XXXII ESCUELA NACIONAL DE OPTIMIZACIÓN Y ANÁLISIS NUMÉRICO

INFERENCIA BAYESIANA PARA EL MODELO VON MISES

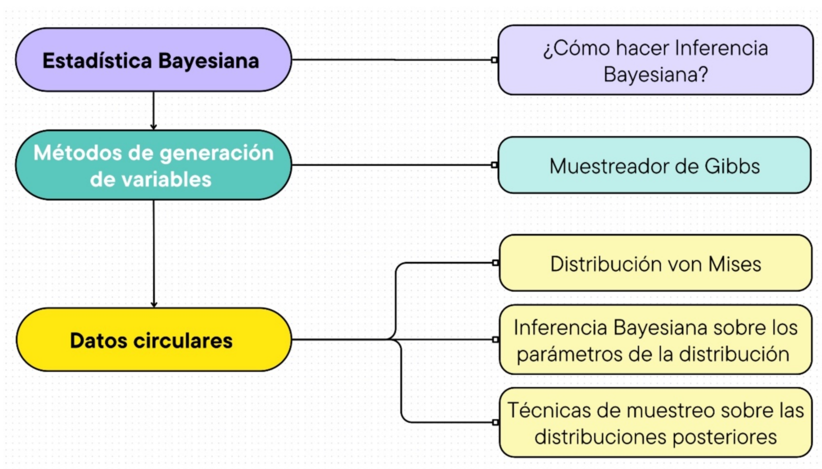
Autor: Lic. Brenda Michelle Dominguez Cruz
Asesor: Dr. Gabriel Nuñez Antonio
Universidad Autónoma Metropolitana
michzurc010@gmail.com, gabnunez@xanum.uam.mx

Casa abierta al tiempo UNIVERSIDAD AUTÓNOMA METROPOLITANA

1. Introducción

Lo datos circulares surgen naturalmente en diferentes fenómenos de la vida real. En consecuencia, su estudio resulta relevante en diferentes áreas de la ciencia y, por lo tanto, es necesario desarrollar técnicas que nos permitan analizar de manera adecuada este tipo de datos. Uno de los modelos más importantes y más usado para describir datos circulares unimodales es el modelo von Mises. La estadística bayesiana es un enfoque bastante útil para el análisis de datos circulares. En este trabajo se describe una metodología empleando técnicas en la estadística bayesiana, para llevar a cabo inferencias sobre los parámetros del modelo von Mises. Estas técnicas están basadas en los métodos de Monte Carlo vía Cadenas de Markov (MCMC).

2. Preliminares



2.1. Estadística e Inferencia Bayesiana

La estadística Bayesiana es una metodología que hace uso del conocimiento previo que se tiene sobre la situación a enfrentar y de manera sistémica poder obtener mejores resultados.

El teorema de Bayes nos plantea que la distribución posterior de los parámetros $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k)$ está dada por

$$f(\boldsymbol{\theta} \mid x_1, ..., x_n) = \frac{f(x_1, ..., x_n \mid \boldsymbol{\theta}) f(\boldsymbol{\theta})}{\int f(x_1, ..., x_n) f(\boldsymbol{\theta}) d\boldsymbol{\theta}} \propto L(x_1, ..., x_n \mid \boldsymbol{\theta}) f(\boldsymbol{\theta}) \quad (1)$$

donde $\{x_1, ..., x_n\}$ es una muestra aleatoria de datos tales que $x_i \sim f(x_i|\boldsymbol{\theta})$ para $i = 1, ..., n$. A $f(\boldsymbol{\theta})$ se le conoce como la distribución inicial o *a priori* de $\boldsymbol{\theta}$ y $L(x_1, ..., x_n \mid \boldsymbol{\theta})$ es la verosimilitud.

2.2. Muestreador de Gibbs general

El muestreador de Gibbs es un método MCMC que nos permite muestrear de la distribución final conjunta mediante sus distribuciones condicionales completas basados en iteraciones que generan Cadenas de Markov que convergen a una distribución de equilibrio.

