



# COLEGIO DE POSTGRADUADOS

---

---

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

## CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA

SERIES DE TIEMPO CON MÚLTIPLES PUNTOS DE CAMBIO Y  
OBSERVACIONES CENSURADAS

**RENÉ CASTRO MONTOYA**

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA  
OBTENER EL GRADO DE:

**DOCTOR EN CIENCIAS**

MONTECILLO, TEXCOCO, EDO. DE MÉXICO  
2009

---

---

La presente tesis, titulada: “**Series de tiempo con múltiples puntos de cambio y observaciones censuradas**”, realizada por el alumno: “**René Castro Montoya**”, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**DOCTOR EN CIENCIAS**

**SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA**

**CONSEJO PARTICULAR**

**CONSEJERO**

\_\_\_\_\_

Dr. Sergio Pérez Elizalde

**DIRECTOR DE TESIS**

\_\_\_\_\_

Dr. Gabriel A. Rodríguez Yam

**ASESOR**

\_\_\_\_\_

Dr. Enrique R. Villa Diharce

**ASESOR**

\_\_\_\_\_

Dr. Gilberto Rendón Sánchez

**ASESORA**

\_\_\_\_\_

Dra. Martha Elva Ramírez Guzmán

# SERIES DE TIEMPO CON MÚLTIPLES PUNTOS DE CAMBIO Y OBSERVACIONES CENSURADAS

René Castro Montoya, Dr.  
Colegio de Postgraduados, 2009

Debido a factores externos a las variables de interés, una serie de tiempo puede presentar cambios en la estructura del modelo o en algunos parámetros, y debido a limitaciones en los instrumentos de medición puede presentar también censura en las observaciones. Por ejemplo, cuando se monitorean contaminantes del aire, como pueden ser hidrocarburos aromáticos (PAHs), monóxido de carbono (CO), dióxido de sulfuro ( $SO_2$ ), etc., las series de tiempo observadas pueden tener mediciones censuradas y cambios en la estructura del modelo. En esta tesis se propone un modelo bayesiano para series de tiempo con un número desconocido de puntos de cambio y con algunas observaciones censuradas, donde cada segmento es un proceso autoregresivo de orden uno. Se consideran distribuciones iniciales conjugadas para las medias y las varianzas en cada segmento, excepto para los coeficientes autoregresivos, ya que se condiciona para que la serie sea estacionaria en los segmentos. Para analizar este modelo se utiliza el algoritmo de cadenas de Markov Monte Carlo con saltos reversibles (RJMCMC, por sus siglas en inglés) desarrollado por [Green \(1995\)](#). Este algoritmo consiste en crear una cadena de Markov irreducible y aperiódica que alterna saltos entre varios modelos con espacios de parámetros de diferente dimensión. El problema de censura se resuelve utilizando el método de [Jung \*et al.\* \(2005\)](#) que consiste en simular los valores censurados con la distribución normal multivariada truncada de la “parte censurada” condicionada sobre la “parte observada”. Para ejemplificar el método propuesto se analiza un conjunto de datos simulados con 10 % y 40 % de las observaciones censuradas.

**Palabras clave:** Imputación múltiple, inferencia bayesiana, algoritmo de Metrópolis, algoritmo de cadenas de Markov Monte Carlo con saltos reversibles.

# TIME SERIES WITH MULTIPLE CHANGE POINTS AND CENSURED OBSERVATIONS

René Castro Montoya, Dr.  
Colegio de Postgraduados, 2009

Due to external factors on the variables of interest a time series can present a change in the structure of the model and/or in some parameters and due to the detection limit of the record device it can present also some type of censure on the observations, e.g., when monitoring air pollutants, such as particulate polycyclic aromatic hydrocarbons (PAHs), carbon monoxide (CO) and sulphur dioxide ( $SO_2$ ), the time series obtained may have censored observations and changes in the structure of the model. In this thesis a bayesian model to a time series with an unknown number of change points and censored observations is proposed, each piece is modeled as an autoregressive process of order one. It is assumed conjugate priors distributions for the mean and the variance in each segment. The posterior distribution of the number and the locations of the change points can not be obtained analytically. To analyze this model the reversible jump Markov chain Monte Carlo algorithm is proposed, this algorithm is based on creating an irreducible and aperiodic Markov chain that can alternate jumps among various models with parameter spaces of different dimensions, while retaining detailed balance. The censored problem is solved by the method of [Jung \*et al.\* \(2005\)](#) by imputing the censored values from a multivariate normal distribution given the observed part. The method is exemplified on a simulated data set with two different percentages of censored observations, namely 10% and 40%.

**Key words:** Multiple imputation, bayesian inference, Metropolis algorithm, reversible jump Markov chain Monte Carlo algorithm.

A mi esposa Deicy Yuniba y a mis hijos René Antonio y Ana Sofia, por la grandeza que me inspira su amor, la fortaleza que me dan sus palabras y por hacer con su ayuda que este proyecto y muchos otros, se hagan realidad.

## AGRADECIMIENTOS

En primer lugar a mi asesor y director de tesis el Dr. Gabriel A. Rodríguez Yam por sus valiosas sugerencias, correcciones y dedicación en el presente trabajo de tesis.

Al Dr. Sergio Pérez Elizalde, por sus observaciones, consejos, confianza y ayuda desinteresada que me ha brindado para desarrollar este trabajo de investigación.

Al Dr. Humberto Vaquera Huerta, con admiración y respeto, por todo su apoyo durante mi estancia aquí en el COLPOS.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado durante la realización de mis estudios.

Al PROMEP-UAS por el otorgamiento de recursos económicos, para la realización de mis estudios de doctorado.

Al Colegio de Postgraduados, por las facilidades que me brindaron durante mi estancia en el COLPOS, así como por la formación académica adquirida en sus aulas.

A la Universidad Autónoma de Sinaloa, en particular a la Facultad de Ciencias Físico-Matemáticas, por darme las facilidades y respaldo para desarrollar este proyecto doctoral.

A los integrantes de mi Consejo Particular: los doctores Martha Elva Ramírez Guzmán, Gilberto Rendón Sánchez y Enrique R. Villa Diharce, por sus comentarios y sugerencias al presente trabajo de tesis.

A Netzahualcóyotl Castañeda Leyva, Leonel Pérez Hernández, Fidel Ulín Montejó, Paulino Pérez Rodríguez y Ema Olivares, por su amistad, apoyo, consejos y ayuda incondicional durante todo el doctorado. Sinceramente mil gracias.

A mis profesores, compañeros de clases y todos aquellos que de alguna u otra manera fueron copartícipes de esta tarea.

# CONTENIDO

<b>1. INTRODUCCIÓN</b>	<b>1</b>
1.1. Objetivos . . . . .	2
1.1.1. Objetivo general . . . . .	2
1.1.2. Objetivos particulares . . . . .	2
1.2. Organización de la tesis . . . . .	3
1.3. Revisión de literatura . . . . .	3
<b>2. MATERIALES Y MÉTODOS</b>	<b>6</b>
2.1. Series de tiempo . . . . .	6
2.1.1. Modelos Autoregresivos . . . . .	8
2.2. El problema de censura . . . . .	8
2.2.1. El algoritmo de Jung <i>et al.</i> (2005) . . . . .	9
2.3. El problema de series de tiempo con puntos de cambio . . . . .	11
2.3.1. Prueba de hipótesis para detección de puntos de cambio . . . . .	12
2.3.2. Estimación de puntos de cambio . . . . .	13

## CONTENIDO

---

2.4. Inferencia bayesiana . . . . .	13
2.5. Cadenas de Markov Monte Carlo . . . . .	16
2.5.1. Algoritmo de Gibbs . . . . .	18
2.5.2. Algoritmo de Metrópolis-Hastings . . . . .	18
2.6. Cadenas de Markov Monte Carlo con saltos reversibles . . . . .	19
2.7. Método para monitorear convergencia en RJMCMC . . . . .	22
<b>3. SERIES DE TIEMPO CON PC Y CEN</b>	<b>25</b>
3.1. Modelo bayesiano para el problema de series de tiempo con PC y CEN . . . . .	25
3.2. Implementación del algoritmo RJMCMC al problema de PC y CEN . . . . .	27
3.3. Análisis de datos simulados . . . . .	30
<b>4. CONCLUSIONES</b>	<b>43</b>
<b>BIBLIOGRAFÍA</b>	<b>44</b>
Apéndice A: Función de verosimilitud y densidad condicional . . . . .	46
Apéndice B: Densidades finales de los parámetros autoregresivos . . . . .	47
Apéndice C: Probabilidad de aceptación de un nuevo punto de cambio . . . . .	49



# LISTA DE CUADROS

3.1. Estimaciones para el caso de dos puntos de cambio (10 % censura) . . . . .	35
3.2. Estimaciones para el caso de dos puntos de cambio (40 % censura). . . . .	39
3.3. Estimaciones para el caso de dos puntos de cambio (0 % censura). . . . .	42

# LISTA DE FIGURAS

3.1. Una realización del proceso en (3.6). . . . .	31
3.2. Observaciones censuradas (10 %) de la realización del proceso (3.6). . . . .	32
3.3. Histograma del número de puntos de cambio (10 % censura) . . . . .	33
3.4. Distribución final de $\tau_1$ y $\tau_2$ , condicionado a $k = 2$ (10 % censura). . . . .	34
3.5. Observaciones censuradas (40 %) de la realización del proceso (3.6). . . . .	36
3.6. Histograma del número de puntos de cambio (40 % censura) . . . . .	37
3.7. Histograma de las localizaciones de los PC (40 % censura) . . . . .	38
3.8. Histograma del número de puntos de cambio . . . . .	40
3.9. Histograma de las localizaciones de los puntos de cambio . . . . .	41

# Capítulo 1

## INTRODUCCIÓN

Debido a factores externos a las variables de interés, una serie de tiempo puede presentar cambios en la estructura del modelo o en algunos de los parámetros, y debido a limitaciones en los instrumentos de medición presentar también censura en las observaciones. Por ejemplo, cuando se monitorean contaminantes del aire, como pueden ser hidrocarburos aromáticos (PAHs), monóxido de carbono (CO), dióxido de sulfuro ( $SO_2$ ), entre otras, las series de tiempo observadas pueden tener mediciones censuradas y cambios en la estructura del modelo. Otro ejemplo es en el estudio de altura de las nubes, una variable importante a considerar para evitar catástrofes aéreas, la cual puede estar sujeta a censura debido a límites de detección en el dispositivo de medición.

El problema de series de tiempo con puntos de cambios y observaciones censuradas es un problema de inferencia estadística, en el que se desconocen los parámetros del modelo y el número de parámetros. Éste problema se puede formular mediante inferencia conjunta de un indicador  $r$  del modelo y el vector de parámetros  $\theta_r$ , donde el indicador del modelo determina la dimensión  $n_r$ , [Green \(2003\)](#) comenta que la manera natural de realizar inferencia acerca de  $(r, \theta_r)$  es mediante un análisis bayesiano. La teoría bayesiana describe un problema combinando sentido común y evidencia observacional en un modelo simple de incertidumbre. Éste consiste en especificar una función de verosimilitud y una distribución inicial.

En este trabajo se propone un modelo bayesiano para series de tiempo con un número desconocido de puntos de cambio y observaciones censuradas, donde cada segmento es un proceso autoregresivo de orden uno. Se consideran distribuciones iniciales conjugadas para las medias y las varianzas en cada segmento. Para los coeficientes autoregresivos se condiciona

## 1.1. Objetivos

---

para que la serie sea estacionaria en los segmentos.

Para estimar el número y las localizaciones de los puntos de cambio se implementa el algoritmo de cadenas de Markov Monte Carlo con saltos reversibles desarrollado por [Green \(1995\)](#), éste algoritmo consiste en crear una cadena de Markov irreducible y aperiódica que alterna saltos entre varios modelos con espacios de parámetros de diferente dimensión, que cumpla la condición de probabilidad de equilibrio, asegurando la convergencia a la distribución final. Los valores censurados se simulan de una distribución normal multivariada dada la parte observada ( [Jung et al., 2005](#)). Para ilustrar el algoritmo se analiza un conjunto de datos simulados de series de tiempo con dos puntos de cambio y distintos porcentajes de observaciones censuradas (10 %, 40 %).

## 1.1. Objetivos

### 1.1.1. Objetivo general

Proponer y analizar un modelo bayesiano para series de tiempo que presentan cambios en su estructura y censura en algunas de las observaciones, se asume que el número de puntos de cambio y sus localizaciones son desconocidos.

### 1.1.2. Objetivos particulares

1. Plantear un modelo bayesiano para series de tiempo con puntos de cambio y observaciones censuradas.
2. Implementar el método de [Jung et al. \(2005\)](#) para resolver el problema de censura.
3. Implementar el algoritmo RJMCMC para generar muestras de la distribución final del número de puntos de cambio y las localizaciones de éstos.
4. Estimar el número y las localizaciones de los puntos de cambio de una serie de tiempo que presenta cambios en su estructura y censura en algunas de las observaciones.
5. Analizar un conjunto de datos simulados.

## 1.2. Organización de la tesis

En el presente capítulo se presenta una introducción al problema de series de tiempo con puntos de cambio y con algunas de las observaciones censuradas. También se presentan los objetivos de esta investigación y se hace una revisión de la literatura del problema de puntos de cambio, especialmente los relacionados con la teoría clásica y bayesiana, así como los métodos para resolver el problema de censura en series de tiempo.

En el Capítulo 2 se presentan ideas básicas de series de tiempo y se plantea el problema de series de tiempo con censura y puntos de cambio y algunos conceptos de inferencia bayesiana. También se presentan propiedades de una cadena de Markov, los algoritmos de Gibbs y Metrópolis-Hastings, así como el algoritmo RJMCMC. El capítulo termina con una breve introducción sobre las especificaciones de la densidad propuesta y la función de salto y así como el método de [Castelloe \(1998\)](#) para monitorear la convergencia de una cadena de Markov Monte Carlo con saltos reversibles.

En el Capítulo 3 se propone un modelo bayesiano para series de tiempo con puntos de cambio y censura en algunas de las observaciones. Se implementa el algoritmo de [Jung et al. \(2005\)](#) para resolver el problema de censura, así como el algoritmo RJMCMC para estimar el número de puntos de cambio y sus localizaciones de una serie de tiempo. Finalmente, se presentan los resultados de analizar un conjunto de datos simulados de series de tiempo con dos puntos de cambio y distintos porcentajes de información censurada (10 %, 40 %). En el Capítulo 4 se presentan las conclusiones del trabajo de tesis y tres apéndices: en el apéndice A se presenta la distribución normal multivariada y sus propiedades básicas, en el apéndice B se obtienen las densidades condicionales para  $\mu$ ,  $\sigma$  y  $\phi$  que se utilizan en la impletación del algoritmo RJMCMC y en el apéndice C se obtiene la expresión analítica de la probabilidad de aceptación en la implementación del algoritmo RJMCMC.

## 1.3. Revisión de literatura

El problema de puntos de cambio ha sido objeto de estudio de muchos autores, e.g., [Davis et al. \(2006\)](#) modelan series de tiempo no estacionarias segmentando la serie en bloques de procesos autoregresivos, se asumen desconocidos el número de puntos de cambio, sus localizaciones y orden de los procesos autoregresivos en cada segmento. Taylor (2000) desarrolla un pro-

### 1.3. Revisión de literatura

---

cedimiento de análisis de puntos de cambio que es extremadamente flexible en diferentes conjuntos de observaciones ordenados en el tiempo, e.g., datos de atributo, distribuciones no normales. [Lavielle y Teyssiere \(2006\)](#) consideran el problema de series de tiempo multivariadas con múltiples puntos de cambio.

Algunos autores han realizado un análisis bayesiano del problema de series de tiempo con puntos de cambio, e.g., [Chib \(1998\)](#) modela el proceso de puntos de cambio mediante una cadena de Markov con las probabilidades de transición restringidas a que los regímenes provienen de una secuencia no reversible, [Barbieri y Ohagan \(1996\)](#) realizan un análisis bayesiano en series de tiempo autoregresivas con puntos de cambio y [Giordani y Khon \(2008\)](#) consideran el problema de modelación e inferencia para procesos sujetos a cambios en los parámetros en tiempos desconocidos. Ellos proponen un algoritmo que mejora la eficiencia de muestreo en modelos de variables discretas y mezclas.

