

Questi sono tutti gli esercizi di statistica che possono essere inclusi nelle prove di esame (collettive o individuali, a seconda del programma del corso e del tipo di esame). Oltre a questi, le prove scritte possono comprendere quiz scelta multipla.

Data la numerosità di un campione, calcolare:

=> numero ideale di classi dell'istogramma

Data una serie di valori, calcolare:

=> media

=> errore standard

=> mediana

=> devianza

=> varianza

=> deviazione standard

=> valori standardizzati

=> 2.5° e 97.5° percentile

=> test di simmetria

=> somma dei ranghi

=> rango medio

Dati una media e deviazione standard, calcolare:

=> numerosità del campione tale da avere limiti fiduciali in un certo rapporto con la media e con un livello di probabilità del 95%

Date due serie di valori (campioni indipendenti), calcolare:

=> test t di Student per campioni indipendenti

=> test di Wilcoxon-Mann Whitney

Date due serie di valori (campioni appaiati), calcolare:

=> test t di Student per campioni appaiati

=> test dei segni di Wilcoxon

Dati più gruppi di valori, calcolare:

=> devianza tra gruppi

=> gradi di libertà tra gruppi

=> varianza tra gruppi

=> devianza entro gruppi

=> gradi di libertà entro gruppi

=> varianza entro gruppi

=> test F per l'analisi della varianza

=> test di Kruskal-Wallis

Date diverse serie di valori di risposta di soggetti sottoposti a diversi trattamenti, calcolare:

=> test di Friedman

Dati due serie di valori, calcolare:

=> codevarianza

=> intercetta

=> errore standard dell'intercetta

=> test t per la significatività dell'intercetta

- => pendenza
- => errore standard della pendenza
- => test t per la significatività della pendenza
- => previsione di un valore y per un dato valore x
- => errore standard della previsione
- => coefficiente di correlazione (parametrico)
- => coefficiente di correlazione di Kendall
- => coefficiente di correlazione di Spearman
- => varianza di y dovuta alla regressione
- => varianza di y dovuta all'errore
- => varianza totale di y
- => test F per la significatività della regressione

Data una tabella 2´2, calcolare:

- => chi-quadro
- => chi-quadro corretto per la continuità
- => tabella delle frequenze attese

Data una tabella m´n, calcolare:

- => gradi di libertà
- => chi-quadro
- => tabella delle frequenze attese

Data una tabella 2´2 per dati appaiati, calcolare:

- => test di McNemar

Data una proporzione (o percentuale), calcolare:

- => errore standard della proporzione (o percentuale)

Date due proporzioni (o percentuali), calcolare:

- => tabella 2x2

Data la probabilità a priori di un singolo evento, calcolare:

- => probabilità di un certo numero di eventi in un numero complessivo di prove (distribuzione binomiale)

Data la prevalenza di una malattia, sensibilità e specificità di un test, calcolare:

Calcolare:

- => probabilità che una persona positiva al test sia effettivamente malata (Bayes)

Data la frequenza media di un evento riferita ad un certo intervallo di spazio/tempo, calcolare:

- => probabilità di frequenza dell'evento in un piccolo intervallo di spazio/tempo (Poisson)

Data la lista dei dati di due o tre variabili di due soggetti, calcolare:

- => distanza euclidea tra i due soggetti (senza standardizzare i dati)

Data la media e deviazione standard di tre o più variabili ed i dati di un soggetto, disegnare:

- => sun-ray plot del soggetto (lunghezza del raggio = ± 3 deviazioni standard)

Data una matrice delle distanze di 4 soggetti, disegnare:

- => dendrogramma con il criterio della distanza minima ed il metodo della media di gruppo non pesata

Nota che

- oltre l'errore standard occorre sempre saper calcolare i limiti fiduciali al 95% di confidenza
- ovviamente i parametri di devianza, varianza, deviazione standard ed errore standard si possono sempre ricavare l'uno dall'altro

Questi sono alcuni quesiti proposti nelle precedenti verifiche di Statistica. E' probabile che anche le future verifiche comprendano esercizi analoghi, ma anche diversi. Le verifiche comprendono generalmente 10 quesiti.

