



**BENEMÉRITA UNIVERSIDAD
AUTÓNOMA DE PUEBLA
FACULTAD DE CIENCIAS FÍSICO
MATEMÁTICAS
POSGRADO EN CIENCIAS
MATEMÁTICAS**

**IMPLEMENTACIÓN DE ALGUNOS
PROCEDIMIENTOS COMPUTACIONALES EN
APLICACIONES PARA LA DETECCIÓN DE PUNTOS
DE CAMBIO**

LUCILA MUÑIZ MERINO

**T E S I S
PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:**

DOCTORADO EN CIENCIAS MATEMÁTICAS

**DIRECTORES DE TESIS
DR. BULMARO JUÁREZ HERNÁNDEZ
Y DR. HUGO ADÁN CRUZ SUÁREZ**

PUEBLA, PUEBLA, ENERO 2021

Agradecimientos

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

A la Benemérita Universidad Autónoma de Puebla, por haberme brindado la oportunidad de seguir mi formación académica en sus aulas.

A mis directores de tesis:

Dr. Bulmaro Juárez Hernández

Y Dr. Hugo Adán Cruz Suárez

A mi hermano:

Dr. Manolo Muñoz Merino

Dedicatoria

A mis **padres**: Fausto Muñiz Merino[†] y Rafaela Merino Cortes, gracias por darme la vida, amor, educación y por estar conmigo siempre.

A mis **hermanos y hermanas**.

A mis **sobrinos**.

Índice general

Lista de tablas	III
Lista de figuras	V
Introducción	1
Objetivos	5
1 Puntos de cambio	6
1.1 Historia de puntos de cambio.	6
1.2 Problemas de puntos de cambio espaciales	10
1.3 Inferencia Bayesiana	13
2 Modelos clásicos espacio tiempo	32
2.1 Puntos de cambio según Brodsky y Darkhovsky	32
2.2 Puntos de cambio abrupto y gradual	33
2.3 Puntos de cambio en espacio-tiempo	34
2.4 Métodos de puntos de cambio en estadística clásica	35
2.5 Aplicaciones	46
3 Detección de puntos de cambio	50
3.1 Factor de Bayes	51
3.2 Proceso de Poisson homogéneo	52
3.3 Método de bisección	53
3.4 Múltiples puntos de cambio temporales	53
3.5 Para 6 puntos de cambio	55
3.6 Múltiples puntos de cambio en espacio tiempo	58
3.7 Método de umbrales	62
3.8 CUSUM con datos de accidentes de tráfico	64
3.9 Otras Pruebas de datos de accidentes de tráfico	68
3.10 CUSUM con datos de coronavirus	68
4 Conclusiones	71
Trabajo a futuro	73

II

ÍNDICE GENERAL

Bibliografía

75

Apéndice

83

Lista de tablas

2.1	Regresión, Kernel, Análisis de espectro singular, Pettitt, Movimiento t, Rangos de Buishand, Homogeneidad normal estándar y Von Newman.	47
2.2	MANOVA, ARIMA, Bayesiano especial, Mann-Kendall, Mediana y ANOVA.	48
2.3	Mann-Kendall, Pettitt, Lepage, Bayesiano, spline, comparativo de datos e imágenes.	49
3.1	Escala de interpretación de B, según Jeffreys.	51
3.2	Puntos de cambio.	55
3.3	6 puntos de cambio.	56
3.4	6 puntos de cambio.	57
3.5	Valores de λ	59
3.6	Puntos de cambio espacio tiempo.	59
3.7	Puntos de cambio espacio tiempo.	60
3.8	Puntos de cambio espacio tiempo.	60
3.9	Puntos de cambio espacio tiempo.	61
3.10	Datos de regresión.	63
3.11	Datos de accidentes de tráfico.	66
3.12	Datos de contagios de coronavirus.	69
3.13	Estadísticas y porcentaje.	69

Lista de figuras

2.1	Gráfica de puntos de cambio.	33
2.2	Cambios espacio-temporales, la Figura fue obtenida de Xun <i>et al.</i> (2014).	35
2.3	Cambios basados en vectores, la Figura fue obtenida de Xun <i>et al.</i> (2014).	35
3.1	Gráfica de 5 puntos de cambio.	55
3.2	Gráfica de 6 puntos de cambio.	57
3.3	Gráfica con punto de cambio en 7.	66
3.4	Gráfica de contagios.	69
3.5	Datos grupos 1, 2 y 3	70

Introducción

Cuando se escucha decir punto de cambio la primer pregunta que surge es ¿qué es un punto de cambio? Chen y Gupta (2012) lo definen como el sitio o punto en el tiempo t , en una sucesión de datos $\{x_{t_i}\}$ $i = 1, \dots, n$ observados y ordenados respecto al tiempo tal que dichas observaciones siguen una distribución F_1 , antes de un punto, y en otro posterior la distribución a seguir es F_2 . Es decir, desde el punto de vista estadístico, la sucesión de observaciones muestra un comportamiento no homogéneo. El problema de punto de cambio es considerado como uno de los problemas centrales de inferencia estadística, pues relaciona a la teoría de control estadístico, a las pruebas de hipótesis (al detectar si existe algún cambio en la sucesión de variables aleatorias observadas), y a la teoría de estimación (al estimar el número de cambios y sus correspondientes localizaciones). Ésto bajo los enfoques clásico y Bayesiano, según Chen y Gupta (2012).

Chen y Gupta (2012) indican que los problemas de puntos de cambio originalmente surgieron en control de calidad y en general pueden ser encontrados en la modelación matemática de diversas disciplinas tales como Medio Ambiente, Epidemiología, Procesos de señal sísmica, Economía, Finanzas, Geología, Medicina, Biología, Física, etc.

En general el problema de puntos de cambio según Chen y Gupta (2012) se visualiza de la forma siguiente:

Sean X_1, X_2, \dots, X_n una colección de vectores (variables) aleatorios independientes con funciones de distribución de probabilidad F_1, F_2, \dots, F_n , respectivamente. Entonces el problema de puntos de cambio consiste en probar la hipótesis nula H_0 de la no existencia de cambio contra la alternativa H_a de que existe al menos un punto de cambio; lo cual se expresa de la siguiente manera:

$$H_0 : F_1 = F_2 = \dots = F_n$$

vs

$$H_a : F_1 = \dots = F_{(k_1)} \neq F_{(k_{1+1})} = \dots = F_{(k_2)} \neq F_{(k_{2+1})} =$$

$$\dots = F_{(k_q)} \neq F_{(k_{q+1})} = \dots = F_n,$$

donde $1 < k_1 < k_2 < \dots < k_q < n$, q es el número desconocido de puntos de cambio y k_1, k_2, \dots, k_q son las posiciones desconocidas respectivas que tienen que ser estimadas. Si las distribuciones F_1, F_2, \dots, F_n pertenecen a una familia paramétrica común $F(\theta)$, donde $\theta \in \mathbb{R}^p$, entonces el problema de puntos de cambio consiste en probar la hipótesis nula H_0 sobre la no existencia de cambio en los parámetros $\theta_i, i = 1, \dots, n$ de la población contra la alternativa H_a de que existe al menos un punto de cambio; lo cual se expresa de la siguiente forma:

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_n = \theta(\text{desconocidos})$$

vs

$$H_a : \theta_1 = \dots = \theta_{(k_1)} \neq \theta_{(k_{1+1})} = \dots = \theta_{(k_2)} \neq$$

$$\theta_{(k_{2+1})} = \dots = \theta_{(k_q)} \neq \theta_{(k_{q+1})} = \dots = \theta_n,$$

donde q y k_1, k_2, \dots, k_q tienen que ser estimados. Estas hipótesis revelan los aspectos de inferencia de puntos de cambio para determinar si existe algún punto de cambio en el proceso, estimar el número de ellos y sus respectivas posiciones.

En diversos casos se asume que las observaciones son independientes e idénticamente distribuidas (i.i.d.), el análisis resulta más complejo si se presenta dependencia entre las observaciones Chen y Gupta (2012). En el caso de series de tiempo la dependencia está presente entre las observaciones dentro de cada segmento de tiempo; en el caso de datos

espacio-temporales, la dependencia sucede sobre el espacio y el tiempo.

En este trabajo de tesis como inicio se plantean los objetivos. En el Capítulo 1 se incluye una pequeña historia de cómo se dieron los avances en puntos de cambio, posteriormente se da una sección de puntos de cambio en Estadística Espacial con razón de verosimilitud y finalmente se da una sección con inferencia Bayesiana, la cual es utilizada en el capítulo 3 para detectar puntos de cambio. En el Capítulo 2 se tratan los diferentes puntos de cambio, ellos son: según Brodsky y Darkhovsky (1993, 2000), visualización de puntos de cambio abrupto y gradual, puntos de cambio espacio tiempo según Xun *et al.* (2014). Posteriormente se tratan algunas metodologías paramétricas y no paramétricas y se presentan algunas aplicaciones. En el Capítulo 3 se hace un análisis de puntos de cambio con el factor de Bayes, umbrales y CUSUM. Para el análisis de puntos de cambio con el factor de Bayes se simularon datos Poisson; el método de umbrales se trabajó con una regresión y para la CUSUM se utilizaron datos de INEGI y de Coronavirus. En el Capítulo 4 se dan las conclusiones. Finalmente se incluye la propuesta de trabajo a futuro.

Objetivos

En este apartado se plantean el objetivo general y los objetivos específicos del trabajo de investigación sobre puntos de cambio.

Objetivo general

Implementación de procedimientos computacionales en el análisis y detección de puntos de cambio, considerando sus estadísticas de prueba, resultados numéricos y aplicaciones.