Algunos autores han desarrollado métodos para analizar series de tiempo con observaciones censuradas, e.g., [Robinson \(1980\)](#) trabaja el caso de procesos autoregresivos con observaciones censuradas, agrupando las observaciones tal que cada segmento incluya una observación censurada (y así se requiera una integral univariada), para simular las observaciones censuradas mediante la esperanza condicional de la parte censurada dada la parte observada. [Jung et al. \(2005\)](#), trabaja el caso de series de tiempo autoregresivas con observaciones censuradas simulando las observaciones censuradas, mediante un vector de valores de la distribución normal multivariada de la parte censurada dada la parte observada, éste método consiste en los siguientes pasos: i) construir una matriz de permutación, ii) seleccionar valores iniciales para la media y la matriz de covarianzas de la distribución normal multivariada, iii) simular los valores censurados mediante la distribución normal multivariada condicional truncada de la parte censurada dada la parte observada, iv) utilizar éstas para completar el conjunto de observaciones, v) estimar la media y la matriz de covarianzas de la distribución normal multivariada. [Ariza-Hernandez y Rodríguez-Yam \(2008\)](#) utilizan el algoritmo EM para estimar los parámetros de un modelo de espacio de estados con observaciones censuradas.

Solo en algunos trabajos se han considerado simultáneamente los problemas de puntos de cambio y censura. Por ejemplo, [Komarkova \(2004\)](#) considera el problema de detectar puntos de cambio en la distribución de los tiempos de vida con algunas observaciones censuradas por la derecha mediante métodos no paramétricos. [Shuangquan \(1998\)](#) considera el problema de puntos de cambio en una sucesión de variables aleatorias independientes con algunas observaciones censuradas por la derecha y propone un procedimiento que consiste en una

### 1.3. Revisión de literatura

---

extensión del estadístico de rangos de Wilcoxon para probar si un punto de cambio ocurre aleatoriamente en una muestra aleatoria. [Gijbels y Gurler \(2003\)](#) consideran el problema de estimar un punto de cambio en una función de riesgo constante en segmentos con algunas observaciones censuradas.

# Capítulo 2

## MATERIALES Y MÉTODOS

En este capítulo se presentan definiciones y resultados que se requieren en este trabajo. En la Sección 2.1 se presentan ideas básicas de series de tiempo, mientras que en la Sección 2.2 se plantea el problema de series de tiempo con censura. En la Sección 2.3 se considera el problema de detección y estimación de puntos de cambio. En la Sección 2.4 se presentan algunos conceptos de inferencia bayesiana, en la Sección 2.5 se presentan conceptos de cadenas de Markov, especialmente los relacionados con la teoría y uso de las técnicas MCMC y se presentan también los algoritmos Metropolis-Hasting y el muestreo de Gibbs. En la Sección 2.6 se presenta el algoritmo de cadenas de Markov Monte Carlo con saltos reversibles. El capítulo termina con una breve introducción acerca de la especificación de la densidad propuesta y la función de salto, así como un análisis de convergencia de la implementación del algoritmo RJMCMC.

### 2.1. Series de tiempo

En esta sección se presentan conceptos de series de tiempo que se utilizan en este trabajo de tesis. En particular los conceptos de estacionariedad, función de covarianza y función de autocovarianza. Detalles más avanzados sobre definiciones y conceptos de series de tiempo se pueden consultar en [Hamilton \(1994\)](#), [Shumway y Stoffer \(2000\)](#) y [Brockwell y Davis \(2002\)](#).

Llamamos serie de tiempo a un conjunto de mediciones registradas secuencialmente en el tiempo. Las series temporales que se manejan en Economía y en otras áreas en donde se



## 2.1. Series de tiempo

---

utiliza el análisis de series de tiempo, están constituidas por observaciones históricas, es decir, no proceden de la experimentación y por tanto, son irrepetibles.

**Definición 2.1** *Un modelo de serie de tiempo para los datos observados  $y_1, y_2, \dots$ , es una especificación de la distribución conjunta de una sucesión de variables aleatorias  $Y_1, Y_2, \dots$ , de las que  $y_1, y_2, \dots$ , es una realización.*

Aunque es importante la distinción entre la realización  $y_1, y_2, \dots$ , de una serie de tiempo y la serie de tiempo  $Y_1, Y_2, \dots$ , a partir de esta sección se usa el término de serie de tiempo para referirse a la realización y a la serie de tiempo sin hacer distinción, lo cuál es muy común en textos de series de tiempo.

**Definición 2.2** *Sea  $Y_1, Y_2, \dots$ , una serie de tiempo con  $E(Y_t^2) < \infty$ ,  $t = 1, 2, \dots$ . La función media se define como  $\mu_Y(t) := E(Y_t)$ ,  $t = 1, 2, \dots$  y la función covarianza de  $Y_t$  se define como  $\gamma_Y(r, s) := Cov(Y_r, Y_s)$  para todo  $r$  y  $s$  enteros.*

**Definición 2.3** *La serie de tiempo  $Y_1, Y_2, \dots, Y_n$  es débilmente estacionaria si  $E(Y_t) = \mu$ ,  $t = 1, 2, \dots, n$ , y  $\gamma_Y(t+h, t)$  es independiente de  $t$  para cada  $h$ .*

**Definición 2.4** *Sea  $Y_1, Y_2, \dots$ , una serie de tiempo débilmente estacionaria. La función autocovarianza de  $Y_t$  con  $t = 1, 2, \dots$ , a una distancia  $h$  es*

$$\gamma_Y(h) := Cov(Y_{t+h}, Y_t). \quad (2.1)$$

**Definición 2.5** *Se dice que la serie de tiempo  $Y_1, Y_2, \dots$  es estrictamente estacionaria si para cualquier colección finita  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$  del proceso se cumple que  $\forall h$ ,  $F_{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}}(c_{t_1}, c_{t_2}, \dots, c_{t_n}) = F_{Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}}(c_{t_1}, c_{t_2}, \dots, c_{t_n})$ .*

La definición anterior dice que si se selecciona  $t_n$  variables aleatorias y éstas se desplazan  $h$  unidades de tiempo, la distribución conjunta de las variables aleatorias no cambia.

## 2.2. El problema de censura

---

### 2.1.1. Modelos Autoregresivos

Los modelos más comunes en series de tiempo son los modelos autoregresivos.

**Definición 2.6** *La serie de tiempo  $Y_{t_1}, Y_{t_2}, \dots$ , sigue un proceso  $AR(p)$  si*

$$Y_t = \mu + \phi_1(Y_{t-1} - \mu) + \dots + \phi_p(Y_{t-p} - \mu) + \epsilon_t, \quad t = 1, 2, \dots \quad (2.2)$$

donde  $\epsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$  y  $\mu, \sigma^2, \phi_1, \phi_2, \dots, \phi_p$  son los parámetros del modelo.

## 2.2. El problema de censura

Un problema frecuente en series de tiempo, debido a límites de detección en los aparatos de medición, es la presencia de observaciones censuradas. Por ejemplo, cuando se monitorean contaminantes del aire, como pueden ser hidrocarburos aromáticos (PAHs), monóxido de carbono (CO) y dióxido de sulfuro ( $SO_2$ ), las series de tiempo obtenidas pueden tener observaciones censuradas. Otro ejemplo es el estudio de la altura de las nubes, la cual pueden estar sujeta a censura debido a límites de detección en el dispositivo de medición. Mas ejemplos se pueden encontrar en series de tiempo en medicina, economía, hidrología, etc. Algunas publicaciones y libros sobre este tema son, [Robinson \(1980\)](#), [Little y Rubin \(1987\)](#), [Hopke et al. \(2001\)](#), [Jung et al. \(2005\)](#) y [Ariza-Hernandez y Rodríguez-Yam \(2008\)](#).

[Dempster et al. \(1977\)](#) introducen el algoritmo EM para simular las observaciones censuradas, mediante la distribución condicional de las observaciones censuradas dada la parte observada. Este método iterativo consiste en dos pasos, en el primero, se reemplazan las observaciones censuradas mediante la esperanza condicional dada la parte observada y se estiman los parámetros, en el segundo paso se actualiza la estimación de los parámetros mediante el método de máxima verosimilitud basado en las observaciones simuladas y observadas. [Robinson \(1980\)](#) trabaja el caso de series de tiempo autoregresivas con observaciones censuradas, agrupa las observaciones tal que cada segmento incluya una observación censurada y así se requiera una integral univariada para simular las observaciones censuradas mediante la esperanza condicional de la parte censurada dada la parte observada, como sigue,

## 2.2. El problema de censura

---

$$E[Y_c|Y_o = x_o, Y_c \in D_c] = \frac{\int_{D_c} y_c f(y_c|y_o) dy_c}{\int_{D_c} f(y_c|y_o) dy_c}. \quad (2.3)$$

donde  $y_c$  es la observación censurada,  $y_o$  representa el vector de mediciones observadas,  $D_c$  es la región de censura y  $f(y_c|y_o)$  es la densidad condicional.

### 2.2.1. El algoritmo de Jung *et al.* (2005)

Frecuentemente el método de imputación múltiple se utiliza para simular observaciones censuradas, éste consiste en utilizar un modelo para extraer información relevante del conjunto de observaciones para simular las observaciones censuradas. Las observaciones simuladas se utilizan para completar el conjunto de observaciones.

El algoritmo de Jung *et al.* (2005) es una adaptación del algoritmo de imputación múltiple, éste método consiste en construir una matriz de permutación, seleccionar valores iniciales para la media y la matriz de covarianzas de la distribución normal multivariada y simular la parte censurada mediante un vector de valores de la distribución condicional normal multivariada de las observaciones censuradas dada la parte observada. Las observaciones simuladas se utilizan para completar el conjunto de observaciones, estimar la media y la matriz de covarianzas de la distribución normal multivariada. El algoritmo de Jung *et al.* (2005) que se utiliza en el Capítulo 4 para simular valores censurados es como sigue:

Sea  $x_1, x_2, \dots, x_n$  una realización de la serie de tiempo  $X_1, X_2, \dots, X_n$  y suponga que esta serie se ha modelado como un proceso  $AR(1)$ , i.e.  $X_t = \mu + \phi(X_{t-1} - \mu) + \epsilon_t$ ,  $1 \leq t \leq n$ , donde  $\mu$ ,  $\sigma^2$  y  $\phi$  son los parámetros del modelo y  $\epsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$ . Algunas observaciones están censuradas por la derecha en  $c_t$ ,  $t = 1, 2, \dots, n$ , i.e., se observa  $y_t := \min(x_t, c_t)$ , en lugar de  $x_t$ .

La media y la matriz de covarianzas del proceso  $X_t$  están dadas por

$$\begin{aligned} \mu_X(t) &= \mu, \\ \Sigma_{ij} &= \frac{\sigma^2}{(1 - \phi^2)} \phi^{|i-j|}. \end{aligned}$$

Considérense las observaciones  $y_1, y_2, \dots, y_n$ , donde  $y_i$  es observado o censurado por la derecha. Sean  $y_o = \{x_i | x_i \leq c_i, i = 1, 2, \dots, n\}$  y  $y_c = \{c_i | x_i > c_i, i = 1, 2, \dots, n\}$ . La idea

## 2.2. El problema de censura

---

es reemplazar los valores censurados por los valores simulados de la densidad condicional  $f(y_c|y_o, \theta_k, k)$  de  $Y_c$ . Como se muestra en [Anderson \(2003\)](#)

$$Y_c|y_o, \theta_k, k \sim N_T(\mu_c^o + \Sigma_{co}^o(\Sigma_{oo}^o)^{-1}(y_o - \mu_o^o), \Sigma_{cc}^o - \Sigma_{co}^o(\Sigma_{oo}^o)^{-1}\Sigma_{oc}^o)$$

donde

$$\Sigma_{oo}^o = P_o \Sigma P_o^t, \quad \Sigma_{oc}^o = P_o \Sigma P_c^t, \quad \Sigma_{co}^o = P_c \Sigma P_o^t \quad \text{y} \quad \Sigma_{cc}^o = P_c \Sigma P_c^t. \quad (2.4)$$

Se selecciona una matriz de permutación  $P = (P_o, P_c)^t$  tal que  $PY = (Y_c, Y_o)^t$ , donde  $P_o$  y  $P_c$  están definidas como sigue

$$P_{o,kj} = \begin{cases} 1, & y_j = y_o & k = 1, 2, \dots, n. \\ 0, & y_j \neq y_o & j = 1, 2, \dots, n. \end{cases} \quad (2.5)$$

$$P_{c,ij} = \begin{cases} 1, & y_j = y_c & i = 1, 2, \dots, n. \\ 0, & y_j \neq y_c & j = 1, 2, \dots, n. \end{cases} \quad (2.6)$$

Así como también se obtienen los vectores  $\mathbf{X}_c$  y  $\mathbf{X}_o$ , mediante las siguientes expresiones

$$\mathbf{X}_{o,k} = \begin{cases} y_k, & y_k \leq c_k & k = 1, 2, \dots, n. \\ 0, & y_k > c_k & k = 1, 2, \dots, n. \end{cases} \quad (2.7)$$

$$\mathbf{X}_{c,i} = \begin{cases} c_i, & y_i > c_i & i = 1, 2, \dots, n \\ 0, & y_i \leq c_i & i = 1, 2, \dots, n. \end{cases} \quad (2.8)$$

Ahora aplicando la matriz de permutación  $P$  a los datos observados  $\mathbf{X} = (X_1, \dots, X_n)^t$ , se tiene

$$PX = (\mathbf{X}_c, \mathbf{X}_o)^t \stackrel{d}{=} (\mathbf{X}_c, \mathbf{Y}_o)^t, \quad (2.9)$$

donde  $\stackrel{d}{=}$  denota igualdad en distribución.

El algoritmo de [Jung et al. \(2005\)](#), para simular las observaciones censuradas consiste en los siguientes pasos:

1. Construir las matrices de permutación  $P_o$  y  $P_c$  y los vectores  $\mathbf{X}_o$  y  $\mathbf{X}_c$ , mediante las expresiones dadas en (2.5), (2.6), (2.7) y (2.8) respectivamente.
2. Obtener valores iniciales para  $\hat{\mu}_o$ ,  $\hat{\phi}_o$  y  $\hat{\sigma}_o^2$ . Entonces calcular la media y la matriz de

### 2.3. El problema de series de tiempo con puntos de cambio

---

covarianza, mediante las siguientes expresiones,

$$\begin{aligned}\mu^{(o)} &= \hat{\mu}_o \mathbf{1}_n. \\ \Sigma_{ij}^{(o)} &= \frac{\hat{\sigma}_o^2}{(1 - \hat{\phi}_o^2)} \hat{\phi}_o^{|i-j|} \quad i, j = 1, 2, \dots, n.\end{aligned}$$

3. Calcular  $\Sigma_{oo}^o$ ,  $\Sigma_{oc}^o$ ,  $\Sigma_{co}^o$  y  $\Sigma_{cc}^o$  mediante las expresiones dadas en (2.4). Entonces calcular la media y varianza condicional de la parte censurada como sigue,

$$\begin{aligned}\nu^{(o)} &= \mu_c^{(o)} + \Sigma_{co}^{(o)} (\Sigma_{oo}^{(o)})^{-1} (\mathbf{y}_o - \mu_o^{(o)}). \\ \Delta^{(o)} &= \Sigma_{cc}^{(o)} - \Sigma_{co}^{(o)} (\Sigma_{oo}^{(o)})^{-1} \Sigma_{oc}^{(o)}.\end{aligned}$$

4. Simular los valores censurados de la manera siguiente,

$$\mathbf{y}_c \sim N_T(\nu^{(o)}, \Delta^{(o)}).$$

5. Obtener el vector de observaciones completas como sigue,

$$\mathbf{y} = P^{-1}(\mathbf{y}_c, \mathbf{y}_o)^t,$$

con  $P$  dada por  $P = (P_o, P_c)^t$ .

