Pontificia Universidad Católica de Chile
Facultad de Matematicas
Departamento De Estadística
Segundo Semestre 2020
Docente: Fernando Quintana

Interrogación 3 : Parte II

Nicolás Sumonte

Parte A

Para esta entrega usamos las librerias `MASS`, `mvtnorm` y `invgamma`. En esta parte que se encontrará en el script de `R` usamos el metodo de la composición para encontrar σ^2 de una distribución $\chi^2(\nu_0, \sigma_0^2)$ luego de esto, utilizamos este σ^2 para calcular la priori de $\beta|\sigma^2$ con la distribución $N_{k\gamma}(b_{0\gamma}, \sigma^2 B_{0\gamma})$ considerando que $b_{0\gamma}$ y $B_{0\gamma}$ están ya normalizados segun el vector γ. Con estos dos parametros a priori encontrados, generamos la verosimilitud $y|X, \sigma^2, b_{0\gamma}, B_{0\gamma}$ con una distribucion $N_n\left(X_\gamma \beta_\gamma, \sigma^2 I_n\right)$ con β_γ, σ^2 encontrados anteriormente. Para el caso de la distribucion a posteriori de β, σ^2 usamos el hecho que $\sigma^2 \mid \boldsymbol{y}, \boldsymbol{X} \sim \chi^{-2}\left(\nu_n, \sigma_n^2\right)$, y que $\boldsymbol{\beta} \mid \sigma^2, \boldsymbol{y}, \boldsymbol{X} \sim N_k\left(\boldsymbol{b}_n, \sigma^2 B_n\right)$ con

$$\begin{aligned}
B_n &= \left(X^T X + B_0^{-1}\right)^{-1} \\
\boldsymbol{b}_n &= B_n\left(\boldsymbol{X}^\top \boldsymbol{X}\hat{\beta} + B_0^{-1}\boldsymbol{b}_0\right) = \left(\boldsymbol{X}^T\boldsymbol{X} + B_0^{-1}\right)^{-1}\left(\boldsymbol{X}^T\boldsymbol{y} + B_0^{-1}\boldsymbol{b}_0\right) \\
\nu_n &= \nu_0 + n \\
\nu_n\sigma_n^2 &= \nu_0\sigma_0^2 + \boldsymbol{b}_0^\top B_0^{-1}\boldsymbol{b}_0 + \boldsymbol{y}^T\boldsymbol{y} - \boldsymbol{b}_n^T B_n^{-1}\boldsymbol{b}_n
\end{aligned}$$

dado, esto repetimos el metodo de la composición encontrando σ^2 y luego β a posteriori. Finalmente se calculan las densidades y realizamos la multiplicación $p\left(\boldsymbol{\beta}, \sigma^2\right) = p\left(\boldsymbol{\beta} \mid \sigma^2\right) p\left(\sigma^2\right)$ para las prioris y posteriores respectivas de σ^2 y β para finalmente calcular la marginal de y según la formula

$$\log(p(y)) = \log(p(y \mid \theta)) + \log(p(\theta)) - \log(p(\theta \mid y))$$

dada en el enunciado.

Todo este procedimiento, está realizado bajo la función `marginal posteriori` en el script entregado.

Parte B

Notemos que para esta parte, bajo una priori para $\gamma^{\mathcal{L}}$ podriamos calcular la probabilidad a posteriori de la siguiente manera:

$$p(\boldsymbol{\gamma} \mid y, X) = \frac{p(\boldsymbol{\gamma})p(y \mid X, \boldsymbol{\gamma})}{\sum_\gamma p(\boldsymbol{\gamma})p(y \mid X, \boldsymbol{\gamma})}$$

y con estas posterioris, bajo razones de equiprobabilidad a priori, es posible contrastar evidencias de los modelos a de la siguiente forma

$$(\gamma_1, \gamma_2 \mid y, \mathrm{X}) = \frac{p(\gamma_1 \mid y, \mathrm{X})}{p(\gamma_2 \mid y, \mathrm{X})} = \frac{p(\gamma_1)}{p(\gamma_2)} \times \frac{p(y \mid X, \gamma_1)}{p(y \mid X, \gamma_2)}$$

con ello, el factor de bayes que nos dice cuanta informacion favorece el modelo γ_1 sobre el modelo γ_2 queda dado por $\frac{p(y|X,\gamma_1)}{p(y|X,\gamma_2)}$ y por consiguiente, sería la division entre dos valorizaciones de nuestra función creada en A, con diferentes valores del vector γ

