivono in alloggi pubblici a Chicago. Per un campione casuale di 30 famiglie, i redditi medi sono

83	90	77	100	83	64	78	92	73	122
96	60	85	86	108	70	139	56	94	84
111	93	120	70	92	100	124	59	112	79

Costruisci un intervallo di confidenza al 95% per μ .



6.53 Il p-valore per verificare $H_0: \mu = 100$ contro $H_a: \mu \neq 100$ è $p = 0.001$. C'è una forte evidenza che

$\mu = 100$

$\mu \neq 100$

$\mu > 100$

$\mu < 100$



7.35 Due nuovi corsi brevi son stati proposti per aiutare gli studenti che soffrivano di fobia anti-matematica, ovvero con punteggio pari o superiore a 8 su una scala con estremi 0 -10. I 10 studenti del campione sono stati casualmente assegnati a due corsi. Alla fine dei corsi, è stata registrata la diminuzione nel punteggio della fobia anti-matematica. I valori ottenuti sono: Corso A {0, 2, 2, 3, 3}, Corso B {3, 6, 6, 7, 8}. Calcola la statistica t e stima il p-valore, per un test unilaterale.



Sia y = numero delle volte che hai letto il giornale in una settimana. Assegnando alle categorie di risposta i valori 7, 3, 1, 0.5, 0 e considerando i seguenti conteggi risultanti per ciascuna modalità, 358, 222, 134, 121, 71, trova y medio.



Riporta il valore del t-Student che moltiplicato per l'errore standard serve a determinare un intervallo di confidenza al 95% con 25 osservazioni



Secondo un accordo sindacale, il reddito medio per tutti i lavoratori di profili senior della linea di assemblaggio di una grande azienda deve essere pari a 500 euro alla settimana. Una rappresentante di un gruppo di donne chiede di analizzare se il reddito medio μ delle lavoratrici è conforme a questo accordo. Per un campione casuale di nove donne occupate sono stati ottenuti i valori, $\bar{y}=410\$$ e $s=90\$$. Riporta il p-valore per $H_a: \mu < 500\$$.



Successivamente al trattamento clinico per la depressione mentale, il miglioramento viene misurato dalla variazione nel punteggio su una scala standardizzata. I punteggi sono 10, 20, 30 per i pazienti che ricevono la terapia A e 30, 45, 45 per quelli che ricevono la terapia B. Assumendo una comune deviazione standard per i due gruppi di terapia, calcola l'errore standard della differenza tra le medie.