6. Estimar  $\mu$ ,  $\phi$  y  $\sigma^2$ .

### 2.3. El problema de series de tiempo con puntos de cambio

Un modelo de puntos de cambio asume que la estructura del modelo o algunos de los parámetros del modelo presentan cambios en puntos desconocidos en el tiempo.

El problema de series de tiempo con puntos de cambios y observaciones censuradas es un problema de inferencia estadística en el que se desconocen los parámetros del modelo y el número de parámetros. Éste problema se puede formular mediante inferencia conjunta de un indicador  $r$  del modelo y el vector de parámetros  $\theta_r$ , donde el indicador del modelo determina la dimensión  $n_r$  de  $\theta_r$ .

## 2.3. El problema de series de tiempo con puntos de cambio

---

Generalmente los problemas de puntos de cambio se dividen en dos categorías: los problemas de puntos de cambio con *muestras completas*, éste requiere tener el conjunto de observaciones para realizar el análisis acerca de la existencia de posibles puntos de cambio. El problema de detección de puntos de cambio *secuencial* incorpora las observaciones a medida que se van recolectando y se realiza la toma de decisión cada vez que se incluye una nueva observación, este tipo de situación se tiene cuando el interés es detectar rápidamente un punto de cambio sin tener que esperar a que se tenga todo el conjunto de observaciones, e.g., en la estimación de parámetros de procesos (filtrado), segmentación de imágenes y algunos procedimientos de detección en control de calidad. Para más información sobre el problema de puntos de cambio *secuencial*, véase [Lai \(1995\)](#) y [Xia \(2006\)](#).

Algunos métodos han sido desarrollados para resolver el problema de puntos de cambio, e.g., [Green \(1995\)](#) modela la intensidad  $x(t)$  de un proceso Poisson, mediante una función escalonada y un número desconocido de saltos, [Troughton y Godsill \(1997\)](#) resuelven el problema de inferencia Bayesiana en procesos autoregresivos con ordenes desconocidos, Müller (1994) utiliza un procedimiento de estimación de puntos de cambio en la función de riesgo bajo censura desde los puntos de vista paramétrico y no paramétrico. Para una revisión mas amplia del problema de puntos de cambio véase [Bhattacharya \(1994\)](#) y [Csörgo y Horvath \(1998\)](#).

### 2.3.1. Prueba de hipótesis para detección de puntos de cambio

El objetivo en la detección de puntos de cambio es determinar si la serie de tiempo es estadísticamente homogénea, y si no, determinar las localizaciones de los puntos de cambio. Hay varios algoritmos para detectar puntos de cambio. Para una introducción a este tema véanse los trabajos de [Basseville y Nikiforov \(1993\)](#) y [Gupta \(2001\)](#).

Sea  $Y_t$ ,  $t = 1, 2, \dots, n$  una serie de tiempo con función de distribución  $F(Y, \theta_i)$  donde  $\theta_i$  es el vector de parámetros, con un punto de cambio en  $\tau$  (desconocido). El problema de determinar la existencia de un punto de cambio se realiza mediante una prueba de hipótesis como sigue

$$H_0 : \theta_1 = \dots = \theta_n$$

$$H_1 : \text{existe un } \tau, \quad 1 \leq \tau \leq n - 1, \quad \text{tal que } \theta_1 = \dots = \theta_\tau \neq \theta_{\tau+1} \dots = \theta_n.$$

## 2.4. Inferencia bayesiana

---

Si  $\tau$  es conocido, se rechaza  $H_0$  para valores grandes de la razón de verosimilitud

$$\Lambda_\tau = \frac{\sup_{\theta \in \Theta} f(y_1, y_2, \dots, y_\tau; \theta) \sup_{\theta \in \Theta} f(y_{\tau+1}, y_{\tau+2}, \dots, y_n; \theta)}{\sup_{\theta \in \Theta} f(y_1, y_2, \dots, y_n; \theta)}. \quad (2.10)$$

Aquí  $f(\cdot)$  es la función de densidad de las variables aleatorias  $Y_1, Y_2, \dots, Y_n$ . Como se muestra en [Gombay y Horvath \(1994\)](#), si  $\tau$  no se conoce, se rechaza  $H_0$  para valores grandes de

$$Q_n = \max_{1 \leq \tau' \leq n-1} (2 \log \Lambda_{\tau'}). \quad (2.11)$$

### 2.3.2. Estimación de puntos de cambio

Sean  $Y_1, Y_2, \dots, Y_n$  variables aleatorias con función de densidad  $f(\cdot, \theta)$  y función distribución absolutamente continua  $F(\cdot, \theta)$ . La función de densidad  $f(\cdot, \theta)$  es dada como sigue,

$$f(y_i, \theta) = \begin{cases} f(y_i, \theta_0), & \theta_0 \in \Theta, \quad 1 \leq i \leq \tau \\ f(y_i, \theta_1), & \theta_1 \in \Theta, \quad \tau + 1 \leq i \leq n. \end{cases}$$

Luego se estima el punto de cambio  $\tau$  con  $\theta_0$  y  $\theta_1$  desconocidos. La función de verosimilitud marginal de las variables aleatorias  $Y_1, Y_2, \dots, Y_n$  está dada por

$$\tilde{L}(\tau) = f(y_1, y_2, \dots, y_\tau; \hat{\theta}_0) f(y_{\tau+1}, y_{\tau+2}, \dots, y_n; \hat{\theta}_1), \quad (2.12)$$

donde  $\hat{\theta}_0$  es el estimador condicional por máxima verosimilitud basado en las primeras  $\tau$  observaciones y  $\hat{\theta}_1$  se obtiene de las  $n - \tau$  restantes. El estimador por máxima verosimilitud está dado por

$$\tilde{V}_n = \max_{1 \leq j \leq n-1} \tilde{L}(j). \quad (2.13)$$

## 2.4. Inferencia bayesiana

En esta sección se presentan algunos conceptos de inferencia bayesiana. Algunos libros sobre el tema son, [Smith y Roberts \(1993\)](#), [Robert y Casella \(2002\)](#) y [Gelman \*et al.\* \(2003\)](#).

Un modelo bayesiano está basado en la especificación de un modelo de probabilidad para

## 2.4. Inferencia bayesiana

---

los datos observados  $y$ , dado un vector de parámetros desconocidos  $\theta$ , encabezado por la función de verosimilitud  $L(y|\theta)$ . Se asume que  $\theta$  es aleatorio y tiene una distribución inicial denotada por  $p(\theta)$ . Cualquier inferencia sobre  $\theta$  está basada en su distribución final, la cual por el teorema de Bayes está dada por

$$p(\theta|y) = \frac{L(y|\theta)p(\theta)}{p(y)}.$$

donde  $L(y|\theta)$  es la función de verosimilitud,  $p(\theta)$  es la distribución inicial y  $p(\theta|y)$  la distribución final. Los parámetros en la distribución inicial se le conoce como hiperparámetros y se consideran conocidos.

### Distribución inicial

La distribución inicial cumple un papel importante en el análisis bayesiano ya que mide el grado de conocimiento inicial que se tiene de los parámetros en estudio. Si bien su influencia disminuye a medida que mas información muestral está disponible, el uso de una u otra distribución inicial determinará ciertas diferencias en la distribución final.

A continuación se presentan dos métodos para especificar la distribución inicial.

**Distribución conjugada** Si existe una familia conjugada, se puede elegir una distribución dentro de esta familia por especificación de, por ejemplo, los primeros momentos de la distribución de  $\theta$  o de la distribución predictiva de  $Y$ .

**Definición 2.7** Sea  $\mathfrak{S}$  una clase de distribuciones muestrales  $f(x|\theta)$  y denótese como  $P$  a la clase de distribuciones iniciales  $f(\theta)$  para  $\theta$ , entonces  $f$  es conjugada para  $\mathfrak{S}$  si  $f(\theta|x) \in P$ ,  $\forall f(\cdot|\theta) \in \mathfrak{S}$  y  $f(\cdot) \in P$ .

A continuación se presentan algunas distribuciones iniciales que se utilizan en el Capítulo 4.

Sea  $y_1, y_2, \dots, y_n$  una realización de la serie de tiempo autoregresiva de orden uno definida en (2.2), que tiene distribución normal multivariada con medias  $\mu$  y matriz de covarianzas  $\Sigma_{ij} = \left[ \frac{\sigma^2}{(1-\rho^2)} \rho^{|i-j|} \right]$ . La función de verosimilitud está dada por

$$L(y_1, y_2, \dots, y_n | \mu, \sigma^2, \phi) = \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu - \phi(y_{j-1} - \mu))^2}.$$



## 2.4. Inferencia bayesiana

---

Si para  $\mu$  se considera la distribución inicial

$$\mu \sim N(\mu_o, \sigma). \quad (2.14)$$

con  $-\infty < \mu < \infty$  y  $\mu_o, \sigma_o^2$  son hiperparámetros,  $\sigma_o^2 > 0$  y  $-\infty < \mu_o < \infty$ , entonces la densidad final de  $\mu$  asumiendo  $\phi, \sigma^2$  son conocidos, es como sigue

$$\mu | \phi, \sigma^2, y_1, y_2, \dots, y_n \sim N \left( \frac{\sum_{i=1}^n (y_i - \phi y_{i-1})(1 - \phi)^2 \sigma_o^2 + \mu_o \sigma^2}{n(1 - \phi)^2 \sigma_o^2 + \sigma^2}, \frac{\sigma_o^2 \sigma^2}{n(1 - \phi)^2 \sigma_o^2 + \sigma^2} \right).$$

Así la distribución normal es una familia conjugada para la familia normal.

Ahora considérese que el parámetro es  $\sigma^2$ , con  $\mu$  y  $\phi$  conocidos, y supóngase que la distribución inicial de  $\sigma^2 \sim GI(\alpha_o, \beta_o)$  con  $\sigma^2$  y  $\alpha_o$  y  $\beta_o$  son hiperparámetros positivos, i.e.,  $f(\sigma^2 | \alpha_o, \beta_o) = \frac{\beta_o^{\alpha_o}}{\Gamma(\alpha_o)} (\sigma^2)^{-\alpha_o-1} e^{-\frac{\beta_o}{\sigma^2}}$ , entonces se obtiene que la distribución final de  $\sigma^2$  está dada por

$$\sigma^2 | \mu, \phi, y_1, \dots, y_n \sim GI \left( \frac{1}{2} n + \alpha_o, \frac{1}{2} \sum_{j=1}^n (y_j - \mu - \phi(y_{j-1} - \mu))^2 + \beta_o \right).$$

Así la distribución gama inversa es una familia conjugada para la gama inversa.

### Distribución inicial no informativa

Cuando no se conoce nada sobre los parámetros, la selección de una distribución inicial adecuada adquiere una connotación especial pues será necesario elegir una inicial que no influya sobre ninguno de los posibles valores de los parámetros en cuestión. Jeffreys (1961) sugiere que si no se conoce nada sobre un parámetro  $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ , entonces la opinión acerca de  $\theta$  dadas las observaciones  $y_1, y_2, \dots, y_n$  debe ser la misma que el de una parametrización para  $\theta$  y una inicial sería:

$$f(\theta) = [\det I(\theta)]^{0.5}. \quad (2.15)$$

donde  $I(\theta)$  es la matriz de información de Fisher, cuya entrada  $(i, j)$  es,  $I_{ij} = -E_{\theta} \left[ \frac{\partial^2 \text{Ln} f(y|\theta)}{\partial \theta_i \partial \theta_j} \right]$ . Como un ejemplo, si  $Y|\theta \sim \text{Bin}(n, \theta)$ , como se puede verificar una distribución inicial no informativa para  $\theta$  es  $\theta \sim \text{Beta}(\frac{1}{2}, \frac{1}{2})$  y si  $Y|\mu, \sigma^2 \sim N(\mu, \sigma^2)$ , como se puede verificar una distribución inicial no informativa para  $\theta := (\mu, \sigma^2)$  es  $f(\mu, \sigma^2) \propto \frac{1}{\sigma^2}$ .

## 2.5. Cadenas de Markov Monte Carlo

Las técnicas de cadenas de Markov Monte Carlo están diseñadas con la finalidad de poder estudiar las características de distribuciones complejas. Básicamente consiste en lo siguiente: dada una distribución  $f(\theta)$  completamente conocida, salvo quizá por su constante de proporcionalidad, se trata de generar una o varias realizaciones de una cadena de Markov cuya distribución estacionaria sea  $f(\theta)$ . Para asegurar convergencia a la distribución estacionaria, se requiere que la cadena de Markov sea irreducible, es decir, cada estado del sistema puede ser alcanzado independientemente del estado inicial y además aperiódica, i.e., la cadena no quede atrapada en un ciclo. Si se cumplen estas dos condiciones, se dice que el proceso es ergódico. Para una información más detallada sobre MCMC veáanse: [Chib y Greeberg \(1995\)](#), [Gilks et al. \(1996\)](#), [Gamerman \(1997\)](#) y [Robert y Casella \(2002\)](#).

Sea  $T \subset \mathfrak{R}$  y  $(\Omega, \mathfrak{S}, P)$  un espacio de probabilidad. Un proceso aleatorio es una función  $X : T \times \Omega \rightarrow \mathfrak{R}$  tal que para cada  $t \in T$ ,  $X(t, \cdot)$  es una variable aleatoria. Si fijamos  $w \in \Omega$  obtenemos una función  $X(\cdot, w) : T \rightarrow \mathfrak{R}$  que se conoce como una trayectoria del proceso. A los valores del proceso se le llama espacio de estados y lo denotamos en lo sucesivo con  $S$ .

**Definición 2.8** *Un kernel de transición es una función  $P$  definida en  $\mathfrak{X} \times B(\mathfrak{X})$  tal que*

$$\forall x \in \mathfrak{X}, P(x, \cdot) \text{ es una medida de probabilidad ; } y \forall A \in B(\mathfrak{X}), P(\cdot, A) \text{ es medible.}$$

Si  $\mathfrak{X}$  es discreto, el kernel de transición es una matriz de permutación  $P$  con elementos  $P_{xy} = P(X_n = y | X_{n-1} = x)$ ,  $x, y \in \mathfrak{X}$ . Si  $\mathfrak{X}$  es continuo, el kernel de transición denota la densidad condicional  $P(x, x')$  de la transición  $P(x, \cdot)$ ; esto es,  $P(X \in A | x) = \int_A P(x, x') dx'$ .

**Definición 2.9** *Una cadena de Markov de tiempo discreto es un proceso de Markov cuyo espacio de estados es un conjunto finito o numerable y cuyo conjunto de índices es numerable, usualmente  $T = 0, 1, 2, \dots$  y cumple la propiedad Markoviana*

$$P(X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i_n) = P(X_{n+1} = j | X_n = i_n) \quad (2.16)$$

para todo  $n$  y cualesquiera estados  $i_0, i_1, \dots, i_n, j$ .

El principio detrás de MCMC es el teorema de ergodicidad que se enuncia a continuación.

## 2.5. Cadenas de Markov Monte Carlo

---

**Teorema 2.1** Sea  $X_1, \dots, X_n$  una cadena de Markov, si  $f$  es una función integrable, entonces

$$\frac{1}{M+1} \sum_{i=0}^M f(X_i) \xrightarrow{a.s.} E_\pi(f(X)), \quad \text{cuando } M \rightarrow \infty. \quad (2.17)$$

donde *a.s.* significa convergencia casi segura.

Sea  $f$  la distribución de interés definida sobre el espacio parametral  $C$  y  $P(x, dx')$  el kernel de transición. Se dice que una cadena Markov es reversible si se cumple la condición de probabilidad de equilibrio siguiente

$$f(x)P(x, dx') = f(x')P(x', dx), \quad \forall x, x' \in \Omega. \quad (2.18)$$

o si

$$\int_A \int_B f(x)P(x, dx') = \int_B \int_A f(dx')P(x', dx) \quad \forall A, B \in \mathbb{A} \quad (2.19)$$

donde  $\mathbb{A}$  es el álgebra de Borel en  $\mathfrak{R}$ .