2.2.1. Algoritmo de Gibbs

1. Hacer $j = 0$ y proponer un valor inicial
$$\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, ..., \theta_k^{(0)})$$
2. Para $j \geq 1$ obtener $\boldsymbol{\theta}^{(j)} = (\theta_1^{(j)}, ..., \theta_k^{(j)})$ de $\boldsymbol{\theta}^{(0)}$ mediante las distribuciones condicionales:
$$\theta_1^{(j)} \text{ de } f(\theta_1|\theta_2^{(j-1)}, \theta_3^{(j-1)}, ..., \theta_k^{(j-1)})$$
$$\vdots$$
$$\theta_k^{(j)} \text{ de } f(\theta_k|\theta_1^{(j)}, \theta_2^{(j)}, ..., \theta_{k-1}^{(j)})$$

A partir de un número h de iteraciones nos aseguramos la convergencia (Teorema Ergódico) de la cadena y buscamos un número k tal que la correlación entre los valores de la cadena $\boldsymbol{\theta}^{(h+k)}, \boldsymbol{\theta}^{(h+2k)}, ..., \boldsymbol{\theta}^{(h+Nk)}$ sea despreciable.

2.3. ¿Qué son los datos circulares?

Los datos circulares surgen de la necesidad de trabajar con datos medidos por la *brújula* y el *reloj*. Tal como en la estadística lineal, podemos establecer ciertas medidas descriptivas para los datos en el círculo unitario y conocer en dónde se concentran los datos, qué tanto se agrupan o dispersan y otras características de la forma en la que se distribuyen.

Los datos circulares pueden considerarse de dos formas:

$$x = (\cos\theta, \operatorname{sen}\theta)^T \qquad z = e^{i\theta} = \cos\theta + i \operatorname{sen}\theta.$$

3. El modelo von Mises

La distribución de von Mises es la distribución más estudiada para datos circulares. Fue descrita por primera vez por Richard von Mises en 1918 como una forma de modelar la distribución de pesos atómicos. Es un modelo simétrico y unimodal, que juega un papel análogo al de la distribución normal. De hecho, en la literatura a menudo se le denomina *distribución normal circular*.

La distribución de von Mises $\boldsymbol{M(\mu, \kappa)}$ [1] tiene como *función de densidad*

$$f(\theta) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)} \quad (2)$$

donde I_0 denota la función de Bessel modificada de primera especie y orden $p = 0$ definida por:

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos\theta} d\theta,$$

4. Inferencia Bayesiana para el modelo von Mises

Sean $\{x_1, ..., x_n\}$ una muestra aleatoria de datos tales que $x_i \sim \boldsymbol{M(\mu, \kappa)}$ y sea $\boldsymbol{\mu} = (\cos\mu, \operatorname{sen}\mu)^T$. Notemos que la distribución von Mises tiene dos parámetros, por lo que podemos hacer inferencia de tres formas: cuando conocemos un parámetro y desconocemos el otro o cuando ambos parámetros son desconocidos. En el caso de que μ sea desconocido y κ conocido, se propone una distribución inicial von Mises y obtenemos que la distribución final es también von Mises, por lo que no es necesario aplicar algún método de muestreo.

4.1. Parámetro μ conocido, κ desconocido

Buscamos conocer la distribución final $f(\kappa \mid \theta_1, ..., \theta_n)$. Forbes y Mardia [2] proponen la siguiente distribución inicial para el parámetro κ

$$f(\kappa) \propto \frac{1}{I_0(\kappa)^a} e^{-b\kappa}$$

y mediante el teorema de Bayes obtenemos que

$$\boldsymbol{f(\kappa \mid \theta_1, ..., \theta_n) \propto \frac{1}{I_0(\kappa)^\eta} e^{-\eta\beta_0\kappa}} \quad (3)$$

donde

$$\eta = n + a \qquad \beta_0 = \frac{1}{\eta}\left[b - \sum_{i=1}^{n} \cos(\theta_i - \mu)\right]$$

Aplicamos el algoritmo de Forbes y Mardia [2] para una muestra de tamaño $n = 2000$, de datos proveniente de una distribución von Mises con parámetros $\boldsymbol{M(\pi/3, 0{,}5)}$. El algoritmo genera una muestra de $N = 1000$ datos cuyo histograma es

