ÉCOLE CENTRALE DE NANTES

BIRATS: A BIVARIATE NORMAL HIERARCHICAL MODEL

Projet Bayes

Élèves :
Lamiae ARIFALLAH
Rim YOUSFI
Meryem Cherqi

Encadrants :
Mathieu RIBATET

15 avril 2024

Table des matières

1 Introduction

Dans l'exemple des rats nous illustrons l'utilisation d'une distribution Normale multivariée (MVN) pour les coefficients de régression de la courbe de croissance pour chaque rat. C'est le modèle adopté par Gelfand et al. (1990) pour ces données, et il suppose a priori que les paramètres d'interception et de pente pour chaque rat sont corrélés. Par exemple, une corrélation positive impliquerait que les rats initialement lourds (interception élevée) ont tendance à prendre du poids plus rapidement (pente plus raide) que les rats plus légers. Le modèle est le suivant :

$$Y_{ij} \sim \text{Normale}(m_{ij}, \tau_c)$$
$$m_{ij} = b_{1i} + b_{2i}x_j$$
$$\mathbf{b}_i \sim \text{MVN}(\mathbf{m}_b, \mathbf{W})$$

où Y_{ij} est le poids du ième rat mesuré à l'âge x_j, et $\mathbf{b}_i$ désigne le vecteur (b_{1i}, b_{2i}). Nous supposons des priors Normaux univariés indépendants 'non-informatifs' pour les composantes séparées m_{b1} et m_{b2}. Un prior Wishart Wishart(R, r) a été spécifié pour $\mathbf{W}$, la matrice de précision de population des coefficients de régression. Pour représenter une connaissance a priori vague, nous avons choisi les degrés de liberté r de cette distribution aussi petits que possible (c'est-à-dire 2, le rang de $\mathbf{W}$). La matrice d'échelle a été spécifiée comme suit :

$$R = \begin{pmatrix} 200 & 0 \\ 0 & 0.2 \end{pmatrix}$$

Cela représente notre estimation a priori de l'ordre de grandeur de la matrice de covariance $\mathbf{W}^{-1}$ pour $\mathbf{b}_i$ (voir le manuel Classic BUGS (version 0.5) section sur les modèles normaux multivariés), et est équivalent à la spécification a priori utilisée par Gelfand et al. Enfin, un prior Gamma Gamma$(0.001, 0.001)$ non informatif a été supposé pour la précision de mesure τ_c.

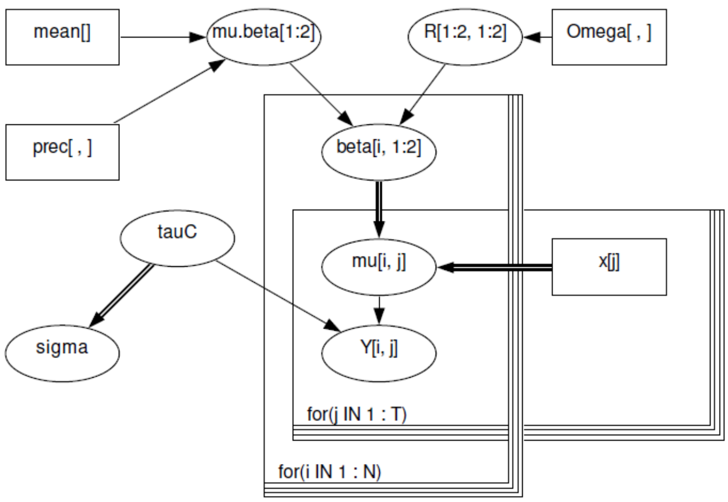


FIGURE 1 – graphe du modèle

2 Modèle mathématique

3 Interprétation des résultats

Pour estimer les variables aléatoires listées précedemment, nous avons utilisé pour celà l'algorithme de Gibbs. Cet algorithme est une méthode d'échantillonnage de Monte Carlo par chaînes de Markov (MCMC) qui permet de générer des échantillons à partir de la distribution a posteriori des paramètres étant donné les données.Afin d'assurer la convergence de l'algorithme, nous avons utilisé une phase de "burn-in" de 1000 échantillons. Durant cette phase, les échantillons générés par l'algorithme sont écartés et non utilisés pour estimer les paramètres. Cette étape permet de réduire l'influence des conditions initiales sur les résultats finaux.

