



Laboratorio informatico di Statistica Matematica su supporto Excel

Anna Maria Paganoni & Laura Pontiggia

Introduzione

Lo scopo di queste dispense è quello di introdurre il lettore all'uso del software *Excel* nell'analisi statistica di dati, ed esse nascono da esperienze didattiche avute in differenti corsi di Laboratorio di Statistica. Ogni tematica statistica affrontata viene trattata partendo da un esempio concreto sul quale si esegue un'analisi guidata, descritta passaggio per passaggio.

Le nozioni di probabilità e statistica necessarie sono presupposte note, e vengono richiamate solo per uniformare le notazioni.

Numerosi sono i testi didattici di Probabilità e Statistica a cui queste dispense di laboratorio possono essere affiancate, fra i tanti ricordiamo, per omogeneità di notazioni,

D.C.Montgomery, G.C.Runger, N.F.Hubele "Engineering Statistics" (Wiley & Sons, 2001).

Nella stesura delle dispense abbiamo utilizzato la versione italiana di *Excel 2000*, alcuni comandi possono essere differenti in altre versioni del medesimo software.

Le dispense sono corredate di un CD contenente i file in cui è riportata tutta l'analisi descritta nei vari esempi e file contenenti dati su cui il lettore può esercitarsi in modo autonomo.

Glossario

Bloccare le celle: quando una formula matematica contiene il riferimento ad una cella che vogliamo rimanga inalterato tale riferimento deve essere scritto con un dollaro davanti alla lettera che indica la colonna e un dollaro davanti al numero che indica la riga.

Comando vettoriale: quando un comando deve produrre un risultato costituito da un vettore costituito da più celle, dobbiamo selezionare tante celle quante corrispondono alla dimensione di tale vettore ed eseguire il comando cliccando contemporaneamente i comandi **Ctrl-Shift-Enter**.

Sovraimposizione della Normale: se vogliamo sovraimporre la distribuzione Normale ad un istogramma, dalla tabella della distribuzione di frequenza copiamo: la colonna delle classi (in cui è indicato solo l'estremo superiore della classe), la colonna delle densità, e quella delle etichette. Costruiamo poi una nuova colonna in cui inseriamo il valore della densità normale in corrispondenza dell'estremo superiore della classe. Per ottenere questa colonna selezioniamo le celle di output e fra le funzioni statistiche scegliamo la funzione **DISTRIB.NORM (x; media; dev_standard; cumulativo)**. Inseriamo come valori **x** la colonna delle classi, e come **media** e **dev_standard** inseriamo quelle ottenute dal campione, ed in **cumulativo** scriviamo FALSO. Essendo questo un comando vettoriale per eseguirlo clicchiamo contemporaneamente i comandi **Ctrl-Shift-Enter**. Procediamo nella costruzione del grafico scegliendo un istogramma in pila, in cui la prima serie è rappresentata dalla densità e la seconda serie dai valori della normale. Clicchiamo con il tasto destro del mouse sulla parte dell'istogramma che rappresenta la serie della Normale e modifichiamo il tipo di grafico, scegliendo grafico per **Linee**. Infine clicchiamo con il tasto destro del mouse sulla linea e in **Formato serie dati - motivo** scegliamo **linea smussata**.

Trascinare: quando vogliamo far eseguire un'operazione analoga a quella impostata in una cella adiacente è sufficiente posizionare il mouse nell'angolo in basso a destra della cella e poi tenendo schiacciato il tasto sinistro del mouse selezionare tutte le celle in cui vogliamo riportare l'operazione.

Ringraziamenti

Ringraziamo con grande affetto il prof. Piercesare Secchi per averci spronato, sostenuto ed aiutato durante la stesura di queste note.

Indice

1. Variabili qualitative	1
2. Indici di posizione, indici di dispersione. Istogrammi e boxplot	3
3. Analisi descrittiva di una popolazione (esempio guidato)	15
4. Analisi descrittiva di due popolazioni (esempio guidato)	19
5. Analisi descrittiva di più popolazioni (esempio guidato)	25
6. Generazione di numeri casuali	27
7. Normal Probability Plot	29
8. Simulazione: la legge dei grandi numeri	33
9. Simulazione: il teorema centrale del limite	35
10. Probabilità di copertura degli intervalli di confidenza per la media	43
11. Probabilità di copertura degli intervalli di confidenza per una proporzione	49
12. Verifica delle ipotesi	55
13. Test sulla proporzione (esempio guidato)	61
14. Inferenza statistica per una e due popolazioni (esempio guidato)	63
15. Confronto tra due popolazioni (esempio guidato)	75
16. Regressione lineare semplice	79
17. Regressione lineare semplice (esempio guidato)	93
18. Regressione lineare multipla	97
19. Regressione lineare multipla (esempio guidato)	107
20. Regressione lineare con predittori categorici	111
21. Regressione lineare con predittori categorici (esempio guidato)	123

File excel allegati:

1. Descrittiva.xls
2. EserciziDescrittiva.xls
3. Simulazione.xls
4. Test.xls
5. EserciziTest.xls
6. RegressioneLineareSemplice.xls
7. EserciziRegressioneLineareSemplice.xls
8. RegressioneLineareMultipla.xls
9. EserciziRegressioneLineareMultipla.xls
10. RegressioneLineareCategorica.xls