Calcolare il t di Student per i due gruppi di dati non-appaiati:

gruppo A: 1, 3, 5, 8, 2, 6, 7

gruppo B: 10, 8, 7, 9, 9

Calcolare la pendenza della retta di regressione dei seguenti dati:

X	Y
2	4
3	6
5	8
6	10

Calcolare il chi-quadro con la correzione per la continuità per la tabella 2×2:

2	9
4	6

Se in un litro di terreno di coltura vi sono 10,000 cellule, qual'è la probabilità (secondo la distribuzione di Poisson) di trovare 6 cellule in un campione di un millilitro di quel terreno?

Qual'è il rango medio di un campione di 423 dati?

Calcolare la deviazione standard del seguente campione: 6, 4, 2, 5, 8.

Quale è la probabilità che i sei figli di una coppia siano tutti e sei maschi?

Supponiamo che 94 pazienti su 100 rispondano ad una certa terapia. Qual'è l'errore standard di questa proporzione?

Un campione di 64 dati ha varianza = 256. Calcolare l'errore standard.

In un campione di 30 adulti si trova che il 40% di essi porta gli occhiali. Calcolare la deviazione standard di questa percentuale.

Il numero di globuli rossi (qui espresso in milioni/mm³ di sangue) in 5 soggetti prima e dopo una certa terapia è il seguente:

Prima	Dopo
3.5	4.5
4.0	4.0
3.0	3.5
3.5	4.5
4.5	5.0

Calcolare il t di Student per campioni appaiati per verificare se la terapia ha fatto aumentare il numero di globuli rossi.

Calcolare il coefficiente di correlazione tra le due serie di dati:

X	Y
2	4
3	5
5	6
7	7
8	8

Un soggetto su un milione è affetto da una certa malattia (prevalenza = 0.000001). I malati di tale malattia mostrano un certo sintomo X nel 99% dei casi. Lo stesso sintomo X è comunque presente anche nell'4% di soggetti sani. Calcolare la probabilità che un individuo che manifesti il sintomo X sia effettivamente malato.

Un certo campione ha media=50, deviazione standard=48 e numerosità $n=144$. Calcolare i limiti fiduciali della media al livello di probabilità del 95% (ponendo $t=1.96$).

Se il risultato di un test statistico è significativo significa che:

- (a) si accetta l'ipotesi zero (o ipotesi nulla)
- (b) il valore del cosiddetto 'P' è maggiore di 0.05
- (c) la probabilità che il risultato sia un falso positivo è inferiore al 5%
- (d) le probabilità a favore dell'ipotesi alternativa sono superiori al 50%
- (e) l'errore di II tipo è inferiore al 5%

Se si somma una costante K a tutti i dati di un campione:

- (a) la media risultata aumentata di K mentre la deviazione standard non varia
- (b) sia la media che la deviazione standard risultano aumentate di K
- (c) la media non varia mentre la deviazione standard risultata aumentata di K
- (d) né la media né la deviazione standard variano

In una certa clinica, su 50 malati, 40 sono guariti. In un'altra clinica, su 30 malati con le stesse caratteristiche, solo 5 sono guariti. Calcolare il chi-quadro per verificare se vi è una differenza significativa nella percentuale di guarigioni tra le due cliniche.

Calcolare la mediana del seguente campione di dati: 6, 8, 12, 23, 34, 35, 35, 38, 46, 48, 54, 57, 60, 75, 87.

Media e mediana sono coincidenti quando

- (a) la varianza è maggiore di zero
- (b) l'errore standard è 1
- (c) la distribuzione è perfettamente normale
- (d) la kurtosis standardizzata è zero

I limiti fiduciali dipendono dalla

- (a) devianza
- (b) varianza
- (c) deviazione standard
- (d) errore standard

Gruppo A: media = 25, devianza = 40, $n = 10$

Gruppo B: media = 30, devianza = 35, $n = 16$

Calcolare il t di Student per campioni non-appaiati

Calcolare il chi-quadro corretto per la continuità della tabella 2x2

3	20
10	4

Calcolare la codevianza dei seguenti dati:

X	Y
1	4
2	5
3	8
4	9

Applicare il teorema di Bayes per conoscere la probabilità che un paziente positivo ad un certo test sia malato, sapendo che il test è positivo sul 85% dei malati e sull'5% dei sani e che la malattia in questione è presente nello 0.01% della popolazione.