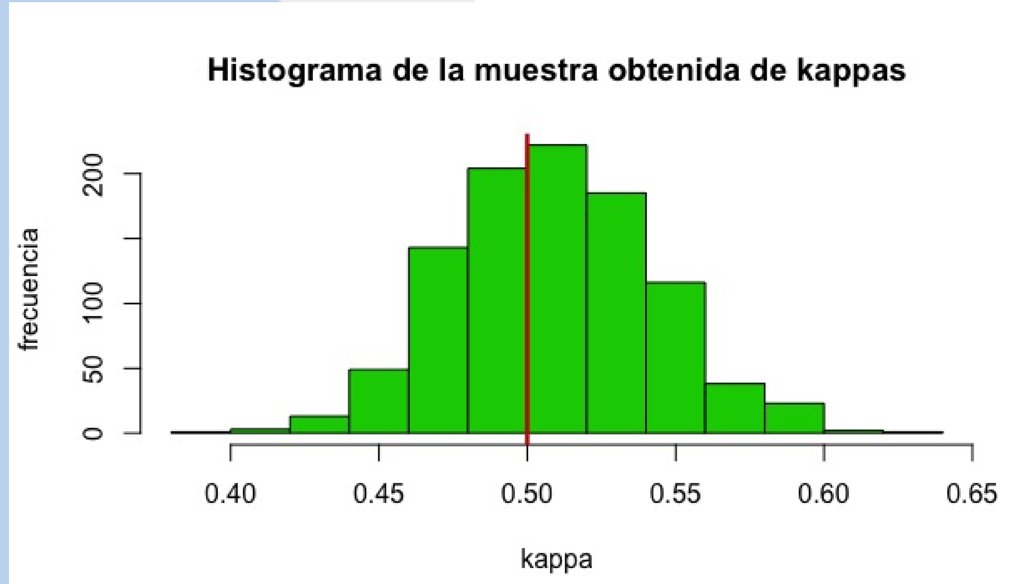


Figura 1: Histograma obtenido usando el algoritmo de Forbes y Mardia [2]

4.2. Parámetros μ y κ desconocidos

Buscamos conocer la distribución posterior de ambos parámetros $f(\mu, \kappa \mid x_1, ..., x_n)$. Guttorp y Lockhart [3] proponen la siguiente distribución inicial conjugada

$$f(\mu, \kappa) \propto \frac{1}{I_0(\kappa)^a} e^{\kappa R_0 \mu^T m_0},$$

donde $c, R_0 \in [0, \infty), m_0 = (\cos\gamma_0, \operatorname{sen}\gamma_0)^T$ son parámetros iniciales.

Y mediante el teorema de Bayes obtenemos que

$$\boldsymbol{f(\mu, \kappa \mid x_1, ..., x_n) \propto \frac{1}{I_0(\kappa)^{n+a}} e^{\kappa \mu^T \{\sum_{i=1}^n x_i + R_0 m_0\}}} \quad (4)$$

Se obtuvieron las densidades condicionales completas de esta distribución:

- Para μ
$$\boldsymbol{f(\mu|\kappa, x_1, ..., x_n) \equiv M(\mu_n, \kappa R_n)} \quad (5)$$
con
$$\mu_n = \left(R_0 m_0 + \sum_{i=1}^{n} x_i\right) R_n^{-1}$$
$$R_n = \left\| R_0 m_0 + \sum_{i=1}^{n} x_i \right\|$$
- Para κ
$$\boldsymbol{f(\kappa|\mu, x_1, ..., x_n) \propto \frac{1}{I_0(\kappa)^\eta} e^{-\eta\beta_0\kappa}} \quad (6)$$
con
$$\eta = n + a$$
$$\beta_0 = \frac{1}{\eta}\left(R_0 \cos(m_o - \mu) - \sum_{i=1}^{n} \cos(\theta_i - \mu)\right)$$

Notemos que al conocer las distribuciones condicionales, pordemos aplicar el Muesreador de Gibbs, dentro del cual aplicamos el algorimo de Forbes y Mardia para muestrear de la densidad condicional de κ.

