

Mètodes Estadístics Bayesians,  
Gener 2012

PART I

1. Quina és la diferència essencial entre l'estadística Bayesiana i l'estadística no Bayesiana? Quins avantatges hi veus a l'estadística Bayesiana?
2. Què és, com es calcula i perquè serveix la distribució a posteriori? Quina relació té amb la funció de versemblança?
3. Partint d'un model estadístic per  $y$ , d'una priori i d'una mostra, com calcularies un interval de credibilitat a posteriori del 90% per predir un valor futur de  $y$ ?
4. Quina diferència hi ha entre afirmar que  $[a, b]$  és un interval de credibilitat del 90% per  $\theta$ , i afirmar que aquest és un interval de confiança del 90% per  $\theta$ ?
5. Quin és el paper de la simulació en la inferència Bayesiana?
6. Si tens una mostra de 1000 observacions de la distribució a posteriori d'un paràmetre real,  $\theta$ , com estimaries les quartiles i la mediana d'aquesta distribució?  
Com estimaries les quartiles i la mediana de la distribució a posteriori d'una funció de  $\theta$ ,  $g(\theta)$ , a partir d'aquesta mostra? Com faries inferència sobre  $g(\theta)$ ?
7. Què són i perquè serveixen els factors de Bayes?
8. Com validaries un model Bayesià?

## PART II

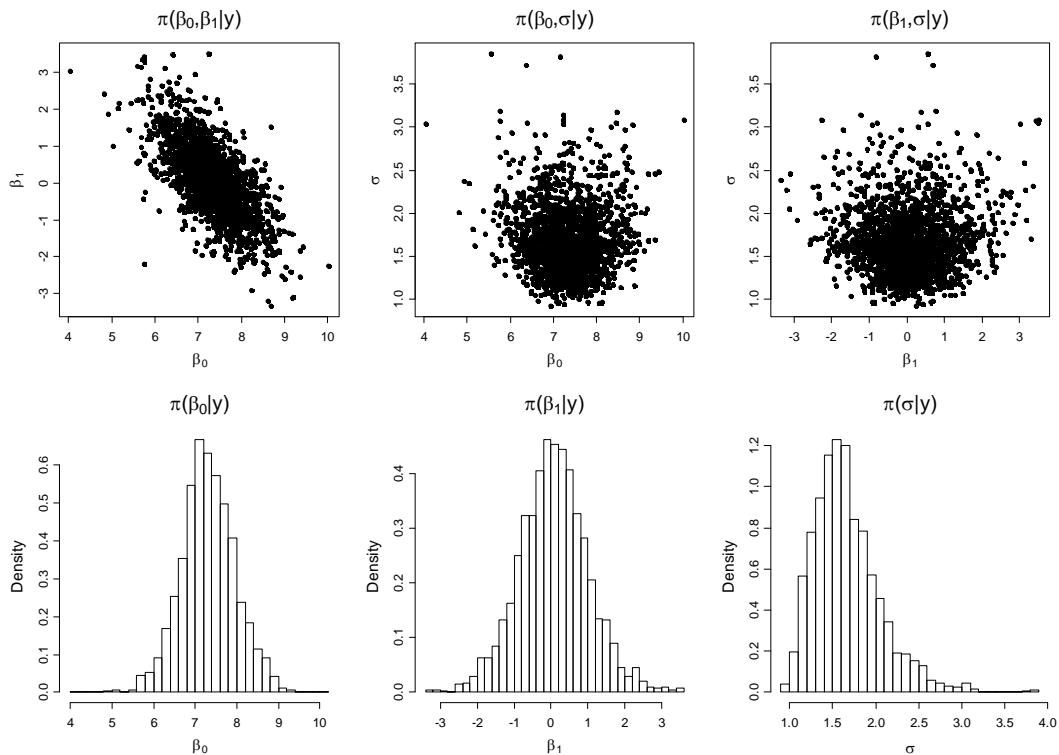
Els professors de l'assignatura d'Inferència Bayesiana volen saber si hi ha diferències significatives en el rendiment dels estudiants a l'assignatura en funció del sexe de l'estudiant. Les notes finals a l'assignatura del curs 2007-2008, excloent els no presentats són:

Nois	Noies
9.6	6.1
7.0	9.1
5.0	8.8
8.0	5.7
8.4	8.9
6.4	6.1
	6.5

Per això formulen el model :

$$\begin{aligned}
 y | \beta_0, \beta_1, \sigma &\sim \text{Normal}(\beta_0 + \beta_1 \text{sexe}, \sigma) \\
 \beta_0 &\sim \text{Normal}(0, 1000) \\
 \beta_1 &\sim \text{Normal}(0, 1000) \\
 \tau = (1/\sigma^2) &\sim \text{Gamma}(0.001, 0.001)
 \end{aligned}$$

on la variable *sexe* val 1 per als nois i 0 per les noies. Els gràfics següents presenten una mostra de 2000 observacions de la distribució a posteriori dels paràmetres del model.



A partir també de les 2000 simulacions de la distribució a posteriori s'ha calculat per a cada paràmetre la mitjana, la desviació i els percentils 5%, 25%, 50%, 75% i 95%.

	<b>mean</b>	<b>sd</b>	<b>5%</b>	<b>25%</b>	<b>50%</b>	<b>75%</b>	<b>95%</b>
$\beta_0$	7.31	0.66	6.27	6.90	7.31	7.73	8.42
$\beta_1$	0.07	0.95	-1.49	-0.55	0.07	0.66	1.64
$\sigma$	1.67	1.11	1.16	1.40	1.61	1.88	2.40

Respon a les següents preguntes:

9. Amb les presses el que ha formulat el model s'ha descuidat d'indicar explícitament quin és l'espai de paràmetres, tot i que implícitament és el suport de la distribució a priori. Creus que aquest espai de paràmetres és l'adequat?
10. Partint del que veus als diagrames adjunts, creus que hi ha diferències significatives entre els nois i les noies. Explica en base a què arribes a aquesta conclusió. Si creus que per contestar aquesta pregunta de forma rigorosa caldria formular un model diferent, explica com ho faries.
11. Com creus que hauria sortit la distribució a posteriori, si en comptes de 13 alumnes n'hi hagués hagut 130?
12. Hi ha una alumna que no va poder fer l'examen per que es va quedar atrapada a un ascensor. Com calcularies un interval de credibilitat del 90% per la nota final d'aquesta alumna?
13. El model bayesià l'han construït de pressa i corrents, sense pensar-s'hi gaire a l'hora de triar la distribució a priori. Quina distribució a priori hauries triat tu i com creus que hauria canviat la distribució a posteriori degut a això?