Su un gruppo di 50 pazienti si controlla l'effetto di due farmaci - tra loro compatibili e sovrapponibili -. Ogni paziente riceve quindi un ciclo terapeutico del farmaco A seguito da un ciclo terapeutico del farmaco B con i seguenti risultati:

10 pazienti migliorano con A e con B

5 pazienti migliorano con A

30 pazienti migliorano con B

5 pazienti non hanno nessun miglioramento

Calcolare il chi-quadro secondo il test di McNemar per stabilire se esiste un effetto migliore di uno dei due farmaci.

Se nell'analisi della regressione si trova che la varianza "dovuta alla regressione" è uguale alla varianza "di errore", si può dire che:

- (a) Y è indipendente dalla variabile X
- (b) Y è dipendente dalla variabile X
- (c) Y è distribuito in modo normale
- (d) la varianza di Y è maggiore della varianza di X

Calcolare l'errore standard del seguente campione di dati: 3, 7, 4, 5, 1.

Supponiamo che nel sangue circolante vi siano 10 cellule 'K' per ml. Se queste cellule fossero distribuite come previsto dalla distribuzione di Poisson, quale sarebbe la probabilità di trovare senza cellule 'K' una frazione di sangue di 10 microlitri?

Calcolare il valore del 97.5° percentile della distribuzione avente media=75 e deviazione standard=10.

Quale è l'errore standard della media ottenuta da un campione di 10 dati con varianza = 900.

Calcolare il coefficiente di correlazione di Kendall (τ) dei seguenti dati:

X	Y
4	6
3	8
8	10
6	9
10	12
15	7
7	5

In una sessione di esami sono promossi 10 studenti su 60. Calcolare i limiti fiduciali della relativa proporzione.

Dato il campione: 9, 3, 3, 6, 8, 6, quale è il valore standardizzato di 9?

Quale è la somma dei ranghi di un campione di 500 dati?

Siano i due campioni (indipendenti):

A: media = 56, devianza = 200, $n = 10$

B: media = 48, devianza = 100, $n = 8$

Calcolare il t di Student per saggiare l'ipotesi che i due campioni provengano dalla stessa popolazione.

In giorni diversi, si valuta il comportamento di alcuni soggetti in diverse condizioni ambientali. Quale test più appropriato per verificare se vi sono differenze significative del comportamento dovute alle diverse condizioni ambientali?

- (a) test F mediante analisi della varianza
- (b) test non-parametrico di Mann-Whitney
- (c) test non-parametrico di Friedman
- (d) test non parametrico di Kruskal-Wallis

Calcolare l'intercetta della retta di regressione dei due campioni:

X=	7	14	21	28
Y=	12	14	16	19

Avendo due gruppi di dati e volendo fare un confronto tra le due medie è meglio usare:

- (a) l'analisi della varianza
- (b) il test t di Student
- (c) il test di Student-Neuman-Keuls
- (d) il test di Friedman

In una pizzeria con 15 tavoli sono presenti 15 persone. Assumendo che la distribuzione dei mangiatori di pizza sia di tipo Poissoniano (e che non vi siano tavoli liberi-prenotati) calcolare la probabilità che nella pizzeria vi sia un tavolo vuoto.

Quale è la probabilità che una coppia abbia figli 0 maschi su un totale di 6 figli?

