TDE 140917

1. a) prima di effettuare un test per misure ripetute per gaussiane, verifico che siano congiuntamente normali. Mschapiro dà un pvalue >0.15, dunque assumiamo la normalità.  
   Utilizzando come contrast matrix

[,1] [,2] [,3] [,4]

[1,] 1 -1 0 0

[2,] 1 0 -1 0

[3,] 1 0 0 -1

testiamo le differenze tra le medie con il test T^2. Il pvalue risulta uguale a 0, dunque rifiuto l’ipotesi di uguaglianza delle medie.

b) Gli intervalli di confidenza simultanei, con correzione di Bonferroni, risultano:

Mar.May 41.49724 42.60177 43.70631

Jun.July 49.60182 50.44403 51.28625

Aug.Oct 44.08892 44.69871 45.30850

Nov.Feb 37.64584 38.66500 39.68416

Ci conviene dunque procedere all’acquisto da novembre a febbraio, come ci aspettiamo che sia, in quanto la domanda sarà molto minore.

1. Per verificare la differenza utilizzo una MANOVA. Prima, verifico che siano congiuntamente gaussiane le tre variabili da testare in entrambi i gruppi, cosa vera per mcshapiro.test.

Poi ho bisogno di verificare che la struttura di covarianza sia uguale in entrambi i casi. Analizzandole, risultano molto simili, quindi procedo con l’analisi, che dà evidenza di una differenza tra il ristorante 1 e il ristorante 2.

Procedo poi con degli intervalli T^2 simultanei per le differenze, che sono i seguenti:

inf sup

total 7.5718826 9.9636729

filling 4.6675027 6.6147750

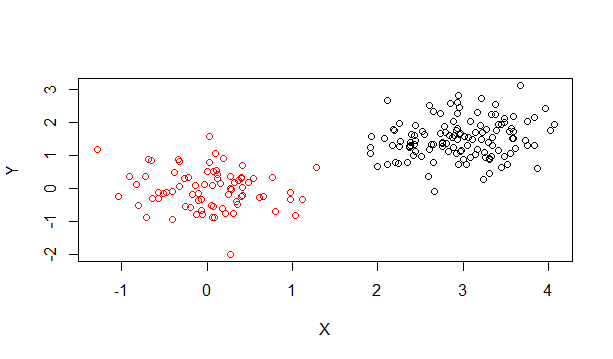
breading 0.3495531 0.9277802

Sembra dunque che il ristorante 1 abbia olive ascolane più grandi, e, soprattutto più ripiene!

1. Effettuo il clustering gerarchico richiesto.  
   Il primo cluster conta 123 punti, con media 2.962927 1.503984.

Il secondo cluster conta 78 punti, con media 0.03269231 -0.02320513.

Il coefficiente cofenetico è .8709168.



Cari archeologi, ho svolto il test statistico da voi richiesto, controllando le medie dei due clusters, che risultano entrambi provenienti da normali bivariate con varianze simili. In particolare, con un test per due popolazioni gaussiane bivariate indipendenti, ottengo prova schiacciante che i due centri siano diversi e, che, dunque ci siano due siti archeologici distinti.  
Distinti saluti,

Filippo Fedeli, statistico capo del consiglio internazionale degli archeologi

1. Procedo a una regressione lineare come richiesto. I coefficienti sono i seguenti:

B0 : 66.49527511

B1 : 19.21718913

B2: 1.69733063

B3 : 0.02371382

Sigma: 8.838284

Gli ultimi due non risultano però significativi. I residui sono però, normali e belli belli con vif bassi.

Utilizzando linearhypothesis, risulta che non c’è evidenza della dipendenza dall’aumento generale del livello del mare. Proviamo dunque a toglierlo dalla regressione.  
Ora è tutto significativo anche se abbiamo perso un po’ di R^2.  
Calcolo i due intervalli di previsione (livello globale 0.9) e ottengo:

int1 = 86.81791 69.01672 104.6191

int2 = 69.44883 51.70745 87.19022

Rischiamo molto nel primo caso, in cui 0.9 è contenuto nell’intervallo, nel secondo caso siamo al sicuro!