Una cadena de Markov es invariante si es reversible. Una cadena de Markov es  $f$  irreducible si es capaz de pasar por todos los eventos que tienen probabilidad positiva bajo  $f$ . Mas formalmente,  $f$  irreducible significa que para cada estado inicial  $x \in \Omega$  y para cada evento  $A \in \mathbb{A}$  con  $f(A) > 0$ , existe un tiempo  $t := t(x, A) > 0$  tal que la probabilidad  $P(x, A)$  de visitar el conjunto  $A$ , en el tiempo  $t$ , dado el punto inicial  $x$ , es positiva.

Una cadena de Markov  $f$  irreducible es aperiódica si ésta no se mueve através del espacio muestral de manera cíclica. Un  $d$  ciclo es una secuencia de  $d$  conjuntos disjuntos no vacíos  $A_0, A_1, \dots, A_{d-1}$ , através de los cuales la cadena siempre pasa en algún orden y de los que no puede escapar debido a que la probabilidad de moverse al siguiente conjunto es igual a uno.

Dos métodos populares para generar muestras de las distribuciones finales, que se describen en las dos secciones que siguen, son el muestreo de Gibbs y el algoritmo de Metrópolis-Hasting. El muestreo de Gibbs no es aplicable cuando la dimensión del espacio parametral no es fija, una generalización del algoritmo de Metrópolis-Hastings, que se describe en la Sección 2.6, es el algoritmo de cadenas de Markov Monte Carlo con saltos reversibles y permite saltar entre modelos con espacios parametrales de diferente dimensión.

## 2.5. Cadenas de Markov Monte Carlo

---

### 2.5.1. Algoritmo de Gibbs

Entre las técnicas MCMC el algoritmo de Gibbs es uno de los métodos más conocido y utilizado. El artículo de [Casella y George \(1992\)](#) proporciona una introducción de este método, y en [Gelfand \*et al.\* \(1990\)](#) se presentan numerosas aplicaciones del mismo.

Sea  $f(\theta)$  una densidad de la que no se pueden generar valores, completamente conocida excepto quizá por la constante de integración, pero demasiado compleja como para calcular sus características: momentos, marginales, predictivas, etc.. Para el estudio de esas características es útil el algoritmo muestreador de Gibbs, si el vector paramétrico puede descomponerse en  $k \geq 2$  subvectores,  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ , de manera que para cada una de las  $k$  distribuciones condicionales  $f(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$  se pueden generar valores de dichas distribuciones sin excesivo costo computacional.

El algoritmo muestreador de Gibbs es el siguiente. Supóngase que partiendo de un valor inicial  $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$  se han generado  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(i-1)}$ . Entonces para generar  $\theta^{(i)} = (\theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_k^{(i)})$  procédase del modo siguiente

- 1) Generar  $\theta_1^{(i)}$  de la distribución condicional  $f_1(\theta_1 | \theta_2^{(i-1)}, \dots, \theta_k^{(i-1)})$
- 2) Generar  $\theta_2^{(i)}$  de la distribución condicional  $f_2(\theta_2 | \theta_1^{(i)}, \theta_3^{(i-1)}, \dots, \theta_k^{(i-1)})$
- ⋮
- k) Generar  $\theta_k^{(i)}$  de la distribución condicional  $f_k(\theta_k | \theta_1^{(i)}, \dots, \theta_{k-1}^{(i)})$ .

Los pasos 1 a  $k$  se repiten hasta que se obtenga convergencia en la cadena de Markov. El resultado es una sucesión de vectores simulados  $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots$  cuya distribución estacionaria es  $f(\theta)$ .

### 2.5.2. Algoritmo de Metrópolis-Hastings

De manera analoga que el muestreador de Gibbs, el algoritmo de Metropolis-Hastings consiste en generar valores de una cadena de Markov cuya distribución estacionaria es la distribución objetivo  $f(\theta)$ . El algoritmo de Metropolis-Hastings proporciona un método para construir innumerables cadenas de Markov con esa propiedad, lo que nos permitirá buscar entre ellas una converja rápidamente a la distribución estacionaria y no posea demasiada autocorrelación.

## 2.6. Cadenas de Markov Monte Carlo con saltos reversibles

---

Básicamente se trata de utilizar una cadena de Markov auxiliar, para la que dispongamos de un generador eficiente de su probabilidad de transición  $q(\theta, \phi)$  (que representa la distribución de probabilidades de pasar en una etapa del punto  $\theta$  al punto  $\phi$ ), y añadir en cada etapa un mecanismo de aceptación-rechazo con probabilidad de aceptación dada por:

$$\alpha(\theta, \phi) = \min \left( 1, \frac{q(\theta, \phi)f(\phi)}{q(\phi, \theta)f(\theta)} \right) \quad (2.20)$$

de manera que si en la etapa  $t$  el valor obtenido es  $\theta^{(t)}$ , en la etapa siguiente se genera  $\phi$  a partir de  $q(\theta, \phi)$  y, con probabilidad  $\alpha(\theta, \phi)$ , definir  $\theta^{(t+1)} = \phi$  y  $\theta^{(t+1)} = \theta^{(t)}$  de otro modo. El algoritmo de Metrópolis-Hastings es como sigue:

1. Seleccionar un punto inicial  $\theta^0$  en el soporte de  $f(\cdot)$
2. Generar  $\phi$  de  $q(\theta^{(t-1)}, \phi)$ , dicha distribución debe ser simétrica en sus argumentos, es decir,  $q(\phi, \theta^{(t-1)}) = q(\theta^{(t-1)}, \phi)$  para todo  $\theta^{(t-1)}, \phi, t$ .
3. Calcular el cociente  $\alpha(\theta, \phi) = \min \left( 1, \frac{q(\theta^{(t-1)}, \phi)f(\phi)}{q(\phi, \theta^{(t-1)})f(\theta^{(t-1)})} \right)$ .
4. Si  $\alpha \geq 1$  entonces  $\theta^{(t)} = \phi$ .
5. Si  $\alpha < 1$  entonces generar  $u \sim Unif(0, 1)$ , si  $u < \alpha$  entonces  $\theta^{(t)} = \phi$ .

Como la distribución objetivo  $f(\theta)$  sólo aparece en el algoritmo a través de un cociente, no es necesario conocer su constante de integración.

Los libros de [Carlin y Louis \(2000\)](#) y [Gelman \*et al.\* \(2003\)](#) tienen abundantes ejemplos en los que se utiliza este algoritmo.

## 2.6. Cadenas de Markov Monte Carlo con saltos reversibles

El algoritmo de cadenas de Markov Monte Carlo con saltos reversibles, es una generalización del algoritmo Metropolis-Hasting, desarrollado por [Green \(1995\)](#) y consiste en crear una cadena de Markov irreducible, aperiódica que alterna saltos entre distintos modelos con espacios parametrales de diferente dimensión, mientras se satisface la condición de probabilidad

## 2.6. Cadenas de Markov Monte Carlo con saltos reversibles

---

de equilibrio, asegurando la convergencia a la distribución límite correcta. Este algoritmo ha sido utilizado ampliamente en la literatura, e.g., [Richardson y Green \(1997\)](#) aplican este algoritmo a distribuciones mezclas con un número desconocido de componentes. Para esto, implementan dos tipos de movimientos: uno en el que se crea un componente nuevo, o se elimina uno existente, y otro en el que un componente es dividido en dos, o dos componentes son combinados en uno. [Troughton y Godsill \(1997\)](#) resuelven el problema de inferencia bayesiana en procesos autoregresivos de orden desconocido. A fin de obtener eficiencia en el algoritmo RJMCMC, propusieron un espacio de movimientos de modelos de las densidades condicionales para los parámetros autoregresivos, los que se obtienen analíticamente.

### Distribución propuesta

Un aspecto importante en la implementación del algoritmo RJMCMC es la especificación de una densidad propuesta  $q$ , de la que se generan vectores de valores para los parámetros adicionales cuando se pretende saltar entre modelos con espacios parametrales de diferente dimensión y una vez que se tienen valores generados se puede construir una función biyectiva entre los espacios parametrales de los respectivos modelos. Una práctica común en la elección de una densidad propuesta consiste en realizar corridas cortas del algoritmo RJMCMC cambiando cada vez de densidad propuesta o cierta característica de ésta.

La eficiencia del algoritmo RJMCMC es sensible a la elección de la densidad propuesta y por consiguiente el número de iteraciones necesarias para alcanzar la convergencia depende de tal elección. Hay algunas sugerencias para seleccionar la densidad propuesta, e.g., [Brooks \*et al.\* \(2003\)](#) proporcionan algunas sugerencias de como seleccionar una densidad propuesta eficiente  $q$ , además [Ehlers y Brooks \(2002\)](#) extienden el trabajo de [Brooks \*et al.\* \(2003\)](#).

La función de salto  $g$  se puede definir de forma determinista y arbitraria, sujeta únicamente a que ésta y su inversa sean derivables.

El algoritmo RJMCMC permite seleccionar muestras de la distribución de interés definida sobre la unión de subespacios  $C_r$  de diferente dimensión  $n_r$ . El algoritmo RJMCMC es una generalización de las técnicas MCMC en el que la dimensión del espacio parametral puede cambiar de una iteración a otra. Sean  $M_r$ ,  $r = 1, 2, \dots, k_{max}$  los modelos candidatos y  $\theta_r$  el vector de parámetros desconocidos de dimensión  $n_r$  del modelo  $M_r$ , los que pueden variar de un modelo a otro, como las encontradas en el problema de selección de variables en regresión, selección de modelos y modelos de series de tiempo autoregresivos de orden desconocido. Esto es posible cubriendo el espacio de movimientos entre modelos especificando

## 2.6. Cadenas de Markov Monte Carlo con saltos reversibles

---

una densidad propuesta, construyendo una función de salto invertible entre los subespacios y calculando la probabilidad de aceptación que satisface la condición de probabilidad de equilibrio. Supóngase que se desea mover del modelo actual  $M_i$  a algún modelo  $M_j$ , entonces el algoritmo RJMCMC consiste en los siguientes pasos:

1. Seleccionar un modelo  $M_j$  de acuerdo a la distribución de probabilidades  $r_{ij}(\theta_i)$ ,  $i = 1, 2, \dots, k_{max}$ , donde  $r_{ij}(\theta_i)$  es la probabilidad de moverse del modelo  $M_i$  al modelo  $M_j$
2. Generar un vector de valores  $\mathbf{u}$  de la densidad propuesta específica  $q(\mathbf{u})$ .
3. Construir una función biyectiva  $g$  que transforma los vectores de parámetros  $g(\theta_i, \mathbf{u}) = (\theta_j, \mathbf{u}')$ , tal que  $dim(\theta_i) + dim(\mathbf{u}) = dim(\theta_j) + dim(\mathbf{u}')$ .
4. Aceptar el movimiento propuesto con probabilidad

$$\alpha = \min \left( 1, \frac{f(M_j, \theta_j | y) r_{ji}(\theta_j) q'(\mathbf{u}')}{f(M_i, \theta_i | y) r_{ij}(\theta_i) q(\mathbf{u})} \left| \frac{\partial g(\theta_i, \mathbf{u})}{\partial (\theta_i, \mathbf{u})} \right| \right). \quad (2.21)$$

Donde los vectores  $\mathbf{u}$  y  $\mathbf{u}'$  son utilizados para hacer los espacios parametrales de  $M_i$  y  $M_j$  de igual dimensión y  $q(\mathbf{u}')$  es la densidad propuesta de  $\mathbf{u}'$  independiente de  $\theta_j$  (dimensión  $m_j$ ) y  $|\cdot|$  es el determinante del jacobiano.

5. Si el movimiento no es aceptado, permanecer en el modelo actual  $M_i$ .

De manera inversa el movimiento de  $M_j$  a  $M_i$  tiene probabilidad de aceptación  $\alpha'$  como sigue:

$$\alpha' = \min \left( 1, \frac{f(M_i, \theta_i | y) r_{ij}(\theta_i) q(\mathbf{u})}{f(M_j, \theta_j | y) r_{ji}(\theta_j) q'(\mathbf{u}')} \left| \frac{\partial g^{-1}(\theta_i, \mathbf{u})}{\partial (\theta_j, \mathbf{u}')} \right| \right). \quad (2.22)$$

En la implementación del algoritmo RJMCMC, frecuentemente se tienen algunas dificultades, en particular permitiendo que la cadena salte de un modelo a otro. Hay dos tipos de elección a realizar cuando se construye movimientos entre modelos de diferente dimensión: la densidad propuesta y la función salto.

## 2.7. Método para monitorear convergencia en RJMCMC

Una manera de monitorear la convergencia de una cadena de Markov Monte Carlo con saltos reversibles, consiste en generar varias cadenas con la misma longitud, pero con diferentes puntos de inicio, y checar en que momento las cadenas han olvidado sus puntos de inicio. Hay varias formas posibles de comparar cadenas paralelas, a continuación se presenta el método de convergencia de [Castelloe \(1998\)](#), éste método se utiliza en el Capítulo 4 para diagnosticar la convergencia de la cadena de Markov en la implementación de RJMCMC al problema de series de tiempo con múltiples puntos de cambio y con observaciones censuradas.

### Método de [Castelloe \(1998\)](#) para monitorear convergencia

El método de [Castelloe \(1998\)](#) para monitorear la convergencia de una cadena de Markov Monte Carlo con salto reversibles es adecuado en situaciones en las que los diferentes espacios parametrales de los modelos son indexados por algún parámetro en la cadena. Este consiste en monitorear las siguientes fuentes de variación: variación entre las cadenas, interacción entre modelos-cadenas y diferencias significativas en las frecuencias de visitas de modelos de una cadena a otra. Cualquiera de esas tres condiciones pueden indicar que las cadenas no provienen de la misma distribución estacionaria y por consiguiente que no se ha alcanzado convergencia a la distribución estacionaria.

Sean  $\theta_1, \theta_2, \dots, \theta_p$  los parámetros, supóngase que se tienen  $C$  cadenas de Markov paralelas, cada una de longitud  $qb$ , con  $\theta_i^{(j)}$ ,  $i = 1, 2, \dots, C$   $j = qb + 1, qb + 2, \dots, 2qb$ , para algún  $q$  y para un tamaño base  $b$ . A continuación se presenta la notación que se utiliza en la



## 2.7. Método para monitorear convergencia en RJMCMC

---

implementación del diagnóstico de convergencia:

$\theta$  : vector de parámetros con la misma interpretación en los modelos.

$C$  : número de cadenas.

$T$  : tamaño ( tamaño de la cadena).

$M$  : número de modelos distintos visitados por la cadena.

$\theta_{cm}^r$  : valor de  $\theta$  para  $r^{th}$  ocurrencia del modelo  $m$  en la cadena  $c$ .

$R_{cm}$  : número de de veces que el modelo  $m$  aparece en la cadena  $c$ .

$$R_{.m} = \sum_{c=1}^C R_{cm}.$$

$$\bar{\theta}_{cm} = \frac{1}{R_{cm}} \sum_{r=1}^{R_{cm}} \theta_{cm}^r.$$

$$\bar{\theta}_{.c} = \frac{1}{T} \sum_{m=1}^M \sum_{r=1}^{R_{cm}} \theta_{cm}^r.$$

$$\bar{\theta}_{.m} = \frac{1}{R_{.m}} \sum_{c=1}^C \sum_{r=1}^{R_{cm}} \theta_{cm}^r.$$

$$\bar{\theta}_{..} = \frac{1}{CT} \sum_{c=1}^C \sum_{m=1}^M \sum_{r=1}^{R_{cm}} \theta_{cm}^r.$$

El método de [Castelloe \(1998\)](#) para monitorear convergencia en RJMCMC se basa en la estimación de las siguientes fuentes de variación.

$$V = \frac{1}{CT - 1} \sum_{c=1}^C \sum_{m=1}^M \sum_{r=1}^{R_{cm}} (\theta_{cm}^r - \bar{\theta}_{..})^2,$$

$$W_c = \frac{1}{C(T - 1)} \sum_{c=1}^C \sum_{m=1}^M \sum_{r=1}^{R_{cm}} (\theta_{cm}^r - \bar{\theta}_{.c})^2,$$

$$W_m = \frac{1}{CT - M} \sum_{c=1}^C \sum_{m=1}^M \sum_{r=1}^{R_{cm}} (\theta_{cm}^r - \bar{\theta}_{.m})^2,$$

$$W_m W_c = \frac{1}{C(T - M)} \sum_{c=1}^C \sum_{m=1}^M \sum_{r=1}^{R_{cm}} (\theta_{cm}^r - \bar{\theta}_{cm})^2.$$

Donde  $V$  es la variación total,  $W_c$  es la variación dentro de las cadenas,  $W_m$  es la variación dentro de los modelos y  $W_m W_c$  es la variación entre cadenas-modelos. Los estadísticos para el

## 2.7. Método para monitorear convergencia en RJMCMC

---

diagnóstico de convergencia son:  $R_1 = \frac{V}{W_c}$  y  $R_2 = \frac{W_m}{W_m W_c}$ . Si las cadenas convergen, entonces, las estimaciones de  $V$  y  $W_c$  son parecidas, así como también las estimaciones de  $W_m$  y  $W_m W_c$ . Valores de  $R_1$  y  $R_2$  mayores que 1 indican convergencia.

