

Cómputo científico para probabilidad y estadística. Tarea 9.

MCMC: Tarea Final

Juan Esaul González Rangel

Diciembre 2023

En ambos problemas hay que diseñar e implementar el MCMC, investigar sobre su convergencia y tener algún grado de certeza sobre si sí se está simulando de la posterior correspondiente. Más aún, recuerde que se trata de un problema de inferencia: Hay que hablar del problema en sí, comentar sobre las posteriores simuladas y posibles estimadores (a partir de la muestra de posterior) que se pueden proporcionar de cada parámetro.

1. (**Problema en ecología**) Sean X_1, \dots, X_m variables aleatorias donde X_i denota el número de individuos de una especie en cierta región. Suponga que $X_i|N, p \sim \text{Binomial}(N, p)$, entonces

$$f(\bar{x}|N, p) = \prod_{i=1}^m \frac{N!}{x_i!(N-x_i)!} p^{x_i} (1-p)^{N-x_i}.$$

Asumiendo la distribución a priori $p \sim \text{Beta}(\alpha, \beta)$ y $N \sim h(\cdot)$, donde h es una dist. discreta en $\{0, 1, 2, \dots, N_{\text{máx}}\}$, se tiene definida la distribución posterior $f(N, p|\bar{x})$.

A partir del algoritmo MH, simule valores de la distribución posterior usando un kernel híbrido. Para ello considere como sugerencia la siguiente distribución inicial para el MCMC

$$p \sim U(0, 1) \text{ y } N \sim U_d \left\{ \max_{i \in \{1, \dots, m\}} (x_i), \max_{i \in \{1, \dots, m\}} (x_i) + 1, \dots, N_{\text{máx}} \right\}$$

y las propuestas

- Propuesta 1: De la condicional total de p (kernel Gibbs).
- Propuesta 2: De la a priori.
- Propuesta 3: Propuesta hipergeométrica (j ?).
- Propuesta 4: Poisson: $N_p \sim \max_{i \in \{1, \dots, m\}} (x_i) + \text{Poisson}(?)$.
- Propuesta 5: Caminata aleatoria

$$N_p = N + \varepsilon, \quad P(\varepsilon = 1) = \frac{1}{2} = P(\varepsilon = -1).$$

Los datos son estos: 7, 7, 8, 8, 9, 4, 7, 5, 5, 6, 9, 8, 11, 7, 5, 5, 7, 3, 10, 3.

A priori, esperamos que sea difícil observar a los individuos entonces $\alpha = 1, \beta = 20$. La especie no es muy abundante y entonces $N_{\text{máx}} = 1000$ y $h(N) = 1/(N_{\text{máx}} + 1); N \in \{0, 1, 2, \dots, N_{\text{máx}}\}$.

Las propuestas y distribución inicial para el MCMC de arriba son **solamente sugerencia**, propongan otras propuestas, experimenten y comenten.

2. (**Estudio de mercado**) Se tiene un producto y se realiza una encuesta con el fin de estudiar cuánto se consume dependiendo de la edad. Sea Y_i el monto de compra y X_i la covariable la cual representa la edad.

Suponga que $Y_i \sim Po(\lambda_i)$ (distribución Poisson con intensidad λ_i)

$$\lambda_i = cg_b(x_i - a)$$

para g_b la siguiente función de liga

$$g_b(x) = \exp\left(-\frac{x^2}{2b^2}\right).$$

O sea, se trata de regresión Poisson con una función liga no usual. Si $\lambda_i = 0$ entonces $P(Y_i = 0) = 1$. a = años medio del segmento (años), c = gasto promedio (pesos), b = “amplitud” del segmento (años).

Considere las distribuciones a priori

$$a \sim N(35, 5), \quad c \sim Gama(3, 3/950), \quad b \sim Gama(2, 2/5).$$

El segundo parámetro de la normal es desviación estandar y el segundo parámetro de las gammas es tasa (rate).

Usando MH simule de la distribución posterior de a , c y b .

Los datos son estos, $n = 100$:

```
X = array([ 25, 18, 19, 51, 16, 59, 16, 54, 52, 16, 31, 31, 54, 26, 19, 13, 59, 48, 54, 23, 50, 59,
55, 37, 61, 53, 56, 31, 34, 15, 41, 14, 13, 13, 32, 46, 17, 52, 54, 25, 61, 15, 53, 39, 33, 52, 65,
35, 65, 26, 54, 16, 47, 14, 42, 47, 48, 25, 15, 46, 31, 50, 42, 23, 17, 47, 32, 65, 45, 28, 12, 22,
30, 36, 33, 16, 39, 50, 13, 23, 50, 34, 19, 46, 43, 56, 52, 42, 48, 55, 37, 21, 45, 64, 53, 16, 62,
16, 25, 62])
```

```
Y = array([1275, 325, 517, 0, 86, 0, 101, 0, 0, 89, 78, 83, 0, 1074, 508, 5, 0, 0, 0, 1447, 0, 0,
0, 0, 0, 0, 0, 87, 7, 37, 0, 15, 5, 6, 35, 0, 158, 0, 0, 1349, 0, 35, 0, 0, 12, 0, 0, 2, 0, 1117, 0,
79, 0, 13, 0, 0, 0, 1334, 56, 0, 81, 0, 0, 1480, 177, 0, 29, 0, 0, 551, 0, 1338, 196, 0, 9, 104, 0,
0, 3, 1430, 0, 2, 492, 0, 0, 0, 0, 0, 0, 0, 0, 1057, 0, 0, 0, 68, 0, 87, 1362, 0])
```

- Investiga y describe muy brevemente los softwares OpenBugs, Nimble, JAGS, DRAM, Rtwalk, Mcee Hammer, PyMCMC.