Objetivos específicos

- Investigar y analizar las metodologías desarrolladas en el estudio de puntos de cambio y sus diferentes aplicaciones.
- Examinar las estadísticas de prueba y metodologías paramétricas y no paramétricas de puntos de cambio.
- Desarrollar programas, estudiar y comparar los resultados de algunos procedimientos aplicados en la detección de los puntos de cambio, en particular el factor de Bayes, el método de umbrales y la estadística de suma acumulativa (CUSUM).

Capítulo 1

Puntos de cambio

En este Capítulo se incluye una pequeña historia de cómo se dieron los avances en puntos de cambio, posteriormente se introduce una sección de puntos de cambio en Estadística Espacial con razón de verosimilitud y finalmente se incorpora una sección con inferencia Bayesiana, la cual es utilizada en el Capítulo 3 para detectar puntos de cambio.

1.1 Historia de puntos de cambio.

Las primeras publicaciones sobre análisis de puntos de cambio aparecieron hacia la mitad del siglo pasado. Uno de los primeros trabajos sobre el problema de punto de cambio retrospectivo (a posteriori) incluye a Page (1957) quien propuso dicho método. Él asumió una sucesión de variables aleatorias independientes observadas con un cambio abrupto, se planteó la hipótesis nula que las observaciones provinieron de una sola distribución y la hipótesis alternativa que las primeras m observaciones provinieron de una distribución y el resto de otra. Supuso además que los parámetros de tales distribuciones eran conocidos. Para probar la hipótesis nula H_0 que se planteó, adoptó un enfoque de discriminante, dividió el espacio muestral en n subconjuntos a los cuales les asignó una hipótesis, la hipótesis H_i , para $i = 1, \dots, n$ fue aceptada para x en la región R_i si la verosimilitud en esa región fue mayor que la de otra región en comparación. El estimador del punto de cambio fue $\hat{t} = \inf\{k : S_{k-1} \geq S_j, j = 2, \dots, n\}$, donde $S_k = \sum_{t=1}^k [\ln f(x_t|\theta_1) - \ln f(x_t|\theta_2)]$, $f(x_t|\theta_1)$ es la función antes del punto de cambio y $f(x_t|\theta_2)$ es la función después del punto de cambio.

Por otra parte Hinkley (1970) mediante el uso de caminatas aleatorias, obtuvo las estadísticas de prueba bajo diferentes suposiciones respecto

a los parámetros, los cuales consideró correspondían a una distribución normal. Obtuvo las distribuciones asintóticas de las estadísticas de prueba y la del estimador de máxima verosimilitud del punto de cambio. Su método lo condujo a distribuciones asintóticas basadas en integrales extremadamente complejas, y al comparar las varianzas del estimador empírico de punto de cambio con los resultados de sus aproximaciones, determinó que sus resultados fueron pobres, ya que los valores no fueron parecidos entre sí, principalmente en muestras de tamaño pequeñas. Los tamaños de muestra que tomó fueron $T = 50, 100, 200$. Por su parte Hawkins (1977) por medio de un proceso de Markov, obtuvo la distribución nula de la estadística de prueba de la razón de \log verosimilitud de un solo punto de cambio sobre la media de una sucesión de variables i.i.d. a una distribución normal, supuso a la varianza σ conocida e igual a 1. Así mismo, proporcionó la distribución nula del punto de cambio, sus resultados de simulación mostraron que la aproximación es buena cuando el tamaño de muestra n y el nivel de significancia α son pequeños pero con errores moderados.

Más tarde Yao y Davis (1986) mostraron que la estadística de prueba cuando cambia la media de una distribución normal con varianza igual a uno, es igual en distribución al máximo de un puente Browniano normalizado. Derivaron también la convergencia débil de la distribución nula de la estadística de prueba de razón de verosimilitud del punto de cambio sobre la media de una distribución normal cuando la varianza es conocida y cuando la varianza es desconocida; por su parte Horváth (1993) obtuvo la distribución nula asintótica de la estadística de prueba de la razón de verosimilitud en el caso de una normal univariada cuando ambas media y varianza pueden cambiar, en éstos tres casos la convergencia fue hacia la distribución doble exponencial. Yao y Davis (1986), determinaron que dicha distribución para n pequeña no provee una buena aproximación a la distribución nula, para esto se basaron en resultados de la teoría de valores extremos expuesta en Hall (1979), la cual dice que si se cuenta con una sucesión de variables aleatorias provenientes de una distribución normal estándar, entonces el máximo de tal sucesión estandarizado por funciones a_n y b_n converge en distribución a la distribución de valores extremos Gumbel, y la tasa de convergencia es aproximadamente $1/\log n$ lo cual indica una convergencia muy lenta principalmente en muestras pequeñas. Horváth (1993) determinó también por medio de simulación que la convergencia es lenta para muestras pequeñas, tomó muestras de tamaños $n = 20, 50$ y 100 . La estadística de prueba para la media, la convergencia en distribución a puentes Brownianos y la distribución Gumbel bajo H_0 presentadas en Yao y Davis (1986), son dadas a continuación:

Estadística de prueba de máxima verosimilitud

Enseguida se incluye la estadística de prueba de máxima verosimilitud, de la cual se obtiene la igualdad en distribución a puentes Brownianos y la distribución asintótica Gumbel.

$$U = \sqrt{V_k} = \max_{1 \leq k \leq n-1} |T_k|, \text{ donde ,}$$

$$T_k = \left(\frac{n}{k(n-k)} \right)^{1/2} \sum_{i=1}^k (x_i - \bar{x}).$$

n es el tamaño de muestra.

Convergencia en distribución a puentes Brownianos

Teorema 1.1. Sean $W_k = X_1 + X_2 + \dots + X_k$, $1 \leq k \leq n-1$, donde las variables aleatorias X_k tienen distribución normal y $\{B(t); 0 \leq t < \infty\}$ un movimiento Browniano estándar; entonces de las propiedades de la variable aleatoria normal, ver Yao y Davis (1986),

$$\left\{ \frac{W_k - k\mu}{\sqrt{n}}; 1 \leq k \leq n \right\} \stackrel{D}{=} \left\{ B\left(\frac{k}{n}\right); 1 \leq k \leq n \right\},$$

($\stackrel{D}{=}$) significa igualdad en distribución. Además,

$$\begin{aligned} U &= \max_{1 \leq k \leq n-1} \left| \frac{W_k}{\sqrt{n}} - \frac{k}{n} \frac{W_n}{\sqrt{n}} \right| / \left[\frac{k}{n} \left(1 - \frac{k}{n}\right) \right]^{1/2} \\ &= \max_{nt=1, \dots, n-1} \left| \frac{W_k}{\sqrt{n}} - t \frac{W_n}{\sqrt{n}} \right| / [t(1-t)]^{1/2} \\ &= \max_{nt=1, \dots, n-1} \left| \frac{W_k}{\sqrt{n}} - \frac{ku}{\sqrt{n}} - t \left(\frac{W_n}{\sqrt{n}} - \frac{nu}{\sqrt{n}} \right) \right| / [t(1-t)]^{1/2} \\ &\stackrel{D}{=} \max_{nt=1, \dots, n-1} |B(t) - tB(1)| / [t(1-t)]^{1/2} \\ &= \max_{nt=1, \dots, n-1} |B_0(t)| / [t(1-t)]^{1/2}, \end{aligned}$$

donde $t = \frac{k}{n}$, $B_0(t) = B(t) - tB(1)$ es un puente Browniano Yao y Davis (1986).

Distribución de valores extremos Gumbel

Teorema 1.2. Para $-\infty < x < \infty$

$$\lim_{n \rightarrow \infty} P[a_n^{-1}(U - b_n) \geq x] = \exp\{-2\pi^{1/2}e^{-x}\},$$

donde $a_n = (2 \log \log n)^{-1/2}$, $b_n = a_n^{-1} + \frac{1}{2}a_n \log_3 n$. U es la estadística de prueba de máxima verosimilitud.

En lo que respecta a la Estadística Multivariada Srivastava y Worsley (1986) usaron una prueba de razón de verosimilitud para probar un cambio en el vector de medias de una distribución normal multivariada suponiendo varianzas iguales pero desconocidas. La prueba de razón de verosimilitud para t (punto de cambio desconocido) está basada en el máximo de la estadística T^2 de Hotelling:

$$T_{\hat{t}}^2 = \text{máx } T_t^2,$$

donde \hat{t} es el punto en el cual ocurre el máximo. En la expresión de la estadística $T_{\hat{t}}^2$ que se presenta más adelante, se tiene a $y_t = \{N_t N_t^* / N\}^2 (\bar{x}_t - \bar{x}_t^*)$ donde N_t es el número de observaciones antes del punto de cambio, N_t^* es el número de observaciones después del punto de cambio y N es el número total de observaciones, lo que se observa entre paréntesis es la diferencia estandarizada entre las observaciones antes y después del punto de cambio, y la estadística T^2 de Hotelling para probar estas diferencias es $T_t^2 = y_t' W_t y_t$, $t = 1, \dots, n-1$, donde,

$$W_t = \left\{ \sum_{i=1}^{N_t} (x_i - \bar{x}_t)(x_i - \bar{x}_t)' + \sum_{i=N_t}^N (x_i - \bar{x}_t^*)(x_i - \bar{x}_t^*)' \right\} / (N-2),$$

es la matriz de varianza muestral unida. Srivastava y Worsley (1986) encontraron una técnica conservativa para la distribución nula de $T_{\hat{t}}^2$, basada en una desigualdad Bonferroni mejorada, para ello trabajaron en términos de la siguiente estadística equivalente:

$$S_t = T_t^2 / (N-2+T_t^2) = y_t' V^{-1} y_t, \text{ donde } V = \sum_{i=1}^N (x_i - \bar{x}_N)(x_i - \bar{x}_N)'$$

Srivastava y Worsley (1986) aportaron la desigualdad Bonferroni mejorada, ellos se basaron en dos propuestas enumeradas a continuación:

1. Desigualdad Bonferroni $P(S_{\hat{t}} > c) \leq \sum_{t=1}^{N-1} P(E_t)$, donde E_t son los eventos tal que $S_t > c$, con $c > 0$ y $c \in Z$.