# Capítulo 3

## SERIES DE TIEMPO CON PC Y CEN

En este capítulo se presenta un modelo bayesiano para series de tiempo con un número desconocido de puntos de cambio y censura en algunas de las observaciones. Éste modelo se aplica a 3 conjuntos de datos simulados. Como se vió en la Sección 2.6 del Capítulo 2, el algoritmo RJMCMC es el adecuado para generar muestras de la distribución final del número de puntos de cambio y las localizaciones de los puntos de cambio. Este capítulo está estructurado del modo siguiente: en la Sección 3.1 se propone un modelo bayesiano para series de tiempo que presentan cambios en sus estructuras y censura en algunas observaciones, en la Sección 3.2 se implementa el método de [Jung \*et al.\* \(2005\)](#) para resolver el problema de censura, así como el algoritmo RJMCMC para estimar el número y las localizaciones de los puntos de cambio de una serie de tiempo. Finalmente, en la Sección 3.3 se presentan los resultados de analizar un conjunto de datos simulados con 0%, 10% y 40% de las observaciones censuradas

### 3.1. Modelo bayesiano para el problema de series de tiempo con PC y CEN

El problema de series de tiempo con puntos de cambio y el problema de series de tiempo con censura en algunas de las observaciones han sido objeto de estudio de varios autores,

### 3.1. Modelo bayesiano para el problema de series de tiempo con PC y CEN

---

e.g., [Robinson \(1980\)](#), [Green \(1995\)](#), [Jung \*et al.\* \(2005\)](#) y [Davis \*et al.\* \(2006\)](#). El problema de puntos de cambio y censura ha sido objeto de estudio únicamente en el área de confiabilidad aplicado a tiempos de vida independientes y a la función de supervivencia, e.g., [Shuangquan \(1998\)](#) y [Gijbels y Gurler \(2003\)](#).

El modelo bayesiano para series de tiempo con puntos de cambio y con observaciones censuradas que se propone en esta sección es como sigue. Sea  $x_1, x_2, \dots, x_n$  la realización de una serie de tiempo con  $k$  puntos de cambio en las localizaciones  $\tau_1, \tau_2, \dots, \tau_k$ , donde  $k$  y  $\tau_1, \tau_2, \dots, \tau_k$  son desconocidos. Además, se asume que algunas observaciones presentan censura. Por conveniencia se considera censura por la derecha en  $c_t, t = 1, 2, \dots, n$ , es decir, en lugar de observar  $x_t$ , se tiene  $y_t := \min(x_t, c_t)$ . Por conveniencia, se define  $\tau_0 := 0$  y  $\tau_{k+1} := n$ . En cada segmento, condicionado a los parámetros, se asume un proceso autoregresivo de orden 1, es decir,

$$\begin{aligned} X_t &= \mu_i + \phi_i(X_{t-1} - \mu_i) + \epsilon_t, & \tau_{i-1} + 1 \leq t \leq \tau_i, \\ & & i = 1, 2, \dots, k + 1, \end{aligned} \quad (3.1)$$

donde  $\epsilon_t \stackrel{iid}{\sim} N(0, \sigma_i^2)$ . Aquí  $k, \tau_1, \tau_2, \dots, \tau_k, (0 < \tau_1 < \tau_2 < \dots < \tau_k < n), \mu_1, \mu_2, \dots, \mu_{k+1}, \sigma_1^2, \sigma_2^2, \dots, \sigma_{k+1}^2, \phi_1, \phi_2, \dots, \phi_{k+1}$  son parámetros. Las distribuciones iniciales para el número de puntos de cambio y sus localizaciones están dadas por,

$$\begin{aligned} K &\sim U(0, 1, 2, \dots, k_{max}), \\ f(\tau_i | \tau_1, \tau_2, \dots, \tau_{i-1}, k) &\sim U(\tau_{i-1} + 1, \dots, n - 1), \quad i = 1, 2, \dots, k, \end{aligned}$$

donde  $k_{max}$  es el máximo número de puntos de cambio que se permite en el modelo. Para las medias  $\mu_i$  y las varianzas  $\sigma_i^2, i = 1, 2, \dots, k + 1$ , se consideran distribuciones iniciales conjugadas, i.e.,

$$\mu_i \sim N(\mu_0, \sigma_0^2), \quad (3.2)$$

$$\sigma_i^2 \sim Ig(\alpha_0, \beta_0), \quad (3.3)$$

donde  $\mu_0, \sigma_0^2, \alpha_0$  y  $\beta_0$  son hiperparámetros. Para asegurar estacionariedad en cada segmento la distribución inicial que se considera para  $\phi_i, i = 1, 2, \dots, k + 1$ , es

$$\phi_i \sim U(-1, 1).$$

### 3.2. Implementación del algoritmo RJMCMC al problema de PC y CEN

---

Inferencia bayesiana sobre  $K$  y  $\theta_k = (\tau_1, \tau_2, \dots, \tau_k, \mu_1, \mu_2, \dots, \mu_{k+1}, \sigma_1^2, \sigma_2^2, \dots, \sigma_{k+1}^2, \phi_1, \phi_2, \dots, \phi_{k+1})$ , se basa en la distribución final  $f(y_c, \theta_k, k | y_o)$ , donde  $y_o := \{x_i | x_i \leq c_i, 1 = 1, 2, \dots, n\}$  y  $y_c := \{c_i | x_i > c, i = 1, 2, \dots, n\}$ , la cual se puede factorizar como sigue,

$$\begin{aligned} f(y_c, \theta_k, k | y_o) &= f(y_c | y_o, \theta_k, k) f(\theta_k, k | y_o) \\ &\propto f(y_c | y_o, \theta_k, k) f(y_o | \theta_k, k) f(\theta_k, k) \\ &\propto f(y | \theta_k, k) f(\theta_k, k), \end{aligned}$$

### 3.2. Implementación del algoritmo RJMCMC al problema de PC y CEN

En esta sección se presenta en detalle la implementación del algoritmo RJMCMC para obtener una muestra de la distribución final del número de puntos de cambio, las localizaciones de los puntos de cambio, las medias, coeficientes autorregresivos y varianzas del modelo bayesiano que se presenta en la sección 3.1. En este caso, se proponen los cuatro tipos de movimientos que siguen: *movimiento tipo 1*) generar un nuevo punto de cambio; *movimiento tipo 2*) eliminar un punto de cambio; *movimiento tipo 3*) actualizar los parámetros autoregresivos y *movimiento tipo 4*) actualizar las observaciones censuradas. En cada iteración se selecciona aleatoriamente uno de estos 4 tipos de movimientos con la misma probabilidad (i.e.,  $\frac{1}{4}$ ). Los movimientos 1 y 2 involucran un cambio de dimensión. En el movimiento tipo 3 se utiliza el algoritmo muestreador de Gibbs y en el movimiento tipo 4 se utiliza el método de [Jung et al. \(2005\)](#).

**Movimiento tipo 1:** Generar un nuevo punto de cambio

Cuando hay un nuevo punto de cambio, se propone pasar de un modelo  $M_i$  con  $k$  puntos de cambio a un modelo  $M_j$  con  $k + 1$  puntos de cambio. Siguiendo el procedimiento propuesto por [Green \(1995\)](#), del modelo con  $k$  puntos de cambio se selecciona aleatoriamente uno de los  $k + 1$  segmentos y después se selecciona aleatoriamente un punto en este segmento. Sea  $i_s$  el segmento seleccionado y  $\tau'_{i_s}$  el punto seleccionado en este segmento. Para decidir si se acepta o no el modelo con  $k + 1$  puntos de cambio se siguen los siguientes 4 pasos.

1. Generar 3 realizaciones independientes  $u_1, u_2$  y  $u_3$  de la densidad normal estándar

### 3.2. Implementación del algoritmo RJMCMC al problema de PC y CEN

---

truncada en el intervalo  $(-1, 1)$ , i.e.,

$$u_i \sim N_{(-1,1)}(0, 1), \quad i = 1, 2, 3. \quad (3.4)$$

2. Denótense por  $\mu_{i_s,1}$ ,  $\mu_{i_s,2}$ ,  $\sigma_{i_s,1}^2$ ,  $\sigma_{i_s,2}^2$  los parámetros de los segmentos que se obtienen al dividir el segmento  $i_s$  del modelo  $M_i$ . Sea  $g$  la función salto que sigue:

$$\begin{aligned} g(\tau_{i_s} - \tau_{i_s-1}, \mu_{i_s}, \sigma_{i_s}^2, u_1, u_2, u_3) &= (\tau'_{i_s} - \tau_{i_s-1}, \tau_{i_s} - \tau'_{i_s}, \mu_{i_s,1}, \mu_{i_s,2}, \sigma_{i_s,1}^2, \sigma_{i_s,2}^2) \quad (3.5) \\ (\tau'_{i_s} - \tau_{i_s-1}) &= (\tau_{i_s} - \tau_{i_s-1})u_1, \\ (\tau_{i_s} - \tau'_{i_s}) &= (\tau_{i_s} - \tau_{i_s-1})(1 - u_1), \\ \mu_{i_s,1} &= \mu_{i_s} - u_2\sigma_i \sqrt{\frac{(\tau_{i_s} - \tau'_{i_s})}{(\tau_{i_s} - \tau_{i_s-1})}}, \\ \mu_{i_s,2} &= \mu_{i_s} - u_2\sigma_i \sqrt{\frac{(\tau_{i_s,1} - \tau_{i_s-1})}{(\tau_{i_s} - \tau'_{i_s})}}, \\ \sigma_{i_s,1}^2 &= u_3 (1 - u_2^2) \sigma_i^2 \frac{(\tau_{i_s} - \tau_{i_s-1})}{(\tau'_{i_s} - \tau_{i_s-1})}, \\ \sigma_{i_s,2}^2 &= (1 - u_3) (1 - u_2^2) \sigma_i^2 \frac{(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s} - \tau'_{i_s})}. \end{aligned}$$

Para proponer esta función salto se ha adoptado la estrategia de momentos en la que se asigna medias y varianzas a los nuevos segmentos que preservan los momentos de primer y segundo orden, esto fue utilizado por [Richardson y Green \(1997\)](#) al problema de determinar el número de componentes en distribuciones mezclas.

3. Obtener la probabilidad de aceptar el modelo con  $k + 1$  puntos de cambio como en (2.21).

#### Movimiento tipo 2: Eliminar un punto de cambio

Se propone pasar de un modelo  $M_i$  con  $k$  puntos de cambio a otro modelo  $M_j$  con  $k - 1$  puntos de cambio. Esto se realiza eliminando aleatoriamente un punto de cambio del modelo  $M_i$ . Se utiliza la transformación  $g^{-1}$  definida en (3.5) del parrafo anterior. Obténgase la probabilidad  $\alpha'$  de aceptar este movimiento como en (2.22).

#### Movimiento tipo 3: Actualizar parámetros mediante el muestreador de Gibbs

Mediante el algoritmo muestreador de Gibbs y las distribuciones condicionales de  $\mu_{i_s}$ ,  $\phi_{i_s}$

### 3.2. Implementación del algoritmo RJMCMC al problema de PC y CEN

---

y  $\sigma_{i_s}^2$  dadas en el Apéndice B, se generan valores para  $\mu_{i_s}$ ,  $\phi_{i_s}$  y  $\sigma_{i_s}^2$  en cada segmento. Las densidades condicionales de  $\mu_{i_s}$ ,  $\phi_{i_s}$  y  $\sigma_{i_s}^2$  para implementar este algoritmo están dadas por

$$\mu_{i_s} | \tau_{i_s}, \phi_{i_s}, \sigma_{i_s}^2, y^{\tau_{i_s-1}+1:\tau_{i_s}} \sim N \left( \frac{A(1 - \phi_{i_s})^2 \sigma_0^2 + \mu_0 \sigma_{i_s}^2}{A \bar{v}_{i_s} (1 - \phi_{i_s})^2 \sigma_0^2 + \sigma_{i_s}^2}, \frac{\sigma_0^2 \sigma_{i_s}^2}{A(1 - \phi_{i_s})^2 \sigma_0^2 + \sigma_{i_s}^2} \right),$$

donde  $y^{\tau_{i_s-1}+1:\tau_{i_s}} = (y_{\tau_{i_s-1}+1}, \dots, y_{\tau_{i_s}})$ ,  $\bar{v}_{i_s} = \frac{1}{n} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \phi_{i_s} y_{j-1})$   
y  $A = (\tau_{i_s} - \tau_{i_s-1})$ ,

$$\phi_{i_s} | \mu_{i_s}, \tau_{i_s}, \sigma_{i_s}^2, y^{\tau_{i_s-1}+1:\tau_{i_s}}, k \sim N \left( \frac{\sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} B(y_j - \mu_{i_s})}{\sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} B}, \frac{\sigma_{i_s}^2}{\sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} B} \right),$$

donde  $B = (y_{j-1} - \mu_{i_s})$ ,

$$\sigma_{i_s}^2 | \mu_{i_s}, \tau_{i_s}, \phi_{i_s}, y^{\tau_{i_s-1}+1:\tau_{i_s}} \sim IG \left( \alpha + \frac{A}{2}, \beta + \frac{1}{2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s} (y_{j-1} - \mu_{i_s}))^2 \right),$$

donde  $A = (\tau_{i_s} - \tau_{i_s-1})$ .

#### Movimiento tipo 4: Simulación de observaciones censuradas

La simulación de las observaciones censuradas se realiza mediante el método de [Jung et al. \(2005\)](#), éste método consiste en construir una matriz de permutación, seleccionar valores iniciales para la media y la matriz de covarianzas de la distribución normal multivariada, simular los valores censurados mediante un vector aleatorio de la distribución normal multivariada truncada de la parte censurada dada la parte observada, utilizar las observaciones simuladas para completar el conjunto de observaciones y estimar la media y la matriz de covarianzas de la distribución normal multivariada. En el Sección 2.2 se describe el método de [Jung et al. \(2005\)](#).

El algoritmo de [Jung et al. \(2005\)](#) para simular las observaciones censuradas, consiste en los siguientes pasos:

1. Construir las matrices de permutación  $P_o$  y  $P_c$ , así como también construir los vectores  $\mathbf{X}_o$  y  $\mathbf{X}_c$ , mediante las expresiones dadas en (2.7) y (2.8) respectivamente.
2. Obtener valores iniciales  $\hat{\mu}^{(o)}$ ,  $\hat{\phi}^{(o)}$  y  $\hat{\sigma}_{(o)}^2$ . Entonces calcular la media y la matriz de

### 3.3. Análisis de datos simulados

---

covarianza, mediante las expresiones siguientes,

$$\begin{aligned}\mu^{(o)} &= \mu^{(o)} \mathbf{1}_n. \\ \Sigma_{ij}^{(o)} &= \frac{\sigma_o^2}{(1 - \phi_o^2)} \phi_o^{|i-j|} \quad i, j = 1, 2, \dots, n.\end{aligned}$$

3. Calcular  $\Sigma_{oo}^o$ ,  $\Sigma_{oc}^o$ ,  $\Sigma_{co}^o$  y  $\Sigma_{cc}^o$  mediante las expresiones dadas en (2.4), y calcular la media y varianza condicional de la parte censurada como sigue,

$$\begin{aligned}\nu^{(o)} &= \mu_c^{(o)} + \Sigma_{co}^{(o)} (\Sigma_{oo}^{(o)})^{-1} (\mathbf{y}_o - \mu_o^{(o)}). \\ \Delta^{(o)} &= \Sigma_{cc}^{(o)} - \Sigma_{co}^{(o)} (\Sigma_{oo}^{(o)})^{-1} \Sigma_{oc}^{(o)}.\end{aligned}$$

4. Sea  $\mathbf{y}_c$  una realización de  $N_T(\nu^{(o)}, \Delta^{(o)})$ .
5. Obtener el vector de observaciones completas como sigue,

$$\mathbf{y} = P^{-1}(\mathbf{y}_c, \mathbf{y}_o)^t,$$

con  $P$  dada por  $P = (P_o, P_c)^t$ .

