Analisi Statistica dei Dati: foglio di lavoro su sintesi della finale e scelta dell'iniziale

Luca La Rocca

11 dicembre 2018

Example 6.3

1. Considerate il modello bayesiano

$$L^\bullet(\pi) = \binom{13}{10}\pi^{10}(1-\pi)^3, \quad 0 < \pi < 1,$$
$$f(\pi) = 1, \quad 0 < \pi < 1,$$

con distribuzione finale $\mathrm{Beta}(11, 4)$ e confrontate graficamente la densità iniziale con quella finale.

2. Calcolate media, mediana e moda finali del parametro π.

3. Calcolate e confrontate gli intervalli credibili per π, al livello di credibilità del 95%, con code equilibrate e con massima densità.

4. Calcolate la probabilità finale dell'ipotesi $\{\pi > 1/2\}$.

Example 6.4

5. In una campagna di screening, sia E l'evento "test diagnostico positivo" (evidenza) e H l'evento "soggetto effettivamente malato" (ipotesi). In virtù del teorema di Bayes, il pronostico finale $\omega = \mathbb{P}(H|E)/\mathbb{P}(\bar{H}|E)$ varrà

$$\omega = \frac{\mathbb{P}(E|H)}{\mathbb{P}(E|\bar{H})}\frac{\pi}{1-\pi},$$

dove $\mathbb{P}(E|H)$ è la sensibilità del test, $\mathbb{P}(E|\bar{H})$ il complemento a uno della sua specificità e $\pi = \mathbb{P}(H)$ la prevalenza della malattia. Supponendo che la sensibilità e la specificità del test siano note e pari a 0.9, mentre la vostra incertezza su π è rappresentata dalla distribuzione finale del modello

$$L^\bullet(\pi) = \binom{100}{1}\pi^{1}(1-\pi)^{99}, \quad 0 < \pi < 1,$$
$$f(\pi) = \frac{1}{B(\alpha_0, \beta_0)}\pi^{\alpha_0 - 1}(1-\pi)^{\beta_0 - 1}, \quad 0 < \pi < 1,$$

scegliete α_0 e β_0 in modo tale che la media iniziale di π valga 0.1 e la mediana iniziale di π valga 0.05.

6. Con la scelta fatta, quanto vale la moda iniziale? Quanto vale la numerosità campionaria iniziale? Con quale probabilità iniziale π è minore di 0.15?

7. Calcolate i parametri della distribuzione finale di π, sia essa $\mathrm{Beta}(\alpha_\star, \beta_\star)$, quindi rappresentate graficamente la vostra incertezza su

$$\theta = \omega/(1+\omega)$$

campionando dalla distribuzione finale di π e tracciando un istogramma di densità per i corrispondenti valori di θ; confrontate infine l'istogramma tracciato con la densità teorica

$$f_\star(\theta) = \frac{c_\star}{\theta^2} f_F(c_\star(1/\theta - 1); 2\beta_\star, 2\alpha_\star), \quad \theta > 0,$$

dove $c_\star = \frac{\mathbb{P}(E|H)}{\mathbb{P}(E|\bar{H})}\frac{\alpha_\star}{\beta_\star}$ ed f_F è la densità di Fisher-Snedecor (`df` in R).

8. Sfruttando il campione generato al punto precedente, valutate la media e la mediana finale di θ; trovate quindi un intervallo credibile per θ, con code equilibrate, al livello di credibilità del 95%.

Testo di riferimento

L. Held & D. Sabanés Bové. *Applied Statistical Inference*. Springer, 2014.