2. $P(S_{\hat{t}} > c) \leq P_{\mu}$, donde $P_{\mu} = \sum_{t=1}^{N-1} P(E_t) - \sum_{t=1}^{N-1} P(E_t \cap E_{t+1})$.

Su aporte:

3. $P(S_{\hat{t}} > c) \leq P_{\mu} \approx 1 - G_{p,v}(c) + q_1 \sum_{t=1}^{N-2} k_t - q_2 \sum_{t=1}^{N-2} k_t^3$, donde

$$k_t = (1 - \rho_t)^{1/2},$$

$$q_1 = g_{p,v} \{2c(1-c)/\pi\}^{1/2} \Gamma\{\frac{1}{2}(p+v-1)\} / \Gamma\{\frac{1}{2}(p+v-1)\},$$

$$q_2 = q_1 \{(p^2 - 1)/c + (v^2 - 1)/(1 - c) - (p + v)(p + v - 1)\}/12(p + v)$$

y $g_{p,v}$ es la densidad de una variable aleatoria beta con parámetros $\frac{1}{2}p$, $\frac{1}{2}v$, $v = N - p - 1$, la expansión es válida si $v \geq 8$.

Su resultado fue obtenido al expandir la probabilidad conjunta en series de Taylor sobre $\rho_t = 1$, $P = P(E_t \cap E_{t+1}) = P(S_t > c, S_{t+1} > c)$.

Encontraron la distribución nula conjunta de S_i y S_j bajo H_0 y mediante integración numérica obtuvieron la distribución de la estadística de prueba $T_i^2 = \max\{y_t'W_t y_t\}$, para $p = 2, 4, 6$ (número de variables usadas) y $N = 10, 20, 40$ (N tamaño de muestra), los niveles de significancia utilizados fueron $\alpha = 0.10, 0.05, 0.01$. Calcularon también los límites para las dos propuestas en las que se basaron y los resultados mostraron que: el límite para la primera debía ser usado para $v < 8$ (v parámetro de forma de la distribución beta usada en su aproximación), su propuesta fue razonablemente exacta para $N = 20, 40$ y $v \geq 8$ y ambos fueron conservativos para $N > 40$, ver Srivastava y Worsley (1986). El segundo límite en el que ellos se basaron fue exacto para $N < 50$ y conservativo para $N > 50$. Además notaron que la exactitud parece incrementar cuando el número de variables p crece. Mencionan además que su método es útil cuando se tienen múltiples puntos de cambio pues puede aplicarse el método de bisección.

Debido a que las distribuciones de las diferentes estadísticas de prueba resultaron ser adecuadas para muestras pequeñas y en algunos casos con errores; para muestras suficientemente grandes en donde la convergencia resultó ser lenta y en otro caso los límites resultaron ser conservativos, se creó una metodología nueva para convergencia asintótica en donde se utilizan puentes Brownianos. Esta nueva metodología demostró tener una convergencia asintótica rápida, ser no conservativa y ser apropiada para tamaños de muestra pequeños, moderados y grandes, ver Srivastava y Worsley (1986).

1.2 Problemas de puntos de cambio espaciales

En esta sección se tratan puntos de cambio espaciales modelados en forma lattice (o modelo areal), donde una lattice es una cuadrícula que puede ser regular o irregular.

Según Xun *et al.* (2014), los tipos de modelado en Estadística Espacial son de tres clases: modelado Geoestadístico; modelado Lattice (o

modelo Areal), en este tipo de modelado se utilizan procesos de los datos tales como el autorregresivo espacial y los campos aleatorios de Markov; y finalmente Procesos Puntuales.

Un trabajo en modelado Lattice con procesos autorregresivos multivariados para estudiar puntos de cambio de forma espacial es el de Otto y Wolfgang (2016), quienes analizan cambios en los parámetros de la media y covarianza de tal proceso. El proceso tiene un punto inicial el cual es considerado como el centro de origen $s_0 \in D_s$; D_s denota la región de estudio y desde el cual evoluciona en toda dimensión del espacio q -dimensional, los cambios en los parámetros ocurren a una cierta distancia desde el centro.

Sean s_1, \dots, s_n , localizaciones en D_s , las cuales son ordenadas con respecto a su distancia desde el origen como $0 < d(s_1, s_0) \leq d(s_2, s_0) \leq \dots \leq d(s_n, s_0)$. Sea $\mathbf{Y}_i(s)$ la i -ésima componente de $\mathbf{Y}(s)$ y $\mathbf{Y}_i(s) = (Y_i(s_1), \dots, Y_i(s_n))^T$. El proceso autorregresivo considerado es dado por:

$$\mathbf{Y}_i = u_i \mathbf{1}_n + \rho_i B(\mathbf{Y}_{i-1} - u_i \mathbf{1}_n) + \boldsymbol{\xi}_i, \quad i = 1, \dots, p,$$

donde $\mathbf{1}_n$ es un vector de unos, $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_p\}$ es una colección de vectores aleatorios independientes, ρ_i es un parámetro autorregresivo de ponderación de $B\mathbf{Y}_i$, u_i es la media de Y_i , B es una matriz de ponderaciones espaciales. Asumiendo que cada $\xi_i \sim N_n(\mathbf{0}, \sigma_{\xi_i}^2 \mathbf{I}_n)$, se sigue que $\mathbf{Y}_i \sim N_n(u_i \mathbf{1}_n, (\mathbf{I}_n - \rho_i B)^{-1} \sigma_{\xi_i}^2 [(\mathbf{I}_n - \rho_i B)^T]^{-1})$, en donde \mathbf{I}_n es la matriz identidad n -dimensional. Se asume que los vectores $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_p$ e $\mathbf{Y}_1, \dots, \mathbf{Y}_p$, son independientes, pero los componentes de cada uno de los \mathbf{Y}_i son correlacionados. Por otra parte, la distribución de la l -ésima observación es $\mathbf{Y}(s_l) \sim N_p(\mathbf{u}, \boldsymbol{\Sigma}(s_l))$, donde $\boldsymbol{\Sigma}(s_l) = \text{diag}(\sigma_1^2(s_l), \dots, \sigma_p^2(s_l))$, $l = 1, \dots, p$ y $\mathbf{u} = (u_1, \dots, u_p)^T$. Todos los componentes de $\mathbf{Y}(s_l)$ se asumen que independientemente siguen un proceso autorregresivo espacial. La varianza $\sigma_i^2(s_l)$ es obtenida como la l -ésima entrada de la diagonal de la matriz de covarianza $(\mathbf{I}_n - \rho_i B)^{-1} \sigma_{\xi_i}^2 [(\mathbf{I}_n - \rho_i B)^T]^{-1}$.

Se considera que puede ocurrir un cambio en los parámetros del modelo a una distancia desconocida δ desde el origen $s_0 = 0$. Se asume que $\delta \in D_n = \{D(s_0), D(s_1), \dots, D(s_n), \infty\}$, y D_n tiene dos elementos si todas las estaciones tienen la misma distancia positiva desde el origen (i.e. $0 = D(s_0) < D(s_1) = \dots = D(s_n) < \infty$) y tiene $n + 1$ elementos si todas las estaciones tienen una distancia positiva diferente desde el origen (i.e. $0 = D(s_0) < D(s_1) < \dots < D(s_n) < \infty$) donde $\delta = \infty$ se refiere al caso en que no existe un cambio estructural. En particular el enfoque es sobre cambios en la media y los parámetros autorregresivos $\boldsymbol{\rho} = (\rho_1, \dots, \rho_p)^T$.

Si $\{\mathbf{X}(s) : s \in D_s\}$ es el proceso observado y se supone que ocurre un cambio $\mathbf{a} = (a_1, \dots, a_p)^T \in \mathbb{R}^p \setminus \{0\}$ a una distancia δ , entonces el proceso observado y su esperanza son:

$$\mathbf{X}(s) = \begin{cases} \mathbf{Y}(s), & \text{si } D(s) < \delta, \\ \mathbf{a} + \mathbf{Y}(s), & \text{si } D(s) \geq \delta, \end{cases}$$

y

$$E(\mathbf{X}(s)) = \begin{cases} \mathbf{u}, & \text{si } D(s) < \delta, \\ \mathbf{u} + \mathbf{a}, & \text{si } D(s) \geq \delta. \end{cases}$$

para cada localización s .

Se define una función indicadora $\psi(d) = (I_{D(s_l) \geq d})_{l=1, \dots, n}$ para la distancia y al incluir el cambio en la media del proceso autorregresivo, el proceso observado se reescribe como:

$$\mathbf{X}_i = u_i \mathbf{1}_n + a_i (\mathbf{I}_n - \rho_i B) \psi(\delta) + \rho_i B (\mathbf{X}_i - u_i \mathbf{1}_n) + \boldsymbol{\xi}_i, \quad i = 1, \dots, p.$$

Además, puede haber cambios en los parámetros autorregresivos del proceso y si $\mathbf{r} = (r_1, \dots, r_p)^T \in \mathbb{R}^p \setminus \{0\}$ denota la magnitud de dicho cambio el cual ocurre a la distancia δ , entonces el proceso se especifica como:

$$\mathbf{X}_i = u_i \mathbf{1}_n + (\rho_i \mathbf{I}_n + r_i \text{diag}(\psi(\delta))) B (\mathbf{X}_i - u_i \mathbf{1}_n) + \boldsymbol{\xi}_i, \quad i = 1, \dots, p.$$

Los cambios en los parámetros autorregresivos conducen a cambios en la matriz de covarianza del proceso.