6. Estimar  $\mu$ ,  $\phi$  y  $\sigma^2$  por mínimos cuadrados.

### 3.3. Análisis de datos simulados

En esta sección se implementa el algoritmo RJMCMC a un conjunto de datos simulados que tendrá diferentes porcentajes de censura. Considerese el proceso siguiente que tiene dos puntos de cambio

$$X_t = \begin{cases} 12 + 0.5(X_{t-1} - 12) + \epsilon_t, & 1 \leq t \leq 200, \\ 12 + 0.79(X_{t-1} - 12) + \epsilon_t, & 201 \leq t \leq 400, \\ 12 - 0.5(X_{t-1} - 12) + \epsilon_t, & 401 \leq t \leq 600, \end{cases} \quad (3.6)$$

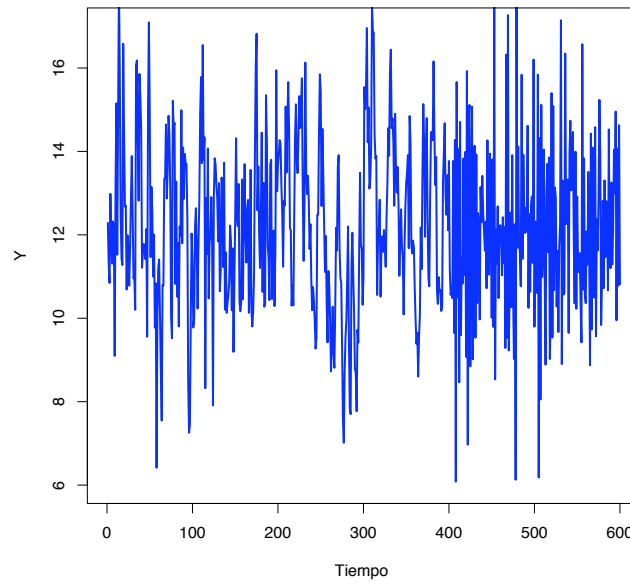
Note que este proceso consiste de tres segmentos autoregresivos de orden uno y los puntos de cambio están localizados en  $\tau_1 = 200$  y  $\tau_2 = 400$ . Los parámetros para los segmentos autoregresivos son  $\mu = (12, 12, 12)$ ,  $\sigma = (1.73, 1.22, 1.73)$  y  $\phi = (0.5, 0.79, -0.5)$ .



### 3.3. Análisis de datos simulados

---

En la figura 3.1 se muestra una realización  $x_1, x_2, \dots, x_n$  del proceso en (3.6).



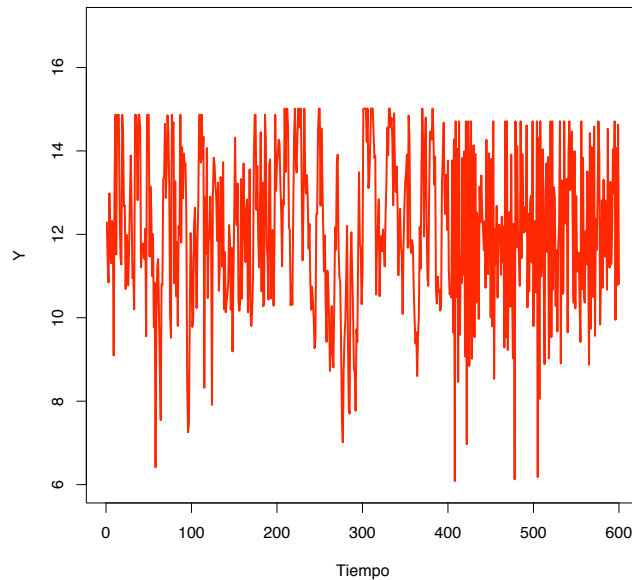
**Figura 3.1:** Una realización del proceso en (3.6).

### 3.3. Análisis de datos simulados

---

i) 10% de censura

Sea  $y_1, y_2, \dots, y_{600}$  observaciones que se obtienen al censurar 10% los valores de la realización  $x_1, x_2, \dots, x_n$  del proceso en (3.6) que se muestra en la figura 3.1. Estas observaciones censuradas se muestran en la figura 3.2.



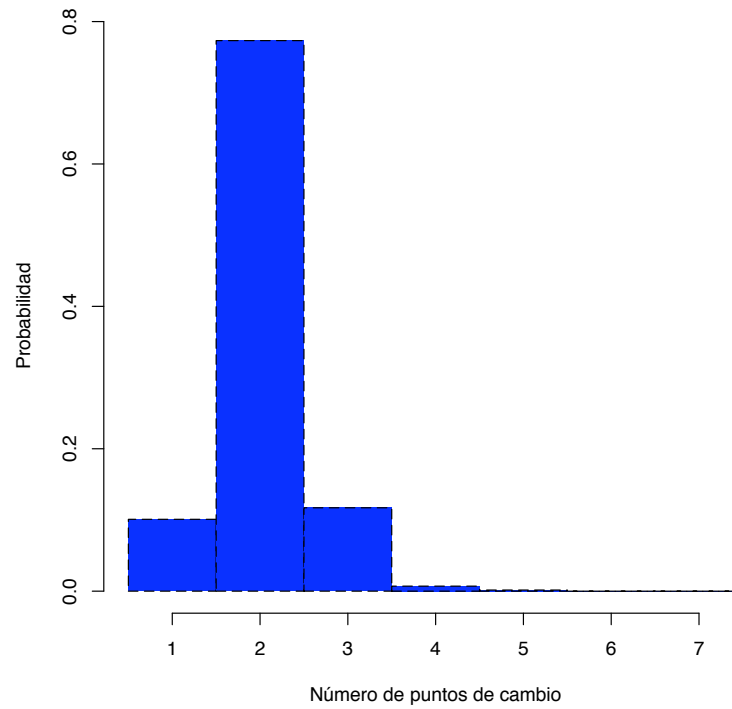
**Figura 3.2:** Observaciones censuradas (10%) de la realización del proceso (3.6).

Suponga que los hiperparámetros de las distribuciones iniciales de las medias y las varianzas que se requieren en (3.2) y (3.3) son:  $\mu_0 = 12$ ,  $\sigma_0^2 = 3$ ,  $\alpha_0 = 5$ ,  $\beta_0 = 1.2$ . También, considérese que las probabilidades de seleccionar cualquiera de los cuatro tipos de movimientos es  $\frac{1}{4}$  y que el valor inicial para el número de puntos de cambio se fija en  $k = 2$ . Se realizan 50,000 iteraciones del algoritmo RJMCMC con la muestra que se obtiene los valores  $R_1$  y  $R_2$  para diagnosticar la convergencia de la cadena vista en la Sección 2.7 son  $R_1 = 1.20$  y  $R_2 = 1.19$  los cuales son mayores que 1, de acuerdo con [Castelloe \(1998\)](#), con las 50000 iteraciones la cadena ha convergido a su distribución estacionaria.

### 3.3. Análisis de datos simulados

---

Usando la cadena se estimó la distribución final de  $K$ , el número de puntos de cambio. En la Figura 3.3 se muestra esta estimación. En base a esta estimación, se observa que el número de puntos de cambio se puede estimar es 2.

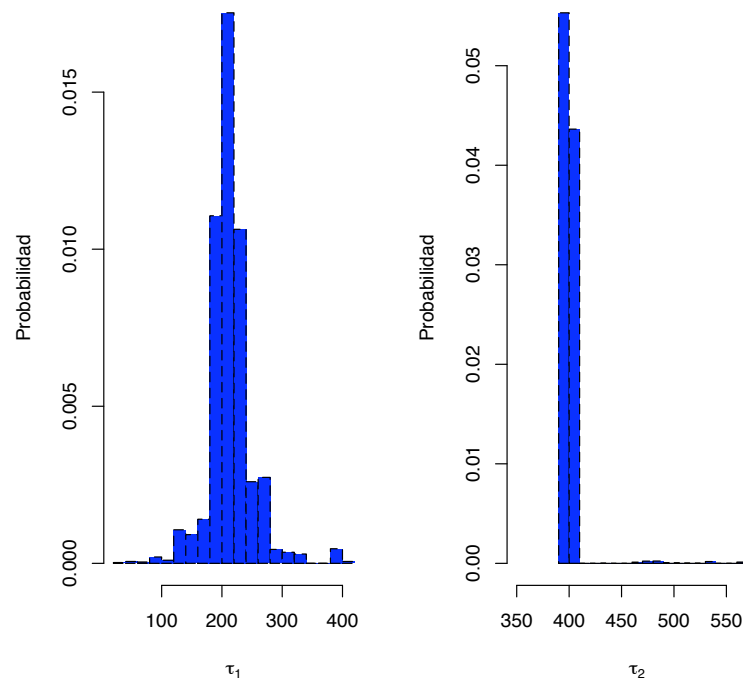


**Figura 3.3:** Histograma del número de puntos de cambio (10% censura)

### 3.3. Análisis de datos simulados

---

Usando la cadena se estimó la distribución final de las localizaciones  $\tau_1$  y  $\tau_2$ , condicionado a  $k = 2$ . Estas estimaciones se muestran en la Figura 3.4.



**Figura 3.4:** Distribución final de  $\tau_1$  y  $\tau_2$ , condicionado a  $k = 2$  (10% censura).

### 3.3. Análisis de datos simulados

---

En la Tabla 3.1 se muestran los valores verdaderos, los valores estimados y la desviación estándar de los puntos de cambio, las localizaciones de estos puntos, así como de las medias, las varianzas y los coeficientes autoregresivos en cada segmento que se obtiene al usar la muestra obtenida con el algoritmo RJMCMC. Estas estimaciones están condicionadas a que se tienen 2 puntos de cambio ( $k = 2$ ).

**Tabla 3.1:** Estimaciones para el caso de dos puntos de cambio (10% censura)

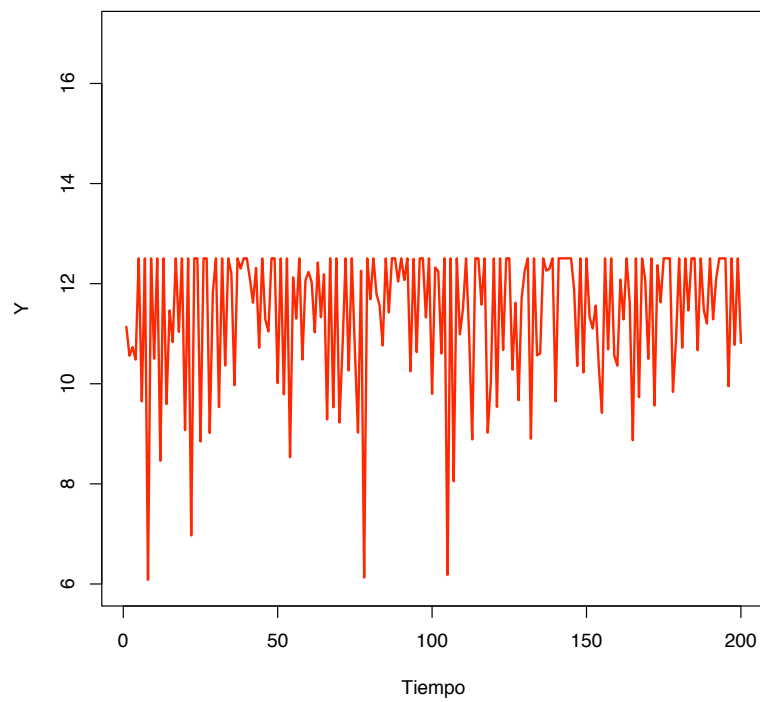
Parámetros	Valor verdadero	Valor estimado	Desviación estándar estimado
$\tau_1$	200	213	36.89
$\tau_2$	400	394	12.86
$\mu_1$	12	12.21	0.20
$\mu_2$	12	12.12	0.38
$\mu_3$	12	11.99	0.12
$\sigma_1$	1.73	1.37	0.16
$\sigma_2$	1.22	1.01	0.16
$\sigma_3$	1.73	1.375	0.17
$\phi_1$	0.50	0.41	0.06
$\phi_2$	0.79	0.77	0.14
$\phi_3$	-0.50	-0.509	0.05

### 3.3. Análisis de datos simulados

---

ii) 40% de censura:

Se censura ahora 40% la realización  $x_1, x_2, \dots, x_{600}$  que se muestra en la figura 3.1. En la Figura 3.5 se muestra las observaciones así censuradas.



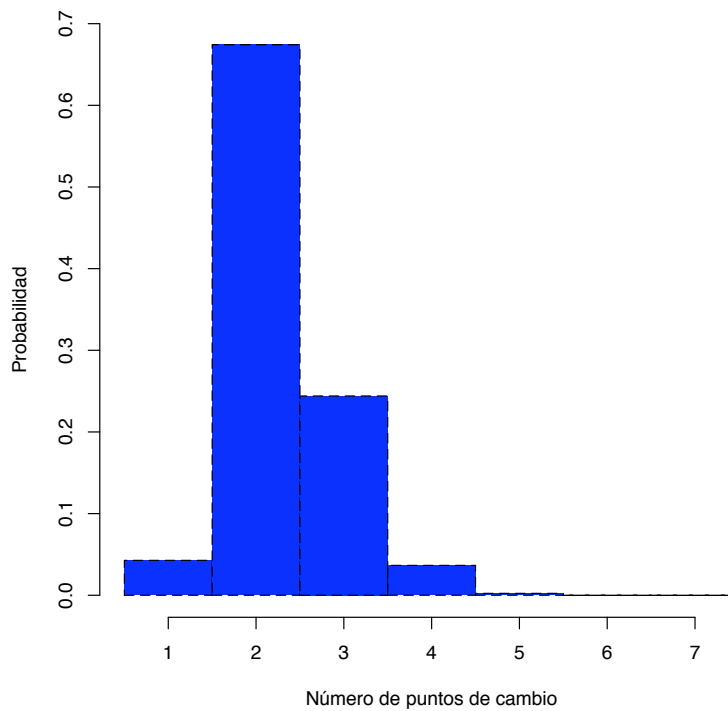
**Figura 3.5:** Observaciones censuradas (40%) de la realización del proceso (3.6).

### 3.3. Análisis de datos simulados

---

Se realizan 50,000 iteraciones del algoritmo RJMCMC con la muestra que se obtiene los valores  $R_1$  y  $R_2$  para diagnosticar la convergencia de la cadena vista en la Sección 2.7 son  $R_1 = 1.003$  y  $R_2 = 1.0018$  los cuales son mayores que 1, de acuerdo con [Castelloe \(1998\)](#), con las 50000 iteraciones la cadena ha convergido a su distribución estacionaria.

Usando la cadena se estimó la distribución final de  $K$ , el número de puntos de cambio. En la Figura 3.6 se muestra esta estimación. En base a esta estimación, se observa que el número de puntos de cambio se puede estimar es 2.