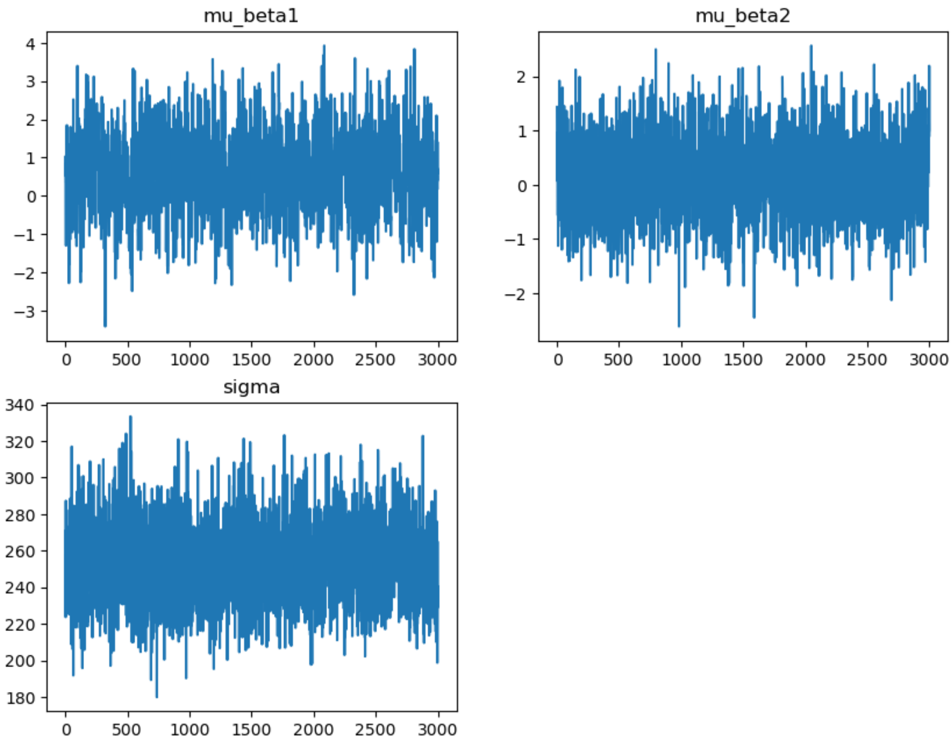
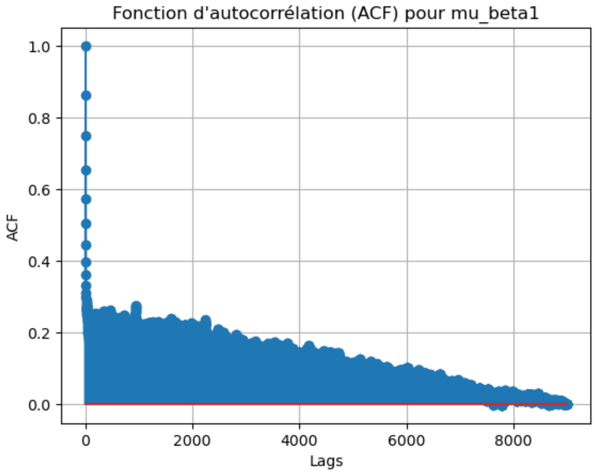
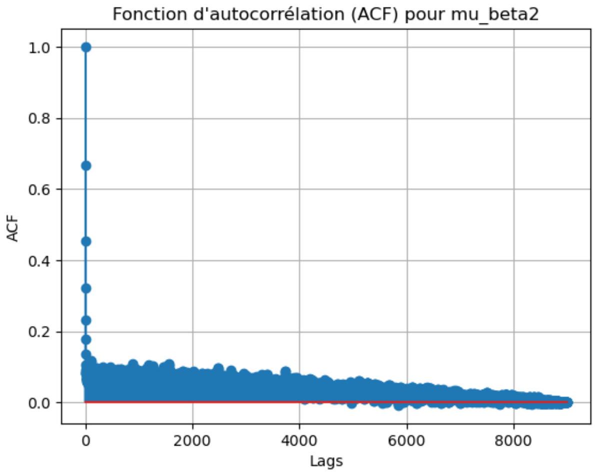


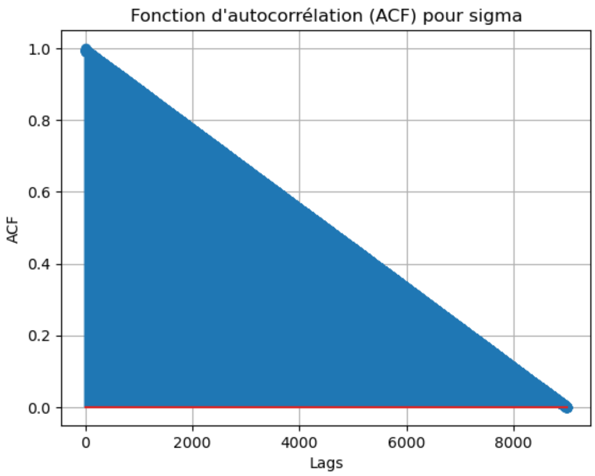
FIGURE 2 – Les chaines de Markov obtenues

La moyenne de μ_{β_1} est 0.6395 et sa variance est 1.0651. Cela signifie que la moyenne de l'interception des courbes de croissance des rats est d'environ 0.6395, avec une variation autour de cette moyenne d'environ 1.0651.
La moyenne de μ_{β_2} est 0.1495 et sa variance est 0.4878. Cela signifie que la moyenne de la pente des courbes de croissance des rats est d'environ 0.1495, avec une variation autour de cette moyenne d'environ 0.4878.
La moyenne de σ est 253.0487 et sa variance est 448.4689. Cela représente la moyenne et la variance de la distribution de mesure de l'erreur, qui est utilisée dans le modèle pour prendre en compte les fluctuations ou les écarts par rapport aux valeurs prédites par le modèle.

Fonctions ACF :


Fonction d'autocorrélation (ACF) pour mu_beta1
ACF
Lags
0
2000
4000
6000
8000
0.0
0.2
0.4
0.6
0.8
1.0


Fonction d'autocorrélation (ACF) pour mu_beta2
ACF
Lags
0
2000
4000
6000
8000
0.0
0.2
0.4
0.6
0.8
1.0


Fonction d'autocorrélation (ACF) pour sigma
ACF
Lags
0
2000
4000
6000
8000
0.0
0.2
0.4
0.6
0.8
1.0