De acuerdo con Otto y Wolfgang (2016), el contraste de hipótesis sobre la decisión de si existe un cambio estructural dentro del proceso a una distancia maximal δ de todas las localizaciones se especifica como:

$$H_0 : \delta > D(s_n) \quad \text{vs} \quad H_1 : \delta \leq D(s_n),$$

donde $D(s_n)$ son todas las distancias posibles medidas a partir del origen a cualesquiera localizaciones de la región de estudio. La decisión de si existe un cambio se basa en la razón de verosimilitud entre el modelo con $\delta = \infty$ y el modelo mejor ajustado con $\delta > 0$. Asumiendo que los vectores $\boldsymbol{\xi}_i$ son distribuidos normalmente, las log verosimilitudes son:

Bajo H_0 :

$$L^{(0)}(\boldsymbol{\rho}, \boldsymbol{\sigma}_\xi, \mathbf{u}; \mathbf{X}_1, \dots, \mathbf{X}_p) = -\frac{n}{2} \sum_{l=1}^p \ln(2\pi\sigma_{\xi_i}^2) \\ - \sum_{l=1}^p \frac{1}{2\sigma_{\xi_i}^2} \boldsymbol{\xi}_i^{(0)T} \boldsymbol{\xi}_i^{(0)} \text{ donde } \boldsymbol{\xi}_i^{(0)} = (\mathbf{I}_n - \rho_i B)(\mathbf{X}_i - u_i \mathbf{1}_n).$$

Para cambio en la media:

$$L^{(1)}(\boldsymbol{\rho}, \boldsymbol{\sigma}_\xi, \mathbf{u}, \mathbf{a}, \delta; \mathbf{X}_1, \dots, \mathbf{X}_p) = -\frac{n}{2} \sum_{l=1}^p \ln(2\pi\sigma_{\xi_i}^2) \\ - \sum_{l=1}^p \frac{1}{2\sigma_{\xi_i}^2} \boldsymbol{\xi}_{i\delta}^{(1)T} \boldsymbol{\xi}_{i\delta}^{(1)} \\ \text{donde } \boldsymbol{\xi}_{i\delta}^{(1)} = (\mathbf{I}_n - \rho_i B)(\mathbf{X}_i - u_i \mathbf{I}_n - a_i \boldsymbol{\psi}(\delta)).$$

Para cambio en el parámetro autorregresivo:

$$L^{(2)}(\boldsymbol{\rho}, \boldsymbol{\sigma}_\xi, \mathbf{u}, \mathbf{r}, \delta; \mathbf{X}_1, \dots, \mathbf{X}_p) = -\frac{n}{2} \sum_{l=1}^p \ln(2\pi\sigma_{\xi_i}^2) \\ - \sum_{l=1}^p \frac{1}{2\sigma_{\xi_i}^2} \boldsymbol{\xi}_{i\delta}^{(2)T} \boldsymbol{\xi}_{i\delta}^{(2)} \\ \text{donde } \boldsymbol{\xi}_{i\delta}^{(2)} = (\mathbf{I}_n - (\rho_i \mathbf{I}_n + r_0 \text{diag}(\boldsymbol{\psi}(\delta)))B)(\mathbf{X}_i - u_i \mathbf{1}_n).$$

La razón de log verosimilitud es:

$$-2\Lambda_n(\mathbf{X}_1, \dots, \mathbf{X}_p) = \max_{\delta \in \mathcal{D}_n} 2(L^{(1)}(\hat{\boldsymbol{\theta}}_\delta^{(1)}; \mathbf{X}_1, \dots, \mathbf{X}_p) \\ - L^{(0)}(\hat{\boldsymbol{\theta}}^{(0)}; \mathbf{X}_1, \dots, \mathbf{X}_p)), \\ \text{donde } \hat{\boldsymbol{\theta}}^{(0)} \text{ denota el punto donde el máximo de la log verosimilitud } L^{(0)} \text{ es alcanzado y } \hat{\boldsymbol{\theta}}_\delta^{(1)} \text{ denota el máximo de la log verosimilitud del modelo de punto de cambio para algún punto de cambio dado } \delta.$$

Dado que la distribución límite exacta de la razón de log verosimilitud no pudo ser obtenida de una manera cerrada, se realizó un estudio de simulación del proceso por medio de Monte Carlo y los cuantiles de la distribución estimada de Λ_n para lo cual se utilizó un kernel Gaussiano, fueron comparados con los de la distribución Gumbel Xun *et al.* (2014).

1.3 Inferencia Bayesiana

En esta sección, se enfocará la atención sobre los temas Bayesianos tomando información de William(2007), Lee (2012) y Bayesiano con puntos de cambio de Altieri (2015). Esto se realizó con el objetivo de resolver alguna de las hipótesis de punto de cambio propuestas en

Altieri en el Capítulo 3 de este trabajo; ya que Altieri propone contrastes de hipótesis de puntos de cambio temporal, espacial y espacio temporal. Los temas Bayesianos analizados serán útiles para el análisis de los casos mencionados anteriormente y si es posible con alguna modificación en cuanto a la distribución a priori, ya que se utiliza para el análisis de los puntos de cambio en Altieri (2015) una distribución a priori uniforme, lo cual permite considerar la posibilidad de utilizar otra distribución a priori.

Así, en el Capítulo 3 donde se analiza el caso de estudio además de utilizar la distribución a priori uniforme, se utilizan la distribución log-gamma, la Gaussiana, la logit-beta y la normal truncada.

En la inferencia Bayesiana generalmente se requiere la distribución a posteriori para los modelos, por ejemplo, la distribución de los parámetros dados los datos, o distribución predictiva a posteriori (para predicción o pronóstico de la distribución de nuevos valores dados los observados).

La distribución a posteriori es igual a la verosimilitud de datos multiplicada por la distribución a priori sobre la constante de normalización (por lo que la distribución a posteriori se integra a uno), esto es, por el teorema de Bayes:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta} \propto p(y|\theta)p(\theta),$$

en un enfoque frecuentista, a menudo se maximiza la verosimilitud de los datos utilizando métodos numéricos como Newton Raphson, para obtener una estimación puntual para un parámetro dado (que es fijo pero desconocido) y se utiliza la idea del remuestreo teórico para estimar un intervalo de confianza correspondiente alrededor de ese parámetro estimado. En un análisis Bayesiano se obtiene una distribución a posteriori para el parámetro para la cual se pueden proporcionar resúmenes estadísticos (mediana, media y moda) y cuantiles para obtener directamente intervalos de alta credibilidad.

Si se usan distribuciones a prioris conjugadas para cada parámetro (una especie de distribuciones canónicas que facilitan los cálculos), entonces se puede determinar la distribución a posteriori en forma cerrada Lee (2012). Generalmente, la integral en el denominador es intratable, por lo que se utilizan métodos numéricos como los Métodos de Montecarlo basados en cadenas de Markov (MCMC), para tomar muestras de las distribuciones condicionales y estimar la distribución marginal para cada parámetro de interés.

Si la distribución a priori pertenece a una familia conjugada, la distribución a posteriori puede calcularse analíticamente, de lo contrario se recurre a métodos numéricos. Una definición de una distribución a priori conjugada se da a continuación.

Definición 1.1. Una distribución a priori es conjugada para una familia de distribuciones si la distribución a priori y la distribución a posteriori son de la misma familia Lee (2012).

Las distribuciones a priori conjugadas son usuales debido a que siempre se obtiene la distribución a posteriori en forma analítica, así existe una conveniencia matemática. La familia exponencial incluye muchas distribuciones comunes (normal, Gamma, Poisson, Binomial, etc.), una distribución a priori conjugada para la familia exponencial tiene la forma de una distribución de la familia exponencial.

Dos ejemplos de distribuciones a prioris conjugadas son la Binomial y la Poisson. La verosimilitud Binomial está dada de la siguiente forma:

$$p(x|\pi) \propto \pi^x (1 - \pi)^{(n-x)},$$

la distribución a priori para π tiene la forma:

$$p(\pi) \propto \pi^{\alpha-1} (1 - \pi)^{(\beta-1)},$$

para $0 \leq \pi \leq 1$, esto es tiene una distribución Beta $\pi \sim Be(\alpha, \beta)$,

entonces la distribución a posteriori tiene la forma:

$$p(\pi|x) \propto \pi^{\alpha+x-1} (1 - \pi)^{\beta+n-x-1},$$

esto es $\pi|x \sim Be(\alpha + x, \beta + n - x)$, es claro que la familia de la distribución Beta es conjugada para una verosimilitud Binomial.

Para una distribución Poisson la verosimilitud es:

$$l(\lambda|x) \propto \lambda^T \exp(-n\lambda),$$

donde T es la estadística suficiente $T = \sum x_i$. La distribución a priori conjugada es :

$$p(\lambda) \propto \lambda^{v/2-1} \exp(-1/2S_0\lambda).$$

La densidad a posteriori es:

$$p(\lambda|x) \propto \lambda^{(v+2T)/2-1} \exp(-1/2(S_0 + 2n)\lambda).$$

Para este ejemplo la familia conjugada es la Gamma.

Además, también para la verosimilitud Poisson se tiene las distribución a priori uniforme que es $g(\mu) = 1$ para $\mu > 0$, y la distribución a priori de Jeffrey $g(\mu) \propto \frac{1}{\sqrt{\mu}}$, para $\mu > 0$.

Teorema de Bayes para la media normal con una distribución a priori continua.