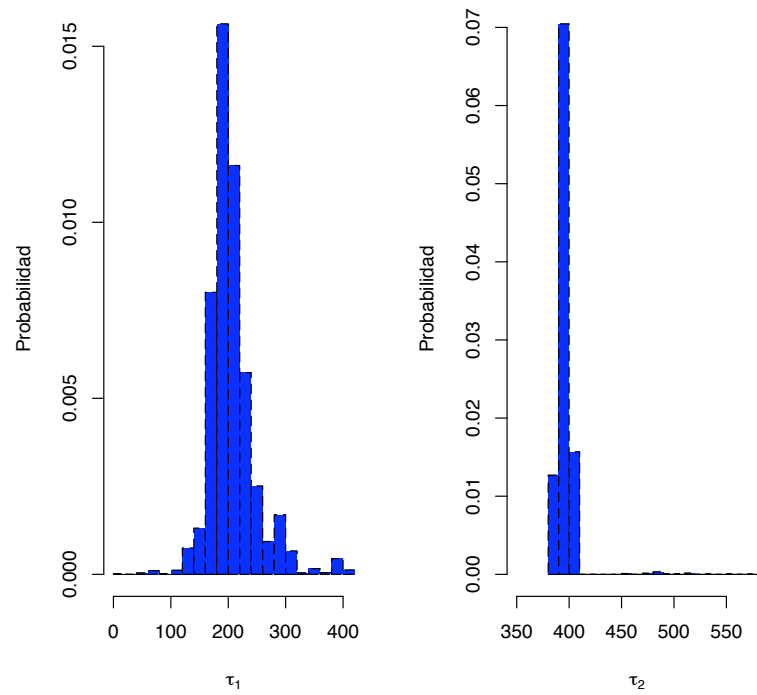


**Figura 3.6:** Histograma del número de puntos de cambio (40% censura)

### 3.3. Análisis de datos simulados

---

Usando la cadena se estimó la distribución final de las localizaciones  $\tau_1$  y  $\tau_2$ , condicionado a  $k = 2$ . Estas estimaciones se muestran en la Figura 3.7.



**Figura 3.7:** Histograma de las localizaciones de los PC (40% censura)



### 3.3. Análisis de datos simulados

---

En la Tabla 3.2 se muestran los valores verdaderos, los valores estimados y la desviación estándar de los puntos de cambio, las localizaciones de estos puntos, así como de las medias, las varianzas y los coeficientes autoregresivos en cada segmento que se obtiene al usar la muestra obtenida con el algoritmo RJMCMC. Estas estimaciones están condicionadas a que se tienen 2 puntos de cambio ( $k = 2$ ).

**Tabla 3.2:** Estimaciones para el caso de dos puntos de cambio (40 % censura).

Parámetros	Valor verdadero	Valor estimado	Desviación estándar estimada
$\tau_1$	200	199	41.38
$\tau_2$	400	398	12.47
$\mu_1$	12	11.69	0.12
$\mu_2$	12	11.76	0.27
$\mu_3$	12	11.47	0.06
$\sigma_1$	1.73	1.02	0.09
$\sigma_2$	1.22	0.67	0.1
$\sigma_3$	1.73	1.20	0.12
$\phi_1$	0.5	0.34	0.079
$\phi_2$	0.79	0.76	0.137
$\phi_3$	-0.5	-0.35	0.059

### 3.3. Análisis de datos simulados

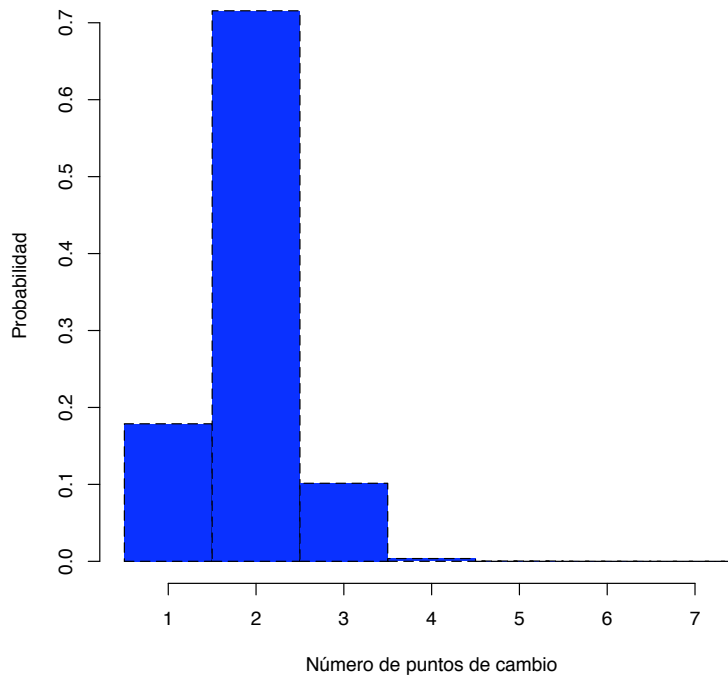
---

iii) *Sin censura:*

Se implementa ahora el algoritmo RJMCMC a la realización  $x_1, x_2, \dots, x_{600}$  que se muestra en la figura 3.1. Esta vez no se censura ninguna observación.

Se realizan 50,000 iteraciones del algoritmo RJMCMC con la muestra que se obtiene los valores  $R_1$  y  $R_2$  para diagnosticar la convergencia de la cadena vista en la Sección 2.7 son  $R_1 = 1.27$  y  $R_2 = 1.18$  los cuales son mayores que 1, de acuerdo con [Castelloe \(1998\)](#), con las 50000 iteraciones la cadena ha convergido a su distribución estacionaria.

Usando la cadena se estimó la distribución final de  $K$ , el número de puntos de cambio. En la Figura 3.8 se muestra esta estimación. En base a esta estimación, se observa que el número de puntos de cambio se puede estimar es 2.

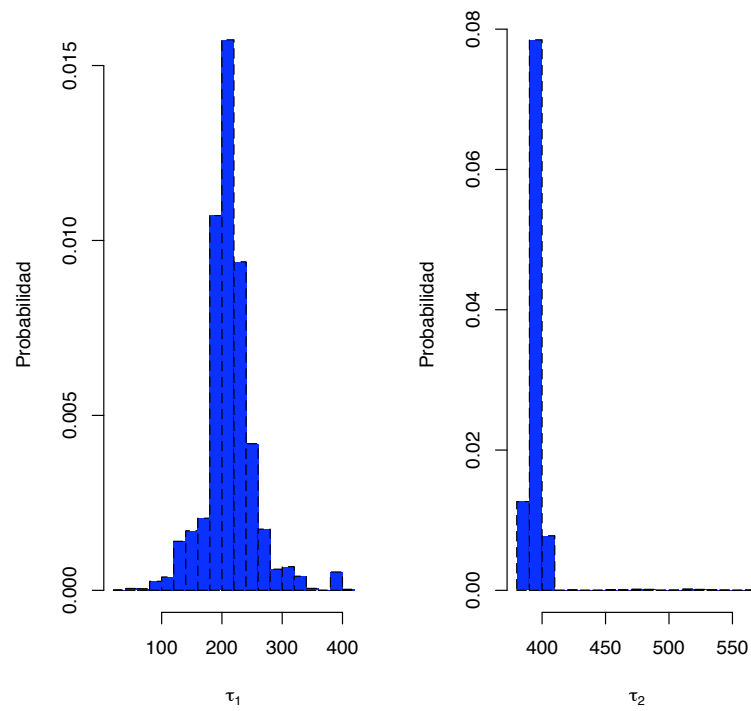


**Figura 3.8:** Histograma del número de puntos de cambio

### 3.3. Análisis de datos simulados

---

Usando la cadena se estimó la distribución final de las localizaciones  $\tau_1$  y  $\tau_2$ , condicionado a  $k = 2$ . Estas estimaciones se muestran en la Figura 3.9.



**Figura 3.9:** Histograma de las localizaciones de los puntos de cambio

### 3.3. Análisis de datos simulados

---

En la Tabla 3.3 se muestran los valores verdaderos, los valores estimados y la desviación estándar de los puntos de cambio, las localizaciones de estos puntos, así como de las medias, las varianzas y los coeficientes autoregresivos en cada segmento que se obtiene al usar la muestra obtenida con el algoritmo RJMCMC. Estas estimaciones están condicionadas a que se tienen 2 puntos de cambio ( $k = 2$ ).

**Tabla 3.3:** Estimaciones para el caso de dos puntos de cambio (0% censura).

Parámetros	Valor verdadero	Valor estimado	Desviación estándar estimada
$\tau_1$	200	209	41.24
$\tau_2$	400	391	13.01
$\mu_1$	12	12.28	0.21
$\mu_2$	12	12.22	0.41
$\mu_3$	12	12.10	0.10
$\sigma_1$	1.73	1.58	0.18
$\sigma_2$	1.22	1.21	0.20
$\sigma_3$	1.73	1.60	0.21
$\phi_1$	0.5	0.44	0.06
$\phi_2$	0.79	0.77	0.15
$\phi_3$	-0.5	-0.52	0.05

## Capítulo 4

# CONCLUSIONES

En este trabajo de tesis se ha propuesto un modelo bayesiano para analizar series de tiempo que presentan cambios en su estructura o cambios en sus parámetros, y debido a limitaciones en los instrumentos de medición puede presentar también censura en las observaciones. Se asume desconocido el número de puntos de cambio, así como las posiciones de los puntos de cambio. En cada segmento se han asumido procesos autoregresivos de orden uno. Para las medias y las varianzas en cada segmento se consideran distribuciones iniciales conjugadas, excepto para los coeficientes autoregresivos. Los métodos clásicos de MCMC no se pueden implementar directamente para analizar el modelo bayesiano propuesto. El algoritmo de Cadenas de Markov Monte Carlo con saltos reversibles (RJCMCMC) desarrollado por [Green \(1995\)](#), el cual es una generalización de las técnicas MCMC en el que la dimensión del espacio parametral puede cambiar de una iteración a otra. Se implementa para obtener una muestra de la distribución final de los puntos de cambio y sus localizaciones. En el ejemplo numérico que se presentó en la Sección 3.3 se observa que las estimaciones del número de puntos de cambio y de las localizaciones de esos puntos de cambio tiene poco sesgo. Además éstas estimaciones son robustas al porcentaje de censura.

# BIBLIOGRAFÍA

- Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*. Third eddition, Wiley.
- Ariza-Hernandez, F. J. y Rodríguez-Yam, G. A. (2008). Analysis of time series with censored observations. Inf. téc., Asociación Mexicana de Estadística.
- Barbieri, M. M. y Ohagan, A. (1996). A reversible jump mcmc sampler for bayesian analysis of arma time series. Inf. téc., Dipartimento di Statistica, Probabilità e Statistiche applicate, Università "La Sapienza", Roma and Department of Mathematics, University of Nottingham.
- Basseville, M. y Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall, Englewood Cliffs, New Jersey.
- Bhattacharya, P. K. (1994). Some Aspects of Change Point Analysis. *In IMS Lecture Notes-Monograph Series*, 23, 28–56.
- Brockwell, P. J. y Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. 2nd Ed. Springer.
- Brooks, S. P., Giudic, P. y Roberts, G. O. (2003). Efficient Construction of Reversible Jump Markov Chain Monte Carlo Proposal Distributions. *Journal of the Royal Statistical Society Series B-statistical Methodology*, 65, 3–55.
- Carlin, B. P. y Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman - Hall/CRC.
- Casella, G. y George, E. I. (1992). Explaining the Gibbs Sampler. *The American Statistician*, 46, 167–174.
- Castelloe, J. (1998). *Issues in reversible jump Markov chain Monte Carlo and composite EM analysis, applied to spatial poisson cluster processes*. Tesis Doctoral, University of Iowa.
- Chib, S. (1998). Estimation and Comparison of Multiple Change Point Models. *Journal of Econometrics*, 86, 221–241.

## BIBLIOGRAFÍA

---

- Chib, S. y Greeberg, E. (1995). Understanding the Metropolis-Hasting algorithm. *The American Statistician*, 49, 167–174.
- Csörgo, M. y Horvath, L. (1998). *Nonparametric Methods for Change Point Problems*. Prentice Hall, Englewood Cliffs, New Jersey.
- Davis, R. A., Lee, T. C. M. y Rodriguez-Yam, G. A. (2006). Structural Breaks Estimation for Non-stationary Time Series Models. *Journal of the American Statistical Association*, 101, 223–239.
- Dempster, A. P., Laird, N. y Rubin, D. (1977). Maximum Likelihood From Incomplete Data via the E. M. Algorithm. *Journal of the Royal Statistical Society, Series B*, 39, 1–38.
- Ehlers, R. y Brooks, S. (2002). Efficient construction of reversible jump mcmc proposals for autoregressive time series models. Inf. téc., University of Cambridge.
- Gamerman, D. (1997). *Markov chain Monte Carlo: Stochastic Simulation Bayesian Inference*. London: Chapman and Hall.
- Gelfand, A. E., Hills, S. E., Racine-Poon, A. y Smith, A. (1990). Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *Journal of American Statistical Association*, 85, 972–985.
- Gelman, A., Carlin, J. B., Stern, H. S. y Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman Hall.
- Gijbels, I. y Gurler, U. (2003). Estimation of a Change Point in a Hazard Function Based on Censored Data. *Lifetime Data Analysis*, 9, 395–411.
- Gilks, W. R., Richardson, S. y (eds.), D. J. S. (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.
- Giordani, P. y Khon, R. (2008). Efficient Bayesian Inference for Multiple Change-point and Mixture Innovation Models. *American Statistical Association*, 26, 66–77.
- Gombay, E. y Horvath, L. (1994). An application of the Maximum Likelihood Test to the Change-point Problem. *Stochastic Process Application*, 50, 161–171.
- Green, P. (1995). Reversible Jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometric*, 82 (4), 711–732.
- Green, P. (2003). *Trans-dimensional Markov Chain Monte Carlo.*, cap. Highly Structured Stochastic System, 179–198. Oxford University Press, Oxford.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

## BIBLIOGRAFÍA

---

- Hopke, P. K., Liu, C. y Rubin, D. B. (2001). Multiple Imputation for Multivariate Data with Missing and Below-Threshold Measurements: Time-series Concentration of Pollutants in the Arctic. *Biometrics*, 57, 22–33.
- Jung, W. P., Mark, G. G. y Sujit, K. G. (2005). Censored Time Series Analysis with Autoregressive Moving Average Models. *Econometric*, 10, 234–256.
- Komarkova, L. (2004). *Change Point Testing Problem for Randomly Censored Data*. Tesis Doctoral, Department of Statistics, Charles University.
- Lai, T. L. (1995). Sequential Changepoint Detection in Quality Control and Dynamical Systems. *Journal of the Royal Statistical*, 57, 613 – 658.
- Lavielle, M. y Teyssiere, G. (2006). Detection of Multiple Change Points in Multivariate Time Series. *Lithuanian Mathematical Journal*, 46, 67–93.
- Little, R. J. A. y Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. New York, Wiley.
- Richardson, S. y Green, P. (1997). On Bayesian Analysis of Mixtures with an Unknown Number of Components (with discussion). *Journal of the Royal Statistical Society*, 59, 731–792.
- Robert, C. P. y Casella, G. (2002). *Monte Carlo Statistical Methods*. Springer, New York.
- Robinson, P. M. (1980). *Estimation and Forecasting for Time Series Containing Censored or Missing Observations*. Time Series, North-Holland Publishing Company, 167-182.
- Shuangquan, L. (1998). *Nonparametric Test for Change Point Problems with Random Censorship*. Tesis Doctoral, Department of Mathematical Sciences, Edmonton, Alberta.
- Shumway, R. H. y Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. Springer.
- Smith, A. F. M. y Roberts, G. (1993). Bayesian Computation via Gibbs Sampler and Related Markov Chain Monte Carlo Methods. *Journal Royal Statistical Society Serie B*, 55, 3–23.
- Troughton, P. T. y Godsill, S. J. (1997). A reversible jump sampler for autoregressive time series, employing full conditionals to achieve efficient model space moves. Inf. téc., Signal Processing and communications laboratory, Department of Engineering, University of Cambridge.
- Xia, B. S. J. (2006). *Optimal Sequentially Planned Change Point Detection Procedures*. Tesis Doctoral, The University of Texas at Dallas.



## Apéndice A: Función de densidad conjunta y densidad condicional

La inferencia estadística básica y, en general, los métodos estadísticos, se basan fundamentalmente en la distribución de las observaciones, y dependiendo del caso, es la distribución que se asume. La distribución más comúnmente usada con este propósito es la distribución normal, la cual ha probado ser un modelo adecuado para varios fenómenos reales. Esta sección está dedicada a presentar la distribución normal multivariada y sus propiedades básicas.

