

[illegible]

Esercizi di statistica descrittiva e di verifica delle ipotesi risolti e commentati con R

[Home](#) [Indice degli esercizi](#) [Links utili e libri](#) [Proponi un articolo](#)

 [Subscribe to Statistica con R via RSS](#)

Or subscribe via e-mail:

Subscribe

Like 160 people like this.

## Recent Posts

Boxplot con media e deviazione standard

## Cochran Q Test in R (Nonparametric Test for Categorical Data)

---

Test di Friedman: confronti multipli non parametrici

## CDC Growth Charts implementate in R

R è anche un ottimo sound editor!

## My Blog List

 **Resources for Medical Students**

Chemical Problems and Solutions

 The Smoking Club

Quale software statistico usi?

R	42 (51%)
Excel	9 (11%)
S.PLUS	0 (0%)
SPSS	19 (23%)
Stata	2 (2%)
SYSTAT	0 (0%)
Mathematica	0 (0%)
Minitab	0 (0%)
Altro	9 (11%)

Voti fino a ora: 81  
Sondaggio chiuso

## Google Ads

### Tag Cloud

Algebra Lineare   Algoritmo  
di Strassen   Analisi del  
trend   Analisi della  
sopravvivenza   Analisi  
delle varianze



## Post più letti

## Indice degli esercizi svolti

## Regressione lineare

## Regressione lineare multivariata

### I test di verifica di normalità

## Analisi della sopravvivenza: curve di Kaplan-Meier e modello di Cox

## Test di analisi di omogeneità della varianza

## Test di Wilcoxon dei ranghi con segno

### Confronto delle medie tra 2 gruppi di campioni appaiati, metodo non parametrico.

Il sindaco di una città vuole verificare se i livelli di inquinamento si riducono chiudendo le strade del centro al traffico. A tal fine viene misurato il tasso di inquinamento ogni 60 minuti (dalle 8 alle 22: totale di 15 rilevazioni) in una giornata in cui il traffico è aperto, e in una giornata di chiusura al traffico. Ecco di seguito i valori di inquinamento atmosferico:

Con traffico: 214, 159, 169, 202, 103, 119, 200, 109, 132, 142, 194, 104, 219, 119, 234

Senza traffico: 159. 135. 141. 101. 102. 168. 62. 167. 174. 159. 66. 118. 181. 171. 112

Siamo di fronte a dati appaiati, perché esiste un vincolo tra le rilevazioni, consistente nel fatto che stiamo considerando la stessa città (con le sue peculiarità atmosferiche, ventilazione, etc.) seppure in due differenti giorni. Non potendo supporre una distribuzione gaussiana per i valori rilevati, dobbiamo procedere con un test non parametrico, il **test di Wilcoxon dei ranghi con segno**.

```
> a = c(214, 159, 169, 202, 103, 119, 200, 109, 132, 142, 194, 104, 219,
119, 234)
> b = c(159, 135, 141, 101, 102, 168, 62, 167, 174, 159, 66, 118, 181, 171,
112)
>
> wilcox.test(a,b, paired=TRUE)
```

Wilcoxon signed rank test

data: a and b

$V = 80$ ,  $p\text{-value} = 0.2769$

alternative hypothesis: true location shift is not equal to 0

Essendo il p-value maggiore di 0.05, significa che le medie delle rilevazioni sono sostanzialmente rimaste inalterate (accettiamo l'ipotesi  $H_0$ ), quindi bloccare il traffico per un solo giorno non ha portato a nessun miglioramento per quanto riguarda l'inquinamento della città.

Il valore  $V = 80$  corrisponde alla somma dei ranghi assegnati alle differenze con segno positivo. Con qualche acrobatismo matematico, possiamo calcolare manualmente la somma dei ranghi assegnati alle differenze con segno positivo, e la somma dei ranghi assegnati alle differenze con segno negativo, per confrontare questo intervallo con l'intervallo tabulato sulle tavole di Wilcoxon per campioni appaiati, e confermare la nostra decisione statistica. Ecco come fare per calcolare le due somme.

```
> a = c(214, 159, 169, 202, 103, 119, 200, 109, 132, 142, 194, 104, 219,
119, 234)
> b = c(159, 135, 141, 101, 102, 168, 62, 167, 174, 159, 66, 118, 181, 171,
112)
>
> diff = c(a - b) #calcolo il vettore contenente le differenze tra tutti i
valori di a e i valori di b
> diff.rank = rank(abs(diff)) #assegno i ranghi alle differenze calcolate
prima, prese in valore assoluto; la funzione abs(x) dà in output il valore
di x senza segno
> diff.rank.sign = diff.rank * sign(diff) #do il segno ai ranghi,
richiamando i segni dei valori delle differenze
> ranghi.pos = sum(diff.rank.sign[diff.rank.sign > 0]) #calcolo la somma dei
ranghi assegnati alle differenze con segno positivo, ossia maggiori di zero
> ranghi.neg = -sum(diff.rank.sign[diff.rank.sign < 0]) #calcolo la somma
dei ranghi assegnati alle differenze con segno negativo, ossia minori di
zero, cambiando di segno il risultato, che sarebbe logicamente negativo
> ranghi.pos #chiamo il risultato: esso corrisponde al valore V calcolato
con il wilcox.test
[1] 80
> ranghi.neg #chiamo il risultato
[1] 40
```

L'intervallo calcolato è (40, 80). L'intervallo tabulato sulle tavole di Wilcoxon per campioni appaiati, con 15 differenze è (25, 95). Poiché l'intervallo calcolato è contenuto nell'intervallo tabulato, accettiamo l'ipotesi H0 di uguaglianza delle medie. Come preannunciato dal p-value, la chiusura delle strade al traffico non ha apportato alcun miglioramento in termini di tasso d'inquinamento.

