Analisi dati in R

Prova di Accertamento dell'apprendimento

Dott. Marco Vicentini

22 aprile 2009

Il candidato esegua i seguenti esercizi, scrivendo la sintassi adeguata in uno script in formato R:

1) Creare un vettore contenente: i numeri primi da 1 a 10, i soli numeri dispari da 1 a 10, la ripetizione del valore 10 per 10 volte.
2) Calcolare utilizzando modalità differenti la somma dei primi 100 numeri interi: si utilizzino almeno 2 delle seguenti funzioni: l'operatore aritmetico di addizione (+), la funzione sum, un ciclo for.
3) Creare un vettore contenente gli interi da 1 a 5, un vettore contenente gli interi da 5 ad 1, ed eseguire la somma e il prodotto dei due vettori
4) Assegnare all'oggetto "punti" un vettore contenente numeri casuali compresi tra 0 e 100, di lunghezza 20. È possibile utilizzare le funzioni di generazione di numeri casuali, quale, ad esempio, runif.
5) Assegnare all'oggetto "sogg" un vettore contenente, anche in ordine casuale, 10 volte il valore "M" e 10 volte il valore "F". Assegnare il formato factor all'oggetto sogg.
6) Creare un dataframe "dati" contenente gli oggetti punti e sogg.
7) Elaborare un opportuno grafico (ad esempio con boxplot) relativo alla distribuzione di punti in funzione dei livelli del fattore sogg. Il grafico puo' essere migliorato graficamente a proprio piacere.
8) Estrarre le statistiche fondamentali (minimo, massimo, distribuzione in quartili) relativamente ai soggetti "M" e "F"
9) Effettuare almeno test statistico sui dati per verificare la significatività della differenza osservata tra "M" e "F" e commentare nello script i risultati ottenuti. È possibile utilizzare la funzione t.test, aov, TukeyHSD, model.frames, senza con questo escludere altre funzioni statistiche conosciute.
10) Caricare il file di dati "dataset4.RData" (scaricabile da http://www.marcovicentini.it/index.php?page=41&menu=4) e ripetere le analisi richieste nei punti 6, 7, e 8 per il dataset Simon, verificando la significatività dei fattori "se" e "re" sulla variabile dipendente "TR"

Al termine della prova di accertamento, inviare il contenuto dello script tramite mail all'indirizzo marco.vicentini@gmail.com. Si raccomanda di accertarsi di aver allegato lo script prodotto.