Para una distribución normal, la verosimilitud de la muestra aleatoria es proporcional a la verosimilitud de la media muestral \bar{y} . Así la distribución a posteriori es

$$g(\mu|y_1, \dots, y_n) = \frac{g(\mu) e^{-\frac{1}{2\sigma^2/n}(\bar{y}-\mu)^2}}{\int g(\mu) \frac{1}{2\sigma^2/n}(\bar{y}-\mu)^2 d\mu}.$$

Esto funciona para cada distribución a priori continua $g(\mu)$. Sin embargo requiere integración que puede hacerse numéricamente, ver William (2007).

La distribución a priori plana da a cada valor μ igual peso. La distribución a priori plana no es realmente una distribución a priori adecuada ya que $-\infty < \mu < \infty$, por lo que no puede integrarse a 1. Sin embargo, esta distribución a priori impropia funciona bien. Aunque la distribución a priori es impropia, la distribución a posteriori integrará a 1, por lo que es apropiada. La distribución a priori de Jeffrey para la media de una distribución normal resulta ser la distribución a priori plana.

La distribución a posteriori para la media se reescribe como $g(\mu|\bar{y}) \propto e^{-\frac{1}{2\sigma^2/n}(\mu-\bar{y})^2}$, así la distribución a posteriori es normal con media \bar{y} y varianza σ^2/n .

Densidad a priori normal para μ .

La densidad a priori es $g(\mu) \propto e^{-\frac{1}{2s^2}(\mu-m)^2}$ y la forma de la verosimilitud es $f(y|\mu) \propto e^{-\frac{(y-\mu)^2}{2\sigma^2}}$. Realizando el producto y eliminando

la constante de proporcionalidad, se tiene que la distribución a posteriori es una distribución normal que tiene media y varianza dadas por:

$$m' = \left(\frac{\sigma^2 m + s^2 y}{\sigma^2 + s^2} \right) \quad y \quad (s')^2 = \frac{\sigma^2 s^2}{(\sigma^2 + s^2)}.$$

Esto demuestra que la distribución normal a priori $N(m, s^2)$ pertenece a la familia conjugada para la distribución de observación normal con varianza conocida.

Intervalo creíble Bayesiano para la media Normal

Teorema 1.3. Cuando la varianza es conocida, el intervalo de confianza Bayesiano de $(1-\alpha)$ % para μ es $m' \pm z_{\frac{\alpha}{2}} s'$, ver William (2007).

Teorema 1.4. Cuando la varianza es desconocida, el intervalo Bayesiano es $m' \pm t_{\frac{\alpha}{2}} s'$, ver William (2007).

Distribución a priori no normal Cuando la distribución a priori es no normal, se encuentra la distribución a posteriori para μ usando el teorema de Bayes y se tiene que integrar numéricamente. La distribución a posteriori será no normal. El intervalo $(1 - \alpha)$ % creíble al encontrar un valor inferior μ_l y superior μ_u , tales que

$$\int_{\mu_l}^{\mu_u} g(\mu|y_1, \dots, y_n) d\mu = 1 - \alpha.$$

Hay muchos de tales valores, la mejor opción μ_l y μ_u nos daría el intervalo más corto posible. Estos valores también satisfacen:

$$g(\mu_l|y_1, \dots, y_n) = g(\mu_u|y_1, \dots, y_n).$$

A veces es más fácil encontrar el intervalo creíble con áreas de cola inferior y superior que son iguales.

Inferencia de la media Bayesiana y frecuentista

En la estadística clásica hay varios tipos distintos de inferencia que se pueden hacer: estimación puntual, estimación por intervalo y pruebas de hipótesis. Cada uno de estos tipos de inferencia se pueden hacer

de forma Bayesiana.

Desde el punto de vista frecuentista, el estimador de la media μ es el promedio \bar{Y} , con esta estadística el estimador es insesgado y de varianza mínima. El estimador Bayesiano para μ es:

$$\hat{\mu}_B = E(\mu|y_1, \dots, y_n) = \frac{1/s^2}{n/\sigma^2 + 1/s^2}xm + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2}x\bar{y}.$$

Se sabe que la media a posteriori minimiza el cuadrado medio a posteriori, esto es $\hat{\mu}_B$ es el estimador óptimo para la media dados nuestros datos de la muestra William (2007). La media a posteriori es una función lineal de la variable aleatoria \bar{Y} , así su valor esperado es

$$E(\hat{\mu}_B) = \frac{1/s^2}{n/\sigma^2 + 1/s^2}xm + \frac{n/\sigma^2}{n/\sigma^2 + 1/s^2}x\mu.$$

El sesgo de la media a posteriori es su valor esperado menos el valor verdadero del parámetro el cual simplifica a $\frac{\sigma^2}{ns^2 + \sigma^2}(m - \mu)$,

la varianza de la media a posteriori es $[\frac{n/\sigma^2}{n/\sigma^2 + 1/s^2}]^2 x \frac{\sigma^2}{n}$.

Esta es más pequeña que $\frac{\sigma^2}{n}$, la cual es la varianza del estimador \bar{Y} ; el estimador Bayesiano tiene un error cuadrado medio más pequeño.

Relación entre un intervalo frecuentista e intervalo creíble Bayesiano de la aproximación a priori plana.

Con una distribución a priori plana para μ , la media a posteriori es igual a $m' = \bar{y}$, y la varianza a posteriori es igual a $(s')^2 = \frac{\sigma^2}{n}$. En este caso el intervalo creíble frecuentista y el intervalo creíble Bayesiano tienen la misma forma:

$$\bar{y} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}.$$

La interpretación frecuentista es que μ es fija. Los puntos finales del intervalo aleatorio se calculan utilizando una declaración de probabilidad sobre la distribución de muestreo de la estadística \bar{y} . En la interpretación Bayesiana se permite que μ sea una variable aleatoria. El intervalo creíble se calcula a partir de la distribución a posteriori dados los datos de muestras que ocurrieron, tiene la probabilidad condicional establecida de contener μ dados los datos.

Inferencia Bayesiana para regresión lineal simple

A veces es necesario modelar una relación entre dos variables x y y . Es posible que se requiera encontrar una ecuación que describa tal relación. A menudo se planea usar el valor de x para ayudar a predecir y usando esta relación. Los datos consisten en n ordenadas (x_i, y_i) para $i = 1, \dots, n$. x es considerada la variable predictora y se conoce sin error y y es una variable de respuesta que depende de x de alguna manera desconocida, cada y observado contiene un término de error.

La verosimilitud de la observación i es:

$$l_i(\alpha_{\bar{x}}, \beta) \propto e^{-\frac{1}{2\sigma^2}[y_i - (\alpha_x + \beta(x_i - \bar{x}))]^2},$$

esto siempre es así.

La verosimilitud de toda la muestra de todas las observaciones es el producto de la verosimilitudes individuales

$$l_{muestra}(\alpha_{\bar{x}}, \beta) \propto \prod_{i=1}^n e^{-\frac{1}{2\sigma^2}[[y_i - (\alpha_x + \beta(x_i - \bar{x}))]^2]}.$$

El exponente se simplifica en $SS_y = \sum_{i=1}^n (y_i - \bar{y})^2$, y $SS_{xy} = \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})$, y $SS_x = \sum_{i=1}^n (x_i - \bar{x})^2$.

La verosimilitud eliminando la constante de proporcionalidad, es el producto de dos funciones de masa asociadas con sendas distribuciones normales y el resultado es el siguiente:

$$l_{muestra}(\alpha_{\bar{x}}, \beta) \propto e^{-\frac{1}{2\sigma^2/SS_x}[\beta - \frac{SS_{xy}}{SS_x}]^2} e^{-\frac{1}{2\sigma^2}[(\alpha_{\bar{x}} - \bar{y})^2]}.$$

Nótese que la pendiente de mínimos cuadrados es $\beta = \frac{SS_{xy}}{SS_x}$ y $\bar{y} = A_x$, el estimador de mínimos cuadrados del intercepto de la línea vertical $x = \bar{x}$. La distribución a priori conjunta para β y $\alpha_{\bar{x}}$ es

$$g(\alpha_{\bar{x}}, \beta) = g(\alpha_{\bar{x}})g(\beta).$$

La distribución a posteriori conjunta para β y α_x es proporcional a la distribución a priori conjunta por la verosimilitud conjunta, esto es:

$$g(\alpha_{\bar{x}}, \beta | \text{datos}) \propto g(\alpha_{\bar{x}}, \beta) l_{\text{muestra}}(\alpha_{\bar{x}}, \beta).$$

Modelos Gaussianos Latentes

Los modelos Gaussianos Latentes son una clase muy general de modelos Bayesianos jerárquicos donde se supone que la variable de respuesta pertenece a una familia exponencial y es condicionalmente independiente dado un campo latente (normalmente distribuido) y algunos hiperparámetros.

El modelo jerárquico puede ser escrito como:

$$y | (\eta, \theta_2) \sim \pi(y | \eta, \theta) = \prod_i p(y_i | \eta_i, \theta_2), \text{ verosimilitud.}$$

$$\eta | \theta_1 \sim p(\eta | \theta_1) = N(0, \Sigma), \text{ campo latente.}$$

$$\theta = [\theta_1, \theta_2]^T \sim p(\theta), \text{ hiperparámetros.}$$

Donde \mathbf{y} es un conjunto de datos observados, η es la distribución conjunta de todos los parámetros en el predictor lineal (incluido el mismo), θ son los hiperparámetros del campo latente que no son Gaussianos.

La distribución marginal de cada parámetro es:

$$(\eta_i | \eta_{-i}) \sim N\left(\mu_i - \frac{1}{Q_{ii}} \sum_{j \neq i} Q_{ij} (\eta_j - \mu_j), Q_{ii}^{-1}\right),$$

donde η_{-i} son los otros parámetros que no son i , combinando los 3 niveles la distribución a posteriori conjunta resulta en:

$$\pi(\eta, \theta | \mathbf{y}) \propto \pi(\theta) N(0, \Sigma) \prod_i \pi(y_i | \eta_i, \theta),$$

el análisis tiene como objetivo encontrar las distribuciones a posteriori marginales para todos los elementos del campo latente.

