

Statistica1, metodi matematici e statistici.

a.a. 2004/2005.

Esercizi in R: parte II.

Gli esercizi possono essere risolti in gruppi formati al massimo da 3 persone e dovranno essere risolti entro Mercoledì 08 Giugno. Entro tale data dovrà essere spedita all'indirizzo `luca.monno@unipv.it` un file di testo con la sequenza dei comandi di R richiesti nell'esercizio.

Esercizio 1. Test del rapporto delle verosimiglianze generalizzato

Supponiamo che il seguente campione

$y = (0.309, -1.306, -0.935, -0.437, -0.658, -0.795, -0.983-0.063, -0.535, -2.153)$

sia stato estratto da una popolazione Normale con media μ e varianza σ^2 entrambi incogniti.

- Disegnare il grafico della log-verosimiglianza profilo $\ell_P(\sigma^2)$
- Aggiungere al grafico una retta parallela alle ascisse con ordinata pari a $\ell_P(\hat{\sigma}^2) - 1/2 \text{ qchisq}(0.95, 1)$. Che informazione ci da questa retta?
- Determinare gli estremi dell'intervallo di confidenza ($1 - \alpha = 0.95$) per σ^2 per il campione y tramite il metodo approssimato dell'inversione del test del rapporto delle verosimiglianze. (Suggerimento: la funzione `uniroot` potrebbe essere utile)

Simulare 1000 campioni *i.i.d.* di numerosità 15 da una v.a. Normale con media 0 e varianza 1

- Assumendo μ incognito, per ogni campione determinare gli estremi dell'intervallo di confidenza ($1 - \alpha = 0.95$) per σ^2 tramite il metodo approssimato dell'inversione del test del rapporto delle verosimiglianze e contare quante volte il vero valore di σ^2 vi appartiene.
- Per ogni campione calcolare il rapporto tra la verosimiglianza $L(\mu, \sigma^2)$ calcolata nel punto di massimo vincolato a $\sigma^2 = 1$ e la verosimiglianza calcolata nel punto di massimo non vincolato. Indichiamo con $\lambda(y)$ questo valore, che informazione ci da?

- Applicare la trasformazione $-2 \log \lambda(y)$ sulle 1000 realizzazioni ottenute, farne l'istogramma e sovrapporvi la densità di una v.a. χ_1^2 .

Esercizio 2. Regressione

1. Costruire una funzione che, dati due campioni y, x , restituisca un vettore con la stima dei parametri di un modello del tipo $Y = \alpha + \beta X$ utilizzando la formula matriciale.
2. Generare un campione di numerosità 30 da una normale con media 0 e varianza 5.
3. Generare 1000 campioni provenienti da un modello lineare normale con media $\alpha + \beta X$ dove $\alpha = 2$, $\beta = 3$, X è il vettore generato al punto precedente e l'errore ha varianza unitaria.
4. Per ogni campione simulato calcolare la stima di α e β utilizzando la funzione che avete costruito al primo punto e verificare che le medie delle 1000 realizzazioni di $\hat{\alpha}$ e $\hat{\beta}$ sono vicine a valori veri di α e β . Rappresentare graficamente le 1000 realizzazioni.
5. Aprire il dataset che si trova nel file "*hubble.dat*" e stimare un modello di regressione per le variabili $y = \text{distanza}$ e $x = \text{velocita}$. Commentare i risultati.