Laboratorio di decisioni statistiche (2024-2025)

Lab. 2 – Analisi bayesiana modelli beta-binomiale e Poisson-gamma – 02 ottobre 2024

Nome e matricola: ______________________________

1. Sia $X_1, \ldots, X_n$ un campione causale (i.i.d.) da una popolazione di Bernoulli di parametro θ, con funzione di massa di probabilità
$$p_\theta^X(x_i) = \theta^{x_i}(1-\theta)^{1-x_i}, \qquad \theta \in [0,1].$$
Si consideri per θ una distribuzione a priori Beta(α, β) con funzione di densità
$$\pi(\theta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \qquad \alpha, \beta > 0$$
e funzione di verosimiglianza
$$\ell(\theta) = \theta^{s_n}(1-\theta)^{n-s_n}, \qquad s_n = \sum_{i=1}^n x_i.$$
Ricordare inoltre che $\Theta|\mathbf{z} \sim \text{Beta}(\overline{\alpha}, \overline{\beta})$, con $\overline{\alpha} = \alpha + s_n$ e $\overline{\beta} = \beta + n - s_n$.
   (a) Disegnare $\pi(\theta)$ con $\alpha = 9.2$ e $\beta = 13.8$.
   (b) Calcolare $\mathbb{P}[\Theta > 0.2]$ e $\mathbb{P}(\Theta \in A)$, con $A = [0.2, 0.6]$.
   (c) Sia $\sum_{i=1}^n x_i = 15$ e $n = 20$. Disegnare nello stesso grafico $\pi(\theta)$, $\ell(\theta)$ e $\pi(\theta|\mathbf{z}_n)$.
   (d) Calcolare le le stime puntuali del parametro e le seguenti probabilità posteriori $\mathbb{P}[\Theta > 0.2|\mathbf{z}_n]$ e $\mathbb{P}(\Theta \in A|\mathbf{z}_n)$, con $A = [0.2, 0.6]$.
   (e) Calcolare l'insieme ET di livello $1 - \gamma = 0.95$ per θ.
   (f) Calcolare l'insieme HPD di livello $1 - \gamma = 0.95$ per θ. [Sugg.: consultare l'help per la funzione hdi() nella Library HDInterval].
   (g) (Analisi non informative). Ripetere l'analisi assumendo $\alpha = \beta = 1$ e poi $\alpha = \beta = 0$
   (h) Ripetere l'analisi assumendo $\sum_{i=1}^n x_i = 30$ e $n = 40$
   (i) Ripetere le analisi non informative.

2. Considerare il modello Poisson-Gamma con
$$\alpha = 2, \quad \beta = 3, \quad \sum_{i=1}^n x_i = 15, \quad n = 5.$$
Ricorda che:
$$\ell(\theta) = c \cdot e^{-n\theta}\theta^{\sum x_i}, \ \theta > 0 \qquad \text{e} \qquad \pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}, \ \theta > 0$$
e che $\Theta|\mathbf{z}_n \sim \text{Ga}(\alpha + \sum x_i, \beta + n)$
   (a) Tracciare i grafici di $\pi(\theta)$ e $\ell(\theta)$ [porre `xlim=c(0,5)` e `ylim=c(0,1.5)`].
   (b) Determinare i valori dei parametri di $\pi(\theta|\mathbf{z}_n)$.
   (c) Aggiungere al grafico precedente quello di $\pi(\theta|\mathbf{z}_n)$ e calcolare stime puntuali e la varianza a posteriori di Θ.
   (d) Determinare gli insiemi ET per θ al 95%.
   (e) Determinare gli insiemi HPD per θ al 95%.
   (f) Confrontare le ipotesi $H_0 : \theta \leq 2$ vs. $H_1 : \theta > 2$.
   (g) Confrontare la densità a posteriori di Θ con l'istogramma di frequenza che si ottiene generando $M = 10000$ valori dalla distribuzione a posteriori.
   (h) Ripetere effettuando le analisi non informative.
   (i) Ripetere con $n = 20$ e $\sum_{i=1}^n x_i = 60$.
   (j) Ripetere con $n = 50$ e $\sum_{i=1}^n x_i = 150$.