Parte C

Para esta parte escogí como b_0 los minimos cuadrados ordinarios, debido a que estos minimizan la suma de cuadrados del exponente, bajo un argumento parecido, escogí B_0 como la varianza de los vectores β. Siguiendo la misma linea, para encontrar σ^2 que se ajustara, utilicé el σ^2 obtenido por a regresión. En cuanto a ν_0 utilicé el minimo de tamaño muestral para que las distribuciones estuvieran definidas y esto es 1. Con esto se obtuvieron los siguientes resultados

Valores de γ	**Coeficientes de Regresión**	**Valores de $p_\gamma(y)$**
$(1,1,1,1,1)$	$\beta_1 \times$ `agricultura` $+ \beta_2 \times$ `examen` $+ \beta_3 \times$ `educación` $+ \beta_4 \times$ `catolico` $+ \beta_5 \times$ `mortalidad`	50,67691
$(1,0,1,1,1)$	$\beta_1 \times$ `agricultura` $+ \beta_3 \times$ `educación` $+ \beta_4 \times$ `catolico` $+ \beta_5 \times$ `mortalidad`	47,66516
$(0,0,1,1,1)$	$\beta_3 \times$ `educación` $+ \beta_4 \times$ `catolico` $+ \beta_5 \times$ `mortalidad`	43,2107
$(0,1,1,1,1)$	$\beta_2 \times$ `examen` $+ \beta_3 \times$ `educación` $+ \beta_4 \times$ `catolico` $+ \beta_5 \times$ `mortalidad`	47,9726

Tabla Nº1: valores encontrados para $p_\gamma(y)$ dado el vector γ

Con estos valores es posible comparar los valores del test de bayes, explicado en el inciso anterior. A continuación se muestran los resultados obtenidos para 6 posibles combinaciones

Modelos comparados	Factor de bayes	Decisción
$\frac{(1,1,1,1,1)}{(1,0,1,1,1)}$	1,063186	Se opta por modelo $(1, 0, 1, 1, 1)$ (Sin examen)
$\frac{(1,1,1,1,1)}{(0,1,1,1,1)}$	1,056	Se opta por modelo $(0, 1, 1, 1, 1)$ (Sin agricultura)
$\frac{(1,1,1,1,1)}{(0,0,1,1,1)}$	1,17272	Se opta por modelo $(0, 0, 1, 1, 1)$ (Sin agricultura y sin Examen)
$\frac{(1,0,1,1,1)}{(0,0,1,1,1)}$	1,103087	Se opta por modelo $(0, 0, 1, 1, 1)$ (Sin agricultura y sin Examen)
$\frac{(1,0,1,1,1)}{(0,1,1,1,1)}$	1,103087	Se opta por modelo $(1, 0, 1, 1, 1)$ (Sin Examen)
$\frac{(0,0,1,1,1)}{(0,1,1,1,1)}$	0,900	Se opta por modelo $(1, 0, 1, 1, 1)$ (Sin Examen y sin agricultura)

Tabla Nº2: Modelos comparados y su respectivo valor del factor de bayes.

Con esto, podemos concluir que el mejor modelo es el que no tiene a los factores `Agricultura` ni `Examen`. Sin embargo quien realizo el análisis de regresión estaba bien al sospechar que `Agricultura` no tenia influencias y al dudar del modelo con `examen` debido a que este ultimo, esta cerca del modelo sin `Examen` y sin `Agricultura`

Parte D

Podemos encontrar $p_\gamma(y)$ con la integral

$$\begin{aligned} p_\gamma(y \mid X, \gamma) &= \iint p(y, \beta)\left(\sigma^2 \mid X, \gamma\right) d\beta d\sigma^2 \\ &= \iint p(y \mid \beta, X) p\left(\beta \mid X, \gamma, \sigma^2\right) p\left(\sigma^2\right) d\beta d\sigma^2 \end{aligned}$$

y ahora si integramos primero respecto a β nos queda que

$$\begin{aligned} p_\gamma(y \mid X, \gamma) &= \int \left(\int p\left(y \mid X, \gamma, \sigma^2, \beta\right) p\left(\beta \mid X, \gamma, \sigma^2\right) d\beta\right) p\left(\sigma^2\right) d\sigma^2 \\ &= \int p\left(y \mid X, \gamma, \sigma^2\right) p\left(\sigma^2\right) d\sigma^2 \end{aligned}$$

llegando así a una formula para esta marginal