La distribución normal en el nivel del campo latente puede sustituirse por algunas otras distribuciones, algunas de ellas son los Campos Aleatorios Gaussianos de Markov Intrínsecos (IGMRF), que a menudo se denominan CAR intrínseco o paseo aleatorio en dos dimensiones (RW2d) y el Besag. La densidad del primero es:

$$\pi(\mu | \tau_\mu) \propto \tau_\mu^{\frac{1}{2}(n-2)} \exp\left(-\frac{1}{2} \mu^T Q \mu\right),$$

donde la matriz de precisión es $Q = \tau_\mu R$. El parámetro τ_μ denota la precisión mientras que R refleja la estructura de vecindad específica del modelo. Mientras el modelo Besag está definido por la densidad conjunta

$$\pi(\mathbf{x}|k) \propto k^{(n-1)/2} \exp\left(-\frac{k}{2} \sum_{i \sim j} w_{ij} (x_i - x_j)^2\right),$$

donde $j \sim i$ significa que η_j pertenece a una vecindad de η_i y w_{ij} denota pesos positivos y simétricos para pares de nodos adyacentes. Se asumen incrementos independientes

$$X_i - X_j \sim N(0, 1/w_{ijk}),$$

y la matriz correspondiente a Q ahora tiene elementos:

$$Q_{ij} = \begin{cases} \sum_{k:k \sim i} w_{ik}, & i = j; \\ -k(w_{ij}), & i \sim j; \\ 0, & d.o.f. \end{cases} \quad (1.1)$$

Campos Aleatorios Gaussianos de Markov(GMRF)

Los Campos Aleatorios Gaussianos de Markov y las distribuciones marginales se sustituyen por distribuciones condicionales. Un campo aleatorio de Markov puede definirse como un conjunto de variables aleatorias que tienen la propiedad de Markov y cuando todas las variables se distribuyen normalmente se tiene un GMRF.

Se puede definir un GMRF por su media condicional y precisión, la media condicional es una suma ponderada de los vecinos, con pesos correspondientes a los valores en Q

$$E[\eta_i | \eta_{-i}] = \mu_i - \frac{1}{Q_{ii}} \sum_{j \sim i} Q_{ij} (\eta_j - \mu_j).$$

La precisión es

$$Prec[\eta_i | \eta_{-i}] = Q_{ii}.$$

Se pueden usar en diferentes tipos de campos aleatorios para modelar el efecto espacial en una red; uno de ellos es el GMRF Intrínseco

(IGMRF). Los modelos CAR a menudo se llaman CAR intrínsecos o paseos aleatorios en dos dimensiones (RW2D), y se caracterizan por una matriz de precisión que no es de rango completo (tiene al menos un valor propio cero).

El análisis de puntos de cambio en Altieri (2015) lo hacen con procesos Cox Log Gaussianos, estos procesos se definen de la siguiente forma:

El proceso Cox Log Gaussiano es un proceso puntual doblemente estocástico formado como un proceso de Poisson no homogéneo con una intensidad estocástica. Estos procesos fueron introducidos por Cox en una dimensión temporal. Su definición en el espacio y en el tiempo es:

1. $\{\Delta(s, t), s \in S, t \in T\}$ es un proceso estocástico no negativo.
2. Condicionando a $\{\Delta(s, t) = \lambda(s, t), s \in S, t \in T\}$, los eventos forman un proceso Poisson no homogéneo con intensidad $\lambda(s, t)$.

Supongamos que $z = \{Z(s, t), s \in S, t \in T\}$ es un proceso Gaussiano evaluado en los reales con media $\mu(s, t) = E[Z(s, t)]$ y función de covarianza $C[(s_i, t_i), (s_j, t_j)] = Cov[Z(s_i, t_i), Z(s_j, t_j)]$. Si la función de intensidad está definida como $\Delta(s, t) = \exp[Z(s, t)]$, entonces el correspondiente proceso Y es también un proceso Cox Log Gaussiano, ver Altieri (2015).

Los procesos Cox Log Gaussianos suponen que la distribución de puntos sobre el espacio se debe a la heterogeneidad ambiental estocástica modelada con una función de intensidad aleatoria $\Delta(s)$; dada $\Delta(s)$ la distribución de puntos sigue un proceso de Poisson no homogéneo. En los procesos Cox log Gaussianos se supone que el logaritmo de la superficie de intensidad sobre una ventana de observación W es un campo Gaussiano (latente) *sobre*(s), es decir $\Delta(s) = \int \lambda(s) ds = \exp(\eta(s))$, y condicionado sobre $\eta(s)$, el número de puntos es $N \sim Poi(\Delta(s))$.

La Estadística Espacial se divide en tres ramas: Geoestadística, Procesos Areales y Procesos Puntuales. Los que son de interés en inferencia Bayesiana son los Procesos Puntuales, estos tienen las siguientes características:

- $N(w)$: Es el número aleatorio de puntos del proceso en la observación de la ventana w .
- X son procesos puntuales definidos sobre un espacio medible y observado en el interior de la ventana w .

- $X_i, i = 1, \dots, N$: punto genérico (evento del proceso).
- $P(N(W = n))$, número de distribuciones univariadas.
- $E(N(w)) = \eta(w)$, número esperado de puntos en W .

La función de intensidad $\lambda(s) = \lim_{ds \rightarrow 0} \frac{E(N(ds))}{|ds|}$,

es decir, el número de puntos en una región extremadamente pequeña. La intensidad puede ser constante, en este caso el proceso es llamado uniforme/homogéneo, o inconstante, en cuyo caso el proceso es no uniforme/no homogéneo.

La intensidad es homogénea cuando el número de puntos en una región es:

$$N(w) \sim Poi(\lambda(w)),$$

por lo tanto,

$$P(N(w) = n) = \exp(-\lambda|w|) \frac{(\lambda|w|)^n}{n!},$$

esto implica que $\lambda(s) = \lambda$, es decir el número medio de eventos por unidad de área no depende de la localización s , la intensidad del proceso es no homogénea si,

$$N(w) \sim Poi(\Delta(w)), \text{ donde, } \Delta(s) = E(N(w)) = \int_w \lambda(s) ds,$$

y $\lambda(s)$ es la intensidad de primer orden en la localización s .

Una clase amplia de Procesos Puntuales está dada por los procesos de Cox. Los Procesos Cox Log Gaussianos son Procesos Puntuales de Cox donde se supone que el logaritmo de la superficie de intensidad es Gaussiana. La inferencia para estos modelos es históricamente muy difícil de tratar, sin embargo el procedimiento de Aproximación de Laplace Anidada Integrada (INLA) abre nuevas posibilidades para poder realizar la inferencia con estos procesos, ver Altieri (2015).

Para obtener estimaciones a posterioris se puede utilizar INLA, que es un método en el que se usa una distribución Gaussiana para aproximar una función de densidad de probabilidad dada (en el análisis Bayesiano la distribución a posteriori).

La INLA explota a la distribución Gaussiana y la expansión de la serie de Taylor de la aproximación de Laplace para obtener una expansión manejable y computacionalmente rápida de la distribución original.

Definición 1.2. Sea $\{\eta(s)\}_{s \in w}$ un campo aleatorio; este es un campo Gaussiano si y sólo si dados s_1, \dots, s_n , un conjunto de localizaciones y b_1, \dots, b_n un conjunto de números reales, tales que, $b_1\eta(s_1) + \dots + b_n\eta(s_n)$ es normalmente distribuido, los vectores $\eta(s_1), \dots, \eta(s_n)$ siguen una distribución normal multivariada para cualquier localización s .

Una transformación para definir una intensidad no negativa para un proceso Cox es:

$$\Delta(s) = \int_w \lambda(s) ds = \exp(\lambda(s)).$$

Por construcción $\{\eta(s)\}_s \in w$, es también un campo aleatorio.

La distribución de un LGCP X , se define a través de la distribución del campo Gaussiano $\{\eta(s)\}_s \in w$ que se especifica por su media, varianza y estructura de correlación (semidefinida positiva). Si el proceso es estacionario e isotrópico, la distribución conjunta (X, η) es invariante bajo movimientos rígidos.

La distribución de los LGCP estacionarios se caracteriza por su intensidad y la función de correlación por pares .

Bajo estacionariedad, $\mu = E(\Delta(0)) = \lambda$ es la media del campo de intensidad. Sea σ^2 la varianza y $C(r)$ la función de covarianza del campo latente a la distancia r (es estacionario, por lo que solo depende de la distancia), es decir $C(r) = \sigma^2 K(r) = Cov\{\eta(s), \eta(s-r)\}$. Por las propiedades de momentos de la distribución Lognormal, la intensidad de primer orden de un LGCP es $\lambda = E(\Delta(0)) = E(\exp(\eta(0))) = \exp(\mu + \frac{1}{2}\sigma^2)$ y la intensidad de covarianza es:

$$g(r) = \lambda^2[\exp(\sigma^2 K(r)) - 1].$$

Un LGCP espacio temporal puede ser definido como un proceso Poisson espacio temporal no homogéneo condicionado sobre una función de intensidad estocástica que varía en espacio y tiempo, la función de intensidad esta dada de la siguiente forma:

$\Delta(s, t) = \exp(\eta(s, t))$ donde $\eta(s, t)$ es un proceso Gaussiano.