NOTA.

Il test di Wilcoxon dei ranghi con segno prevede la NON assegnazione del rango a quelle differenze pari a zero. Pertanto con il metodo illustrato, si avrebbero risultati falsati, in quanto anche a differenze pari a zero, viene assegnato un rango. Si consiglia pertanto di verificare sempre i contenuti dei vari vettori che vengono calcolati. Se qualcuno ha tempo, può provare a costruire una funzione con le sequenza di calcoli da me utilizzata, aggiungendo la condizione di eliminare dall'assegnazione dei ranghi quelle differenze pari a zero. Io per ora non ho tempo, ma magari in futuro scriverò una funzione completamente automatizzata.




EDIT.

La risposta al dubbio espresso nella nota è molto più semplice di quanto avessi previsto! Per eliminare le differenze pari a zero dal calcolo, è sufficiente aggiungere la seguente istruzione:

```
diff <- diff[ diff != 0 ]
```

e il gioco è fatto!

Etichette: [Campioni appaiati](#), [Due campioni](#), [Test di Wilcoxon](#)

CONDIVIDI    

Articoli Correlati:

- Repeated measures ANOVA: confronto tra gruppi di misure ripetute
- Test di Wilcoxon dei ranghi con segno
- Test t di verifica della media di due campioni appaiati
- Test di Wilcoxon-Mann-Whitney della somma dei ranghi
- Test t a due campioni #2
- Test t a due campioni #1
- Cochran Q Test in R (Nonparametric Test for Categorical Data)
- Test di Friedman: confronti multipli non parametrici

ANOVA ANOVA a due vie Bernoulli CA Campioni appaiati Campioni indipendenti CDC Choropleth map Cluster Analysis Cochran Q test Cochran-Mantel-Haenszel Confronti multipli Correlazione Curve di Kaplan-Meier Dataframe Diagramma a torta Diamond Chart Distribuzione binomiale Distribuzione F-max Due campioni Editor Excel For..To..Do Friedman test Funzioni di R ggplot2 Google Analytics Google Chart API Grafici Growth charts Hierarchical Clustering IEEE-754 If..Then..Else Immagini InkBlot graph Latin Squares Design Log-rank test Matrici McNemar test Modello di Cox Odd-ratio Omoschedasticità OR Ortogonale p-value Partitional clustering Più campioni Post-hoc tests Probabilità Quadrati latini Regressione Regressione lineare multipla Regressione logistica Regressione polinomiale Regressione segmentale Repeated measures ANOVA Residui Stagionalità Standardizzazione Suoni T-test Tabelle di contingenza Test del chi-quadro Test di Armitage Test di Bartlett Test di Cox-Stuart Test di Fligner-Killeen Test di Kendall Test di Kruskal-Wallis Test di Levene Test di Pearson Test di Spearman Test di verifica di normalità Test di Wilcoxon Test F di Fisher Trend analysis Un campione Verifica di proporzioni Welch test Z-test

ANOVA a due vie

Repeated measures ANOVA: confronto tra gruppi di misure ripetute

Quando le assunzioni dell'ANOVA sono violate: tests e post-hoc tests

Cluster Analysis in R #1: Hierarchical Clustering

Contact me at:

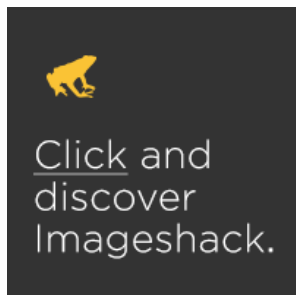
T. Martino: todoslogos@Gmail.com

Blog Archive

febbraio 2014 (1)  
ottobre 2011 (2)  
settembre 2011 (2)  
agosto 2011 (2)  
giugno 2011 (1)  
novembre 2010 (1)  
ottobre 2010 (5)  
settembre 2010 (1)  
luglio 2010 (1)  
giugno 2010 (1)  
aprile 2010 (1)  
febbraio 2010 (2)  
gennaio 2010 (1)  
dicembre 2009 (3)  
settembre 2009 (1)  
agosto 2009 (10)  
luglio 2009 (15)  
novembre 2008 (16)

Blog's reader info

96 readers BY FEEDBURNER

[1 Commento:](#)[martina ha detto...](#)

Grazie mille, mi dava un errore e ora seguirò il tuo metodo!

22 giugno 2012 16:06

**Posta un commento**

Inserisci il tuo commento...

Commenta come: Account Googl ▼

Pubblica

Anteprima

[<< Post più recente](#)[Home page](#)[Post più vecchio >>](#)**Links utili**[Homepage di R](#)[Download R](#)[R Data Analysis Examples at the UCLA university](#)[Multivariate analysis at the OxfordJournals](#)**Statistiche... del blog!**

In questo blog ci sono 66 posts e 292 commenti.

**Visualizzazioni totali (dal 01.06.2010)****733,858****Follow me on...**[Facebook](#)