Se obtuvo una muestra de $N = 1000$ datos para μ y κ, usando una muestra inicial de tamaño $n = 50$ proveniente de una distribución von Mises $\boldsymbol{M(45\pi/180, 10)}$

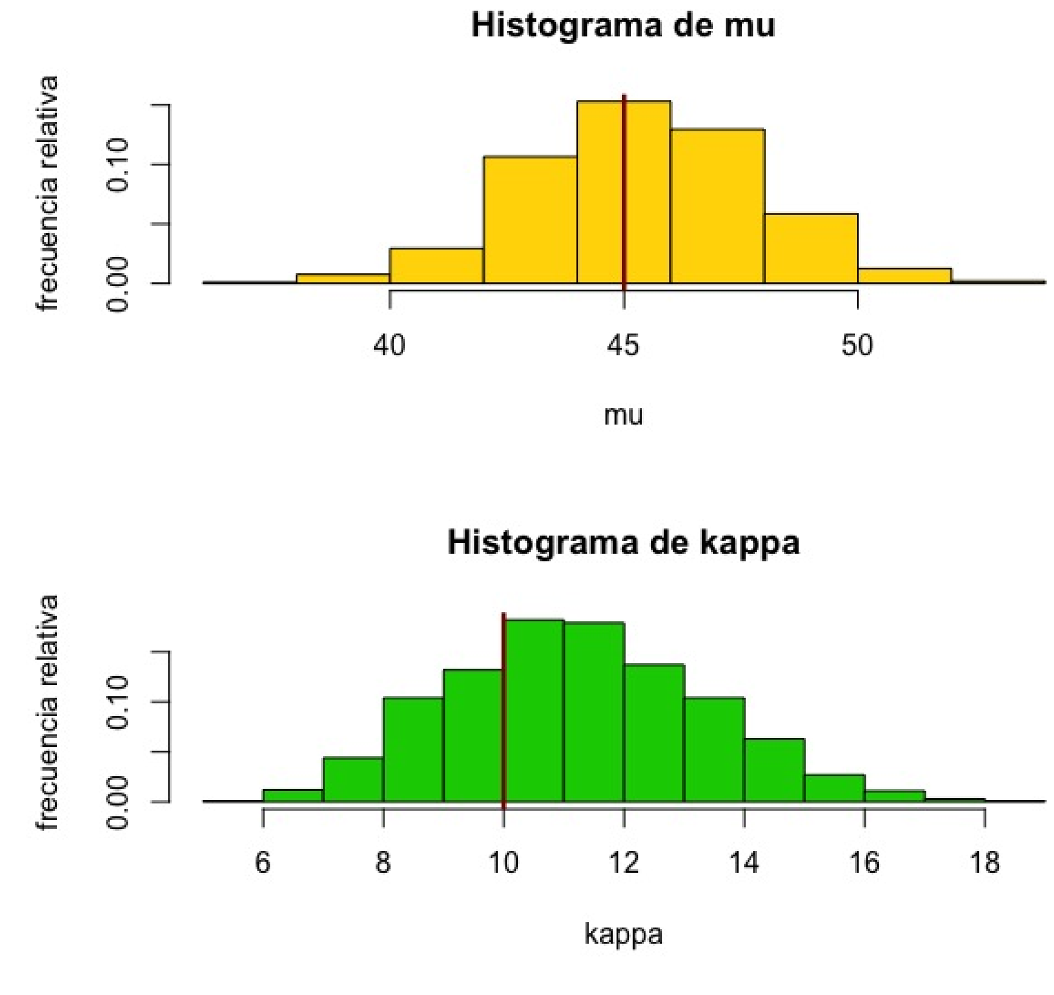


Figura 2: Distribuciones finales de μ y κ usando un Gibbs sampler junto al algoritmo de Forbes y Mardia

Referencias

[1] K. V. Mardia and P. E. Jupp. *Directional Statistics*. John Wiley & Sons, Ltd, Chichester, 1999.

[2] P. G. F. y Kanti V. Mardia. A fast algorithm for sampling from the posterior of a von mises distribution. *Journal of Statistical Computation and Simulation*, 85(13), Mayo 2014.

[3] P. G. y Richard A. Lockhart. Finding the location of a signal: A bayesian analysis. *Journal of the American Statistical Association*, 83(402), Junio 1988.