La intensidad en modelos separables es descompuesta como

$$\lambda(s, t) = \lambda_0(s, t)R(s, t) = \lambda_0(s)\mu_0(t) \exp(\eta(s, t)),$$

donde $\lambda_0(s, t) = \lambda_0(s)\mu_0(t)$ es un producto de una componente puramente espacial y otra puramente temporal.

La correlación espacio temporal es $K(r, v) = K_s(r)K_t(v)$.

Obteniendo estimaciones a posterioris con INLA

Para obtener estimaciones a posterioris se utiliza la aproximación Laplace (INLA). El objetivo de INLA es proporcionar aproximaciones determinísticas precisas y rápidas a todas o al menos a algunas de las a posterioris marginales para los η_{is} .

La aproximación aplicada por INLA tiene como objetivo encontrar la distribución a posteriori marginal, es decir, la distribución condicional del campo latente η , dados los datos $\pi(\eta_i|y)$ y de θ , $\pi(\theta_i|y)$ sin integrarse sobre η_i . Los pasos para esto son los siguientes:

1. La aproximación de Laplace $\hat{\pi}$ de la a posteriori conjunta de los hiperparámetros:

$$\pi(\eta|y) = \frac{\pi(\eta, \theta, y)}{\pi(\eta|\theta, y)\pi(y)} = \frac{\pi(\eta, \theta|y)}{\pi(\eta|\theta, y)} \propto \frac{\pi(\eta, \theta|y)}{\pi_G(\eta|\theta, y)} \Big|_{\eta=\eta_\theta^*} \cong \pi(\theta, y),$$

π_G es una aproximación Gaussiana (Laplace) de $\pi(\eta|\theta, y)$ con la característica de hacer coincidir la verdadera distribución a posteriori en la moda. La verdadera distribución condicional completa del campo latente se aproxima mediante una distribución Gaussiana evaluada en la moda η_θ^* . Esta aproximación es muy precisa en la mayoría de los casos ya que $\pi(\eta|\theta, y)$ parece casi Gaussiana debido a la suposición de que la distribución a priori $\pi(\eta|\theta)$ es Gaussiana.

2. Se integra numéricamente para encontrar la marginal a posteriori (aproximada) $\tilde{\pi}(\theta_i|y)$.

3. Se integra fuera de θ para encontrar las marginales a posterioris $\pi(\eta_i|y)$. En la integración de Laplace de $\pi(\theta|y)$ (ya obtenida) se explota numéricamente para encontrar los puntos del soporte θ_p para la

integración numérica. Estos puntos tienen pesos de área Δ_p que están conectados a la suma que se aproxima a la integral:

$$\tilde{\pi}(\eta_i|y) = \int \tilde{\pi}(\eta_i|\theta, y)\tilde{\pi}(\theta|y)d\theta \approx \pi_G(\eta_i|\theta_p, y)\tilde{\pi}(\theta_p|y)\Delta_p.$$

La Aproximación de Laplace Anidada Integrada primero explora la distribución marginal conjunta a posteriori de los hiperparámetros θ para localizar la moda que se convertirá en la media de la distribución a posteriori Gaussiana aproximada.

Estimando modelos con INLA

Los modelos Gaussianos Latentes son una subclase de modelos de regresión aditiva, es decir, un subconjunto de todos los modelos aditivos Bayesianos con un predictor aditivo estructurado Altieri (2015).

Definición 1.3. Un modelo aditivo estructurado puede ser escrito como: $y_i = \mu_i + \xi_i$ con $g(\mu_i) = \beta_0 + \sum \beta_k z_{ki} + \sum w_{ji} f_j(z_{ji}) + \mu_i$, donde el campo latente es $\eta = \{\beta_0, \beta_k (K = 1, \dots, k), f_j (j = 1, \dots, J), \mu_i (i = 1, \dots, N)\}$, i.e. una colección de todos los parámetros aleatorios y específicamente.

1. β_0 es un intercepto.
2. β_k son efectos lineales fijos de covariables.
3. $f_j(z_{ji})$ son efectos espaciales suaves.
4. μ_i es un término de error no estructurado.
5. $\eta \sim N(0, Q^{-1})$ todos los efectos son asumidos normales con media cero.
6. Q debe tomarse en cuenta para cuando se tienen un numero de hiperparámetros menor o igual que 7.
7. $\beta \sim N(0, \tau_\beta^{-1}I)$, se supone que todos los efectos lineales, son independientes y dependen del mismo hiperparámetro.
8. La matriz de precisión general Q tiene valores diagonales no nulos que solo correspondan a los efectos estructurales no lineales.

9. $\mu \sim N(0, \tau_\mu^{-1}I)$, los términos de error lineal μ_i son iid y $\tau_\mu \sim Ga(\alpha, \beta)$. Note que $E(\tau_\mu) = \alpha\mu$ y $\sigma_\mu^2 = \frac{1}{\tau_\mu} \sim INVGA(\alpha, \beta)$, donde INVGA es una distribución Gamma Inversa. Esto significa que $E(\sigma_\mu^2) = \beta/\alpha - 1$ y $v(\sigma_\mu^2) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$.

Estimando Procesos Cox Log Gaussianos (LGCP) con INLA.

En el caso particular de LGCP, la log intensidad del proceso Poisson puede ser descrita por un predictor lineal como $y_i = \lambda_i + \xi_i \sim Poi(\lambda_i)$ donde $y_i = N(C_i)$. La aproximación del Proceso Puntual X por un proceso discreto Poisson y es bueno cuando las celdas son lo suficientemente pequeñas con:

$$\log(\lambda_i) = \beta_0 + \sum_{k=1}^K \beta_k z_{ki} + \sum_{j=1}^J w_j f_j(z_{ji}) + \mu_i,$$

donde β_0 es una intersección común y β_k son intersecciones específicas de tiempo y explican la variación en la intensidad a través del tiempo/espacio. El campo latente Gaussiano es $\eta = \{\beta_0, \beta_k, (k = 1, \dots, K), f_j(j = 1, \dots, J)\}$ y la distribución de todos los parámetros del campo se aproxima mediante un GMRF.

Introducción al análisis de puntos de cambio

Un punto de cambio se define como un lugar o un punto de tiempo τ en una serie y de manera que las observaciones siguen una distribución F_0 hasta ese punto y otra, F_1 después de ese punto. Se suponen datos ordenados de 1 a T y el modelo que los describe presenta algunos cambios abruptos, los datos se dividen en segmentos que siguen el mismo modelo pero diferentes especificaciones de parámetros. Los problemas de puntos de cambio se pueden desarrollar en diferentes niveles de complejidad por ejemplo: un punto de cambio con localización desconocida en τ_0 , número desconocido de puntos de cambio, con localizaciones desconocidas, y el número desconocido de puntos de cambio pueden ser $m = 1, \dots, M$, con localizaciones desconocidas τ_1, \dots, τ_m .

Distribuciones a priori

Hay dos clases posibles de escenarios de distribuciones a priori. El primero está estructurado en dos niveles: en primer lugar, se define una distribución a priori $\pi(m)$, $m = 0, \dots, M$, en el número de puntos de cambio y luego condicionado en m , se establece un conjunto de distribuciones a priori para las posiciones de los puntos de cambio, construidas donde cada punto de cambio en la distribución a priori depende de la siguiente posición del punto de cambio:

- 1) $\pi(m)$.
- 2) $\pi_m(\tau_1, \dots, \tau_m) = \pi_m(\tau_m)\pi_m(\tau_{m-1}|\tau_m) \dots \tau_m(\tau_1|\tau_2)$,

por definición se tiene $\tau_0 = 0$ y $\tau_{m+1} = T$.

El segundo ajuste de la distribución a priori consiste en una distribución a priori conjunta sobre el número y posición de puntos de cambio, construido modelando la ocurrencia de puntos de cambio mediante un proceso de punto discreto, se tendría un Proceso Puntual en dos niveles: un proceso temporal en el nivel de la distribución a priori, y un proceso espacio temporal en el nivel de los datos, el Proceso Puntual en el nivel de la distribución a priori se construye observando la función de densidad de masa $g(v)$ entre el tiempo V entre dos puntos de cambio sucesivos. Como la distribución es discreta su función de distribución acumulativa $G(v) = \sum_{\mu=1}^v g(\mu)$ será gradual y la distribución a priori es:

$$\pi_m(\tau_1, \dots, \tau_m) = g_0(\tau_1)(\prod_{j=2}^m g(\tau_j - \tau_{j-1}))(1 - G(\tau_{m+1} - \tau_m)),$$

donde $g_0(\tau_1)$ es la función de masa en el primer punto de cambio, como valor inicial para la serie de productos una opción natural para la distribución de v es la distribución Binomial negativa.

$$v \sim BN(k, p), g_0(v) = \sum_{i=1}^k \binom{v-i}{i-1} \frac{p^i (1-p)^{v-i}}{k}, g(v) = \binom{v-k}{k-1} p^k (1-p)^{v-k},$$

donde k es el número de puntos de cambio hasta que se detiene la sucesión y p es la probabilidad de cada juicio Bernoulli i.i.d..

Distribución a posteriori

Para cada m se obtiene una distribución a posteriori siguiendo la regla de Bayes y utilizando la verosimilitud aproximada dado un número m de puntos de cambio $L_0^{(m)}$, se obtiene:

$$\pi(m|y) \propto \pi(y|m)\pi(m) \approx L_0^{(m)}\pi(m).$$

La distribución a posteriori de m permite elegir el mejor número de puntos de cambio dados los datos, digamos \hat{M} . Si $\hat{M} \geq 1$ se eligen las mejores posiciones de puntos de cambio de la distribución a posteriori. Las posiciones más probables se encuentran utilizando las recursiones (ya calculadas) para $m = \hat{M}$ a través de la distribución condicional.

