

Mètodes Bayesians

Gener 2015

1. Que utilitza un freqüentista alhora de fer inferència? i un Bayesià?
2. Suposa que tens una mostra de 1.000 observacions de la distribució a posteriori d'un paràmetre real, θ . Com estimaries la mitjana i la mediana de la distribució a posteriori d'una funció de θ , $g(\theta)$, a partir d'aquesta mostra?
3. Donat un contrast de tres hipòtesis (H_1 , H_2 i H_3), demostra o desmenteix que si disposes dels Factors de Bayes de la hipòtesis 1 contra la 3, FB_{13} , i de la hipòtesis 3 contra la 2, FB_{32} , llavors pots obtenir el Factor de Bayes de la hipòtesis 1 contra la 2 a partir del càlcul $FB_{12} = FB_{13} \times FB_{32}$.

Uns investigadors amb l'objectiu de mesurar la relació entre la producció de patates (y) i la quantitat de fertilitzant utilitzada (x) han dividit el camp en 8 parcel·les comparables, i en cadascuna d'elles han utilitzat un nivell de fertilitzant diferent. La següent taula recull les dades obtingudes:

fertilitzant (x)	1	1.5	2	2.5	3	3.5	4.5
producció (y)	25	31	27	28	36	35	34

Després d'implementar el següent model Bayesià :

$$Y_1, \dots, Y_7 | \beta_0, \beta_1, \sigma \sim \prod_{i=1}^7 \text{Normal}(\beta_0 + \beta_1 x_i, \sigma)$$
$$\beta_0 \sim \text{Normal}(0, 1000), \beta_1 \sim \text{Normal}(0, 1000), \sigma \sim \text{Uniforme}[0, 100]$$

i mitjançant l'algoritme *Gibbs Sampling* s'obtenen 100.000 simulacions de la distribució a posteriori a partir de les quals es construeix la següent taula:

paràmetre	mitjana	desviació	2.5%	mediana	97.5%
β_0	23.9	4.5	15.2	23.9	32.7
β_1	2.7	1.6	-0.4	2.7	5.8
σ	4.2	2.2	2.1	3.7	9.6

4. Que faries per assegurar-te de que les simulacions amb les que faràs la inferència corresponen realment a simulacions de la distribució a posteriori?
5. Com construïries, a partir de les simulacions, un interval del 95% per a la producció de patates obtinguda quan la quantitat de fertilitzant utilitzada fos $x = 4$?

Assumim que el nombre de gols per partit que fa el Barça al seu camp en un partit de lliga segueix un model $Poisson(\lambda)$,

$$p(y|\lambda) = e^{-\lambda} \frac{\lambda^y}{y!},$$

amb espai de paràmetres $\Omega = [0, \infty)$, i per tant amb $E[y|\lambda] = \lambda$ i $Var[y|\lambda] = \lambda$. Les distribucions a priori conjugades per aquest model són les $Gamma(a, b)$,

$$\pi(\lambda) = \frac{b^a \lambda^{(a-1)} e^{-b\lambda}}{\Gamma(a)},$$

i per tant amb $E[\lambda] = a/b$ i $Var[\lambda] = a/b^2$. A la darrera lliga els gols per partit que va fer el Barça a casa han sigut 5, 1, 2, 2, 3, 3, 5, 4, 4, 5, 6, 2, 2, 3, 5, 1, 4, 2 i 4. Suposant que el nombre de gols per partit són condicionalment independents i idènticament distribuïts:

6. Calcula la funció de versemblança.
7. Tria i justifica una distribució a priori conjugada per λ i calcula la distribució a posteriori.
8. Explica com triaries entre $H_1 : \lambda < 2$, $H_2 : 2 \leq \lambda < 3$ i $H_3 : 3 \leq \lambda$.