

Corso Integrato di Statistica Informatica e Analisi dei Dati Sperimentali

A.A 2009-2010

Esercitazione F

Scopo dell'esercitazione

Applicazione dei test statistici z-test, f-test, χ^2 -test

Si ricorda che per mandare via mail le relazioni bisogna:

- salvare i dati (qualora ci fossero) in un file ASCII,
- analizzare i dati utilizzando un foglio elettronico (o in script Gnuplot),
- riassumere i risultati in una relazione (1 pagina in formato doc),
- Salvare il file dei dati, il file con i calcoli (xls) e la relazione (.doc) in un unico file compresso da inviare a:

CI.biologia[at]gmail.com

specificando il nome dei partecipanti, l'esercizio e l'esercitazione di riferimento.

1 Esercizi

1.1 Esercizio G1

Il file ASCII **T_test1.dat** riporta i risultati di due set di misure di concentrazione di un inquinante A in due diversi fiumi (campioni indipendenti). Stabilire con un livello di confidenza del 95% se le differenze osservate sono significative. Assumere i campioni omoschedastici.

Nel file EXCEL **t-student.xls** è mostrata una possibile soluzione.

Per lo svolgimento:

- importare i dati in un foglio EXCEL
- calcolare i principali indici statistici: media, deviazione standard, errore sulla media, per i due ste di misure e riportarli in una tabella con le unità di misura corrette e con le cifre significative corrette.
- Calcolare il valore della variabile t_{oss} associata alle misure e confrontarlo con il valore limite per una confidenza del 95%
- e/o: calcolare il valore della significatività per la variabile t_{oss} .
- commentare i risultati dell'esperimento in una breve relazione da inviare come descritto sopra.

1.2 Esercizio G2

Il file ASCII **T_test2.dat** riporta i risultati delle misure di concentrazione di un inquinante A in uno stesso fiume effettuando i prelievi X_1 al mattino e i prelievi X_2 la sera, negli stessi punti. Stabilire con un livello di confidenza del 95% se le differenze osservate nei livelli medi di inquinante sono significative. Assumere i campioni appaiati.

Nel file EXCEL **t-student.xls** è mostrata una possibile soluzione.

Per lo svolgimento:

- importare i dati in un foglio EXCEL
- calcolare i principali indici statistici: media, deviazione standard, errore sulla media, per i due ste di misure e riportarli in una tabella con le unità di misura corrette e con le cifre significative corrette.
- Calcolare il valore della variabile t_{oss} associata alle misure e confrontarlo con il valore limite per una confidenza del 95%
- e/o: calcolare il valore della significatività per la variabile t_{oss} .
- commentare i risultati dell'esperimento in una breve relazione da inviare come descritto sopra.

1.3 Esercizio G3

Per verificare la sensibilità di una specie di piante ad un virus si espongono al contagio M gruppi ($M=90$) di $n = 20$ individui ciascuno. Dopo un tempo certo tempo si conta la frequenza di contagi per ogni gruppo e si riportano in tabella le frequenze N_k con cui sono osservati k contagi per gruppo (file **Chi_test1.dat**). Si vuole verificare l'ipotesi che la probabilità di contagio sia p e che quindi il numero di contagi per gruppo (k) segua una distribuzione binomiale $f(20, p)$. Verificare se si può escludere che la probabilità di contagio sia $p = 0.3$; verificare se si può escludere che la probabilità di contagio sia $p = 0.35$.

Nel file EXCEL **test_chi.xls** è mostrata una possibile soluzione.

Per lo svolgimento:

- importare i dati in un foglio EXCEL
- Calcolare il valore della variabile χ_{oss}^2 associata alle misure e confrontarlo con il valore limite per una confidenza del 95%
- e/o: calcolare il valore della significatività per la variabile χ_{oss}^2 .
- commentare i risultati dell'esperimento in una breve relazione da inviare come descritto sopra.

1.4 Esercizio G4

E' stato effettuato un campionamento per stabilire la densità di piantine in una data regione. Si effettuano 43 campionamenti su aree di $100 m^2$ ciascuna e nel file (file **Chi_test2.dat**) si riportano le frequenze assolute di osservazione. Stabilire mediante un test del χ^2 se la distribuzione di piantine può essere considerata una distribuzione di Poisson o se l'esperimento è più compatibile con una distribuzione uniforme.

Nel file EXCEL **test_chi.xls** è mostrata una possibile soluzione.

Per lo svolgimento:

- importare i dati in un foglio EXCEL
- Determinare i principali indici statistici (media, dev.st. errore sulla media)
- Determinare le distribuzioni di frequenza teoriche nell'ipotesi di una distribuzione di Poisson o nell'ipotesi di distribuzione uniforme. Utilizzare il valore medio trovato per la distribuzione di Poisson.
- confrontare le frequenze teoriche e sperimentali nei due casi (distribuzione uniforme e di Poisson) utilizzando un istogramma di frequenza.
- Calcolare il valore della variabile χ_{oss}^2 associata alle misure e confrontarlo con il valore limite per una confidenza del 95%
- e/o: calcolare il valore della significatività per la variabile χ_{oss}^2 .
- commentare i risultati dell'esperimento in una breve relazione da inviare come descritto sopra.

1.5 Esercizio G5

E' stato effettuato un test di un farmaco somministrando il farmaco o un placebo a un gruppo di pazienti. I risultati sono riassunti nella tabella di contingenza del file: **Chi_test3.dat**.

Stabilire mediante un test del χ^2 se i risultati sono significativamente correlati.

Nel file EXCEL **test_chi.xls** è mostrata una possibile soluzione.

Per lo svolgimento:

- importare i dati in un foglio EXCEL
- Calcolare il valore della variabile χ_{oss}^2 associata alle misure e confrontarlo con il valore limite per una confidenza del 95%
- e/o: calcolare il valore della significatività per la variabile χ_{oss}^2 .
- commentare i risultati dell'esperimento in una breve relazione da inviare come descritto sopra.

1.6 Esercizio G6

Utilizzare i dati dell'esperimento F7 (misura dei tempi di reazione) per stabilire, mediante un test t-student, se i tempi di reazione dei due studenti sono significativamente diversi.

1.7 Esercizio G7

Utilizzando una bilancia da cucina, una caraffa e una bacinella piena d'acqua.

1. prelevare (ad occhio) mezza caraffa d'acqua e pesarla. Ripetere la misura 50 volte
2. Calcolare i principali indici di sintesi statistica.
3. Stabilire se la distribuzione dei pesi si può considerare Normale con valore atteso deviazione standard eguali a quelli ottenuti sperimentalmente utilizzando un test del χ^2
4. Far ripetere l'esperimento (punti 1-3) ad un collega e stabilire, utilizzando un test t-student, se i valori medi ottenuti dai due sperimentatori sono distinguibili con un livello di confidenza del 90% .
5. confrontare i risultati dei due sperimentatori utilizzando un istogramma di frequenza.
6. Preparare una relazione da inviare come descritto sopra. Allegare alla relazione il file EXCEL, il file dei dati sperimentali e il file della relazione.