Statistica1, metodi matematici e statistici.
a.a. 2003/2004.
Homework su *Verosimiglianza ed R.*

Gli esercizi possono essere risolti in gruppi formati da al massimo 3 persone e dovranno essere risolti entro Lunedì 10 Maggio. Entro tale data dovrà essere spedita agli indirizzi `andrea.tancredi@uniroma1.it` o `lmonno@ciop.mat.uniroma3.it` un file di testo con la sequenza dei comandi di R richiesti nell'esercizio. Si consiglia anche di svolgere l'esercizio preliminare di natura teorica la cui soluzione NON deve essere consegnata.

Esercizio preliminare

Dato un campione casuale $x_1, \ldots, x_n$ estratto da una popolazione X avente distribuzione gamma con densità

$$f(x) = \frac{\lambda^\nu}{\Gamma(\nu)} e^{-\lambda x} x^{\nu-1}$$

- scrivere la verosimiglianza per ν, λ
- scrivere la logverosimiglianza per ν, λ
- Trovare una statistica sufficiente
- scrivere la v.a. score
- scrivere le equazioni di verosimiglianza
- è possibile trovare analiticamente le stime di max verosimiglianza?
- calcolare la matrice di informazione attesa e quella osservata e verificare che coincidono e sono pari a

$$I(\nu, \lambda) = \begin{pmatrix} \frac{n\nu}{\lambda^2} & -\frac{n}{\lambda} \\ -\frac{n}{\lambda} & n\psi''(\nu) \end{pmatrix}$$

dove $\psi''(\nu) = \frac{d^2}{d\nu^2} \log \Gamma(\nu)$

Esercizio 1

- Generare un numero casuale da una distribuzione uniforme con supporto $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ e chiamare λ il numero così generato.
- Generare un numero casuale da una distribuzione con supporto $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ e con probabilità proporzionali a quelle di una Poisson con media 5 e chiamare ν il numero così generato.
- Simulare un campione di numerosità 20 da una v.a. di tipo gamma con densità
$$f(x) = \frac{\lambda^{\nu}}{\Gamma(\nu)} e^{-\lambda x} x^{\nu-1}$$
con i valori di λ e ν ottenuti in precedenza.
- Calcolare le stime di massima verosimiglianza per ν e λ attraverso la procedura `nlm` prendendo come punto iniziale il valore vero dei parametri
- Calcolare la media campionaria m_1 e il momento secondo campionario m_2 (la media delle osservazioni al quadrato)
- Calcolare $m_1^2/(m_2 - m_1^2)$ e $m_1/(m_2 - m_1^2)$. Sono vicini ai valori ν e λ? Perche'? (Suggerimento: sono le stime ottenute con il metodo dei momenti)
- Calcolare media aritmetica del campione e la stima di massima verosimiglianza della media della popolazione.
- Disegnare la funzione di verosimiglianza e la funzione di logverosimiglianza associate al campione per il parametro ν assumendo noto il valore di λ.
- Disegnare la funzione di verosimiglianza e la funzione di logverosimiglianza associate al campione per il parametro λ assumendo noto il valore di ν.
- Disegnare la funzione di verosimiglianza e la funzione di logverosimiglianza associate al campione per il parametro λ, ν.
- Simulare 1000 campioni di numerosità 20 sempre da una v.a. di tipo gamma con densità
$$f(x) = \frac{\lambda^{\nu}}{\Gamma(\nu)} e^{-\lambda x} x^{\nu-1}$$
con i valori di λ e ν ottenuti in precedenza e calcolare su ogni campione la stima di massima verosimiglianza del parametro di forma della gamma e lo stimatore dei momenti $\tilde{\nu} = m_1^2/(m_2 - m_1^2)$.
- Costrure un grafico con l'istogramma della distribuzione della SMV $\hat{\nu}$ e la distribuzione dello stimatore $\tilde{\nu}$ in base ai 1000 campioni ottenuti.

- Calcolare la varianza delle 1000 realizzazioni di $\hat{\nu}$ e la varianza asintotica di $\hat{\nu}$ (ovvero trovare l'elemento realtivo alla varianza di $\hat{\nu}$ nell'inversa della matrice di informazione di Fisher).
- Stimare in base alle mille realizzazioni ottenute gli errori quadratici medi dei due stimatori per i veri valori dei parametri.
- Rappresentare graficamente l'errore quadratico medio della stima di massima verosimiglianza della media della popolazione quando è noto il parametro λ ed è pari ad 1.