La mediana è un parametro preferibile alla media quando la distribuzione mostra

- (a) kurtosis elevata
- (b) skewness elevata
- (c) deviazione standard elevata
- (d) errore standard elevato

Un certo campione di 100 dati ha media=10 e devianza=9900. Ricavare la deviazione standard

Confrontando tra loro diversi campioni, a varianza maggiore corrisponde sempre

- (a) devianza maggiore
- (b) deviazione standard maggiore
- (c) errore standard maggiore
- (d) skewness maggiore

Un piccolo campione di 100 dati risulta avere media=50 e deviazione standard=400. Calcolare quanto dovrebbe essere grande il campione perché i limiti fiduciali non si discostino più 10% del valore della stessa media, con un margine di confidenza del 95% (ponendo $t=1.96$)

Calcolare il chi-quadro della seguente tabella 3x3

4	3	16
5	6	10
7	9	2

Calcolare la distanza euclidea tra i due soggetti definiti dalle variabili v1, v2, e v3

	v1	v2	v3
soggetto A	2	4	2
soggetto B	1	2	4

Siano i due campioni

A: 2, 5, 9, 3, 7, 6, 1, 23, 5

B: 12, 8, 13, 25, 45

Calcolare il rango medio dei dati del campione B

La deviazione standard di un campione è 72. L'errore standard è 12. Qual'è la numerosità del campione?

Due campioni da confrontare hanno i seguenti valori:

campione A media = 10 devianza = 20 n = 60

campione B media = 18 devianza = 18 n = 50

Calcolare il t di Student.

Di 150 pazienti trattati con una certa terapia (A) 110 giungono a guarigione. Di altri 180 pazienti trattati con una certa diversa terapia (B) 125 giungono a guarigione. Calcolare il chi-quadro (corretto per la continuità) per verificare se le frequenze di guarigione ottenute con le due diverse terapie siano significativamente diverse.

Supponiamo che un certo allele nella popolazione sia presente in 1 soggetto su 10 ($p=0.1$). Qual'è la probabilità di trovare questo stesso allele in 3 persone su un totale di 5 persone esaminate (e non imparentate tra loro)?

Consideriamo il seguente campione:

5, 7, 7, 4, 5, 6, 3, 4, 6, 9, 10, 11, 15, 12, 14, 19, 12, 10, 20.

Qual'è il rango del valore 9?

La probabilità di falsi negativi è pari:

- (a) 1 – probabilità di falsi positivi
- (b) 1 - probabilità di veri negativi
- (c) probabilità di veri negativi - 1
- (d) 1 - probabilità di veri positivi

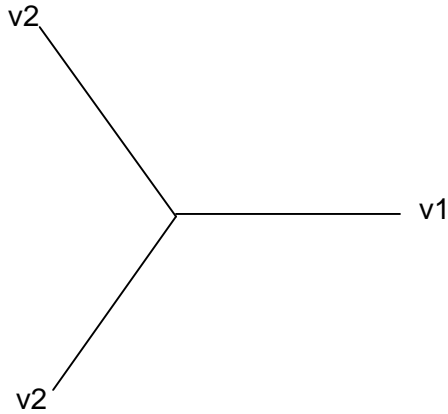
Una certa malattia colpisce un soggetto su mille. Supponiamo che un certo test sia positivo sul 99% dei malati ma anche sul 5% dei sani. Qual'è la probabilità che un soggetto positivo al test sia effettivamente malato?

Calcolare i limiti fiduciali al 95% di probabilità della media del seguente campione di dati: 6, 7, 9, 11, 10, 8, 5 ($t = 1.96$)

Tre variabili (v_1 , v_2 e v_3) hanno i seguenti valori di media e deviazione standard:

v_1 :	media = 22	dev. standard = 6
v_2 :	media = 20	dev. standard = 10
v_3 :	media = 50	dev. standard = 5

Costruire il sun-ray plot di un soggetto con i valori: $v_1 = 10$; $v_2 = 15$; $v_3 = 55$ attribuendo agli assi un range di ± 2 deviazioni standard



Calcolare la probabilità teorica che, lanciando il dado 10 volte, il due esca 2 volte.

I granulociti acidofili sono circa il 5% dei leucociti circolanti. Nello striscio di sangue di un paziente il numero dei granulociti acidofili è di 25 su un totale di 150 leucociti esaminati. Calcolare il chi-quadro (non corretto per la continuità) per saggiare l'ipotesi che tale valore rappresenti un significativo incremento dei granulociti acidofili di quel paziente.