Sea el proceso autoregresivo de orden uno como sigue,

$$Y_i = \mu + \phi(Y_{i-1} - \mu) + \epsilon_i, \quad i = 1, 2, \dots, n. \quad (.1)$$

donde  $\epsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$ . Aquí  $\mu$ ,  $\sigma^2$  y  $\phi$ , son parámetros. Sea  $\mathbf{y}=(y_1, y_2, \dots, y_n)$  las observaciones provenientes del modelos anterior con vector de medias  $\mu$  y matriz de covarianzas  $\Sigma_{ij} = \frac{\sigma^2}{(1-\phi^2)}\phi^{|i-j|}$ . Esto es

$$\mathbf{y} \sim N_n(\mu, \Sigma)$$

La función de verosimilitud tiene la siguiente forma:

$$\begin{aligned} f(\mathbf{y}|\mu, \sigma^2, \phi) &= \frac{1}{\sqrt{(2\pi)^n(1-\phi^2)^{-1}}} e^{-\frac{1-\phi^2}{2\sigma^2}(y_1-\mu)^2 - \frac{1}{2\sigma^2} \sum_{j=2}^n [(y_j-\mu) - \phi(y_{j-1}-\mu)]^2} \\ &= \frac{\sqrt{1-\phi}}{\sigma} \phi\left(\frac{y_1-\mu}{\sigma}\sqrt{1-\phi^2}\right) \prod_{j=2}^n f(y_j|y_{j-1}) \end{aligned}$$

## Apéndice B: Densidad condicional de los parámetros autoregresivos

En este apéndice se obtienen las densidades condicionales para  $\mu$ ,  $\sigma$  y  $\phi$  que se utilizan en el movimiento tipo 3 de la implemetación del algoritmo RJMCMC. Supongase que se desean actualizar los parámetros  $\mu$ ,  $\tau$  y  $\phi$  en el  $i_s$  segmento con extremos en los puntos de cambio en  $\tau_{i_s-1}$  y  $\tau_{i_s}$ . La obtención de las densidades condicionales es como sigue,

## BIBLIOGRAFÍA

---

Densidad condicional de  $\mu$

$$\begin{aligned}
 f(\mu|\tau, \phi, \sigma^2, y_1, y_2, \dots, y_n) &= \frac{f(\mu, \tau, \phi, \sigma^2, y_c|y_o)}{f(\tau, \phi, \sigma^2, y_c|y_o)} \\
 &\propto f(\mu, \tau, \phi, \sigma^2, y_c|y_o) \\
 &\propto f(y_c, y_o|\mu, \tau, \phi, \sigma^2) f(\mu, \tau, \phi, \sigma^2) \\
 &= \prod_{i=1}^{k+1} \prod_{j=\tau_{i-1}+1}^{\tau_i} f(y_j|\mu, \tau, \phi, \sigma^2, y_{j-1}) \prod_{i=1}^{k+1} f(\mu_i|\mu_0, \sigma_0^2) \\
 f(\mu_{i_s}|\tau_{i_s}, \phi_{i_s}, \sigma_{i_s}^2, y_{\tau_{i_s-1}+1}, \dots, y_{\tau_{i_s}}) &= \prod_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} f(y_j|\mu_{i_s}, \tau_{i_s}, \phi_{i_s}, \sigma_{i_s}^2, y_{j-1}) f(\mu_{i_s}|\mu_0, \sigma_0^2) \\
 &= e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s}(y_{j-1} - \mu_{i_s}))^2} e^{\frac{-1}{2\sigma_0^2} (\mu_{i_s} - \mu_0)^2} \\
 &= e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \phi_{i_s} y_{j-1} - \mu_{i_s}(1 - \phi_{i_s}))^2} e^{\frac{-1}{2\sigma_0^2} (\mu_{i_s} - \mu_0)^2} \\
 &= e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (\Delta_{j i_s} - \mu_{i_s} B_{i_s})^2} e^{\frac{-1}{2\sigma_0^2} (\mu_{i_s} - \mu_0)^2} \\
 &= e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} ([\Delta_{j i_s} - \bar{\Delta}_{i_s}] + [\bar{\Delta}_{i_s} - \mu_{i_s} B_{i_s}])^2} e^{\frac{-1}{2\sigma_0^2} (\mu_{i_s} - \mu_0)^2} \\
 (\mu_{i_s}|\tau_{i_s}, \phi_{i_s}, \sigma_{i_s}^2, y_{\tau_{i_s-1}+1}, \dots, y_{\tau_{i_s}}) &\sim N \left( \frac{A \bar{v}_{i_s} (1 - \phi_{i_s})^2 \sigma_0^2 + \mu_0 \sigma_{i_s}^2}{A(1 - \phi_{i_s})^2 \sigma_0^2 + \sigma_{i_s}^2}, \frac{\sigma_0^2 \sigma_{i_s}^2}{A(1 - \phi_{i_s})^2 \sigma_0^2 + \sigma_{i_s}^2} \right)
 \end{aligned}$$

donde  $\bar{v}_{i_s} = \frac{1}{n} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \phi_{i_s} y_{j-1})$  y  $A = (\tau_{i_s} - \tau_{i_s-1})$ .

$$\begin{aligned}
 f(\sigma^2|\mu, \tau, \phi, y_o) &= \frac{f(\mu, \tau, \phi, \sigma^2, y_c|y_o)}{f(\tau, \phi, \mu, y_c|y_o)} \\
 &\propto f(\mu, \tau, \phi, \sigma^2, y_c|y_o) \\
 &\propto f(y_c, y_o|\mu, \tau, \phi, \sigma^2) f(\mu, \tau, \phi, \sigma^2) \\
 &= \prod_{i=1}^{k+1} \prod_{j=\tau_{i-1}+1}^{\tau_i} f(y_j|\mu, \tau, \phi, \sigma^2, y_{j-1}) \prod_{i=1}^{k+1} f(\sigma_i^2|\alpha, \beta) \\
 f(\sigma_{i_s}^2|\mu_{i_s}, \tau_{i_s}, \phi_{i_s}, y_{\tau_{i_s-1}+1}, \dots, y_{\tau_{i_s}}) &= \prod_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} f(y_j|\mu, \tau, \phi, \sigma^2, y_{j-1}) Ig(\sigma_{i_s}|\alpha, \beta) \\
 &= \sigma_{i_s}^{-(\tau_{i_s} - \tau_{i_s-1})} e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s}(y_{j-1} - \mu_{i_s}))^2} \sigma_{i_s}^{-2(\alpha-1)} e^{\frac{-\beta}{\sigma_{i_s}^2}} \\
 (\sigma_{i_s}^2|\mu_{i_s}, \tau_{i_s}, \phi_{i_s}, y_{\tau_{i_s-1}+1}, \dots, y_{\tau_{i_s}}) &\sim IG \left( 0.5A + \alpha, 0.5 \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s} B^2) + \beta \right).
 \end{aligned}$$

## BIBLIOGRAFÍA

---

donde  $A = (\tau_{i_s} - \tau_{i_s-1})$  y  $B = (y_{j-1} - \mu_{i_s})$ .

$$\begin{aligned}
 f(\phi|\mu, \tau, \phi, \sigma^2, y_1, y_2, \dots, y_n) &= \frac{f(\mu, \tau, \phi, \sigma^2, y_c|y_o)}{f(\mu, \tau, \sigma^2, y_c|y_o)} \\
 &\propto f(\mu, \tau, \phi, \sigma^2, y_c|y_o) \\
 &\propto f(y_c, y_o|\mu, \tau, \phi, \sigma^2) f(\mu, \tau, \phi, \sigma^2) \\
 &= \prod_{i=1}^{k+1} \prod_{j=\tau_{i-1}+1}^{\tau_i} f(y_j|\mu, \tau, \phi, \sigma^2, y_{j-1}) \prod_{i=1}^{k+1} f(\phi) \\
 f(\phi_{i_s}|\mu_{i_s}, \tau_{i_s}, \sigma_{i_s}^2, y^{\tau_{i_s-1}+1:\tau_{i_s}}, k) &\propto \sigma_{i_s}^{-(\tau_{i_s}-\tau_{i_s-1})} e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s}(y_{j-1} - \mu_{i_s}))^2} \frac{1}{2} \\
 &\propto e^{\frac{-1}{2\sigma_{i_s}^2} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_{j-1} - \mu_{i_s}) \left( \phi_{i_s} - \frac{(y_j - \mu_{i_s})}{(y_{j-1} - \mu_{i_s})} \right)^2} \\
 (\phi_{i_s}|\mu_{i_s}, \tau_{i_s}, \sigma_{i_s}^2, y^{\tau_{i_s-1}+1:\tau_{i_s}}, k) &\sim N \left( \frac{\sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} B(y_j - \mu_{i_s})}{\sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} B}, \frac{\sigma_{i_s}^2}{\sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} B} \right)
 \end{aligned}$$

donde  $B = (y_{j-1} - \mu_{i_s})$ .

## Apéndice C: Probabilidad de aceptación de un nuevo punto de cambio

En el algoritmo de RJMCMC, la probabilidad de aceptación de un nuevo punto de cambio, se puede expresar como sigue,

$$\alpha = \min(1, R).$$

con  $R =$  (razón de verosimilitud) (razón inicial) (razón de propuesta) (jacobiano)

La razón de verosimilitud ( $RV$ ) está dada por

$$\begin{aligned}
 RV &= \frac{\prod_{i=1}^{k+2} \prod_{j=\tau_{i-1}+1}^{\tau_i} f(y_j|\mu'_i, \tau'_i, \phi'_i, \sigma'^2_i, y_{j-1}, k+1)}{\prod_{i=1}^{k+1} \prod_{j=\tau_{i-1}+1}^{\tau_i} f(y_j|\mu_i, \tau_i, \phi_i, \sigma^2_i, y_{j-1}, k)} \\
 &= \frac{\prod_{j=\tau_{i_s-1}+1}^{\tau'_{i_s}} f(y_j|\mu'_{i_s}, \tau'_{i_s}, \phi'_{i_s}, \sigma'^2_{i_s}) \prod_{j=\tau_{i_s}'+1}^{\tau_{i_s}} f(y_j|\mu_{i_s}, \tau_{i_s}, \phi_{i_s}, \sigma^2_{i_s})}{\prod_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} f(y_j|\mu_{i_s}, \tau_{i_s}, \phi_{i_s}, \sigma^2_{i_s}, y_{j-1}, k)} \\
 &= \frac{C e^{\frac{-1}{2\sigma'^2_{i_s}} \sum_{j=\tau_{i_s-1}+1}^{\tau'_{i_s}} (y_j - \mu'_{i_s} - \phi'_{i_s}(y_{j-1} - \mu'_{i_s}))^2} D e^{\frac{-1}{2\sigma^2_{i_s}} \sum_{j=\tau_{i_s}'+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s}(y_{j-1} - \mu_{i_s}))^2}}{\sigma_{i_s}^{-(\tau_{i_s}-\tau_{i_s-1})} e^{\frac{-1}{2\sigma^2_{i_s}} \sum_{j=\tau_{i_s-1}+1}^{\tau_{i_s}} (y_j - \mu_{i_s} - \phi_{i_s}(y_{j-1} - \mu_{i_s}))^2}}
 \end{aligned}$$

donde  $C = \sigma'^{-(\tau'_{i_s}-\tau_{i_s-1})}$  y  $D = \sigma'^{-(\tau_{i_s}-\tau'_{i_s})}$

## BIBLIOGRAFÍA

---

La razon inicial ( $RI$ ) es como sigue

$$\begin{aligned} RI &= \frac{f(\mu', \tau', \phi', \sigma'^2, k+1)}{f(\mu, \tau, \phi, \sigma^2, k)} \\ &= \frac{f(\tau'|k+1)f(\mu'|k+1)f(\phi'|k+1)f(\sigma'^2|k+1)f(k+1)}{f(\tau|k)f(\mu|k)f(\phi|k)f(\sigma^2|k)f(k)} \end{aligned}$$

$$\frac{f(\tau'|k+1)}{f(\tau|k+1)} = \frac{f(\phi'|k+1)}{f(\phi|k+1)} = 1$$

Ahora

$$\begin{aligned} \frac{f(\mu'|k+1)}{f(\mu|k)} &= \frac{\prod_{i=1}^{k+2} f(\mu'_i|\mu_0, \sigma_0^2)}{\prod_{i=1}^{k+1} f(\mu_i|\mu_0, \sigma_0^2)} \\ &= \frac{f(\mu_{i_s}'|\mu_0, \sigma_0^2)f(\mu_{i_s}'|\mu_0, \sigma_0^2)}{f(\mu_{i_s}|\mu_0, \sigma_0^2)} \\ &= e^{\frac{-1}{2\sigma_0^2}(\mu_{i_s}'-\mu_0)+(\mu_{i_s}'-\mu_0)-(\mu_{i_s}-\mu_0)^2} \end{aligned}$$

luego

$$\begin{aligned} \frac{f(\sigma_i^2|k+1)}{f(\sigma_i^2|k)} &= \frac{\prod_{i=1}^{k+2} Ig(\sigma_i^2|\alpha, \beta)}{\prod_{i=1}^{k+1} Ig(\sigma_i^2|\alpha, \beta)} \\ &= \frac{Ig(\sigma_{i_s}^2|\alpha, \beta)Ig(\sigma_{i_s}'^2|\alpha, \beta)}{Ig(\sigma_{i_s}^2|\alpha, \beta)} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma_{i_s}'^{-2}\sigma_{i_s}^{-'2}\sigma_{i_s}^{-2})^{\alpha+1} e^{-\beta(\sigma_{i_s}'^{-2}+\sigma_{i_s}^{-'2}+\sigma_{i_s}^{-2})} \end{aligned}$$

Por lo tanto

$$RI = e^{\frac{-1}{2\sigma_0^2}(\mu_{i_s}'-\mu_0)+(\mu_{i_s}'-\mu_0)-(\mu_{i_s}-\mu_0)^2} \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma_{i_s}'^{-2}\sigma_{i_s}^{-'2}\sigma_{i_s}^{-2})^{\alpha+1} e^{-\beta(\sigma_{i_s}'^{-2}+\sigma_{i_s}^{-'2}+\sigma_{i_s}^{-2})}$$

## BIBLIOGRAFÍA

---

El jacobiano ( $J$ ) es

$$\left( \begin{array}{cccccc} u_1 & 0 & 0 & (\tau_i - \tau_{i-1}) & 0 & 0 \\ (1 - u_1) & 0 & 0 & -(\tau_{i_s} - \tau_{i_s-1}) & 0 & 0 \\ 0 & 1 & \frac{-1}{2}u_2\sqrt{\frac{\sigma_{i_s}^{-1}(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s}' - \tau_{i_s-1})}} & 0 & \sqrt{\frac{\sigma_{i_s}^2(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s}' - \tau_{i_s-1})}} & 0 \\ 0 & 1 & \frac{1}{2}u_2\sqrt{\frac{\sigma_{i_s}^{-1}(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s} - \tau_{i_s}')}} & 0 & \sqrt{\frac{\sigma_{i_s}^2(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s} - \tau_{i_s}')}} & 0 \\ \frac{u_3(1-u_2^2)\sigma_{i_s}^2}{(\tau_{i_s}' - \tau_{i_s-1})} & 0 & \frac{u_3(1-u_2^2)(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s}' - \tau_{i_s-1})} & 0 & \frac{-2u_3u_2\sigma_{i_s}^2(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s}' - \tau_{i_s-1})} & \frac{-(1-u_2^2)\sigma_{i_s}^2(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s}' - \tau_{i_s-1})} \\ \frac{(1-u_3)(1-u_2^2)\sigma_{i_s}^2}{(\tau_{i_s} - \tau_{i_s}')} & 0 & \frac{(1-u_3)(1-u_2^2)(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s} - \tau_{i_s}')} & 0 & \frac{2u_2(1-u_3)\sigma_{i_s}^2(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s} - \tau_{i_s}')} & \frac{-(1-u_2^2)\sigma_{i_s}^2(\tau_{i_s} - \tau_{i_s-1})}{(\tau_{i_s} - \tau_{i_s}')} \end{array} \right)$$