Puntos de cambio

Para los puntos de cambio se utiliza una distribución a priori en dos niveles, por lo tanto se establece una distribución a priori en el número (0 o 1) de puntos de cambio. Para obtener detalles de esta configuración de la distribución a priori, se da la misma masa de probabilidad a ambos valores.

$$\pi(M) = \begin{cases} 0.5, & M = 0, 1 \\ 0, & M > 1. \end{cases} \quad (1.2)$$

Se debe tener en cuenta que una distribución a priori uniforme implica que la razón de las distribuciones a priori $\frac{\pi(1)}{\pi(0)}$ es 1, por lo tanto cualquier cálculo que involucre la razón de las distribuciones a posteriori que es producto de razones de verosimilitud y distribuciones a priori puede simplificarse.

En el trabajo de Altieri (2015) se encuentran los puntos de cambio sobre la intensidad homogénea λ considerando los contrastes siguientes:

Modelo con efectos fijos

$$H_0 : \log\lambda(t) = \mu + \xi_t \text{ para } t = 1, \dots, T$$

vs

$$H_1 : \log\lambda(t) = \mu_1 + \xi_t \text{ para } t \leq T^*; \quad \log\lambda(t) = \mu_2 + \xi_t \text{ para } t > T^*,$$

donde μ_t es una intercepción específica en el tiempo, constante en el espacio pero que puede variar en el tiempo, su variación ocurre en correspondencia del punto de cambio, en una posición en el intervalo $[t_1, t_2]$ y ξ es un término de error no estructurado.

Modelo con efecto temporal

Se mantiene el supuesto de homogeneidad espacial, los datos pueden mostrar dependencia temporal del patrón de puntos de tiempo. La dependencia solo se permite dentro de segmentos de tiempo no entre segmentos. El contraste es:

$$H_0 : \log\lambda(t) = \mu + \phi + \xi_t, \text{ para, } t = 1, \dots, T$$

vs

$$H_1 : \log\lambda(t) = \mu_1 + \phi_1 + \xi_t, \text{ para, } t \leq T^*; \quad \log\lambda(t) = \mu_2 + \phi_2 + \xi_t, \text{ para, } t > T^*.$$

Dentro de cada segmento de tiempo, ϕ es un efecto aleatorio modelado como un modelo autorregresivo de orden 1 (AR(1)), es decir se supone que el logaritmo de la función de intensidad en cada punto de tiempo dependa de su propio valor en el momento anterior: $\psi_t = \rho\psi_{t-1} + \mu_t$ donde como es usual $|\rho| < 1$. Los hiperparámetros para la precisión cumplen que $v_\psi \sim \text{Gamma}(\alpha_\psi, \beta_\psi)$.

Este contraste de hipótesis es utilizado en el Capítulo 3 de este trabajo donde se trata sobre el tiempo y se programa para obtener los puntos de cambio con el factor de Bayes.

Modelo con efectos espaciales

El contraste es:

$$H_0 : \log\lambda(t, s) = \delta + \phi_s + \xi_{ts} \text{ para } t = 1, \dots, T \text{ y } s = 1, \dots, S$$

vs

$$H_1 : \log\lambda(t, s) = \delta + \phi_{1s} + \xi_{ts} \text{ para } t \leq T^* \text{ y } s = 1, \dots, S$$

$$\log\lambda(t, s) = \delta + \phi_{2s} + \xi_{ts} \text{ para } t > T^* \text{ y } s = 1, \dots, S,$$

donde δ es un intercepto común y ϕ_s describe la dependencia espacial, está indexado por s ya que puede tomar diferentes valores en la cuadrícula. Bajo H_1 un solo valor define la intensidad de cada celda en todo el primer segmento de tiempo, y después del punto de cambio cambia el valor de cada celda. El efecto espacial como una caminata aleatoria en dos dimensiones en una red (CAR) intrínseco. Aquí también el parámetro de precisión se puede modelar como $v_\phi \sim \text{Gamma}(\alpha_\phi, \beta_\phi)$.

Modelo general

Se incluyen ambos efectos, se permite la falta de homogeneidad espacial, la dependencia temporal dentro de segmentos y la dependencia espacial en cada momento. El contraste es:

$$H_0 : \log\lambda(t, s) = \lambda + \psi + \phi_s + \xi_{ts}, \text{ para, } t = 1, \dots, T \text{ y } s = 1, \dots, S.$$

vs

$$H_1 : \log\lambda(t, s) = \delta + \psi_1 + \phi_{1s} + \xi_{ts}, \text{ para, } t \leq T^* \text{ y } s = 1, \dots, S;$$

$$\log\lambda(t, s) = \delta + \psi_2 + \phi_{2s} + \xi_{ts}, \text{ para, } t > T^* \text{ y } s = 1, \dots, S.$$

Se supone que la dependencia temporal está dentro de los segmentos no a través de ellos. En un proceso no homogéneo, la dependencia AR(1) concierne a las celdas, la intensidad de cada celda depende de su valor anterior. El valor final estimado para ϕ es una síntesis de los valores de las celdas en el espacio. El parámetro de precisión para los efectos temporales y espaciales tiene una distribución a priori Gamma que está configurada por defecto

como no informativa.

Capítulo 2

Modelos clásicos espacio tiempo

En este Capítulo se tratan las diferentes clasificaciones de puntos de cambio: según Brodsky y Darkhovsky, visualización de puntos de cambio abrupto y gradual, puntos de cambio espacio tiempo según Xun *et al.* (2014). Posteriormente se tratan algunas metodologías paramétricas y no paramétricas y finalmente se presentan algunas aplicaciones.

2.1 Puntos de cambio según Brodsky y Darkhovsky

De acuerdo con Brodsky y Darkhovsky (1993, 2000), los problemas y métodos de diagnóstico de puntos de cambio se pueden clasificar de la manera siguiente:

- 1. Por el carácter de la información sobre el objeto de diagnóstico:** Análisis retrospectivo (a posteriori) y análisis secuencial.
- 2. Por el carácter de métodos de diagnóstico estadístico:** Métodos paramétricos, no paramétricos y semiparamétricos.
- 3. Por el carácter del objeto de diagnóstico:** Problemas de diagnóstico estadístico para procesos aleatorios (en tiempo discreto o continuo) y problemas de diagnóstico estadístico para campos aleatorios.

4. **Por el carácter de dependencia estadística entre observaciones:** se pueden formular problemas de puntos de cambio para sucesiones aleatorias con observaciones independientes, y problemas de puntos de cambio para observaciones dependientes en el tiempo o espacio, en forma unidimensional y multidimensional, con un solo punto de cambio o múltiples puntos de cambio.

5. **Por el mecanismo de cambio en el estado del objeto de diagnóstico:** Detección de cambio abrupto (problemas de puntos de cambio), detección de cambio gradual, detección en relaciones de regresión. Dichos cambios son observados en las características probabilísticas de las observaciones.

Los tipos de cambio que se tratan en este trabajo de tesis son, por el carácter de métodos de diagnóstico estadístico: paramétricos y no paramétricos y por el mecanismo de cambio en el estado del objeto de diagnóstico: Detección de cambio abrupto.

2.2 Puntos de cambio abrupto y gradual

Los puntos de cambio abrupto y gradual se visualizan en la Figura 2.1 la cual fue obtenida de Brodsky y Darkhousky (1993), $\psi_\theta(t)$ es la función con punto de cambio referente a los θ :

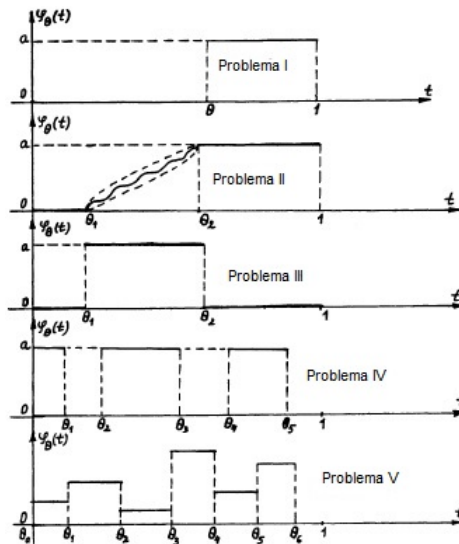


Figura 2.1: Gráfica de puntos de cambio.

Observando la Figura 2.1 de arriba hacia abajo se presentan los puntos de

cambio de la siguiente forma:

Problema I) Muestra un punto de cambio abrupto.

Problema II) Muestra un cambio gradual.

Problema III) Muestra dos cambios abruptos.

Problema IV) Muestra cinco cambios abruptos.

Problema V) Muestra seis cambios abruptos.

En el último problema en la Figura 2.1 el cambio ya no es homogéneo.

2.3 Puntos de cambio en espacio-tiempo

El problema de puntos de cambio en espacio-tiempo analiza diferentes tipos de cambios, según Xun *et al.* (2014), los cambios espacio temporales se clasifican de diferentes formas: cambio en parámetros estadísticos; cambio en el valor, es decir la diferencia entre un valor de un dato y sus vecinos en localización o tiempo; cambio en el modelo ajustado a los datos, el cual se refleja en el cambio del comportamiento de la tendencia la cual puede ser lineal y polinomial; y cambio en los atributos de la derivada, es decir la diferencia entre la predicción y el valor actual es considerada un cambio, también se analizan cambios solo temporales en series de tiempo.

En una serie de tiempo: cambio en la media de una serie de tiempo, cambio en el valor de una serie de tiempo y cambio en modelos lineales en una serie de tiempo.

Tipos de cambios espaciales

Basado en imágenes: Los cambios en imágenes incluyen una sola imagen o un conjunto de imágenes (más de 2 imágenes).

Basado en *ráster*: Cambio en el patrón de escala la cual puede ser local, focal o zonal, estos tipos de cambio se muestran en la Figura 2.2.

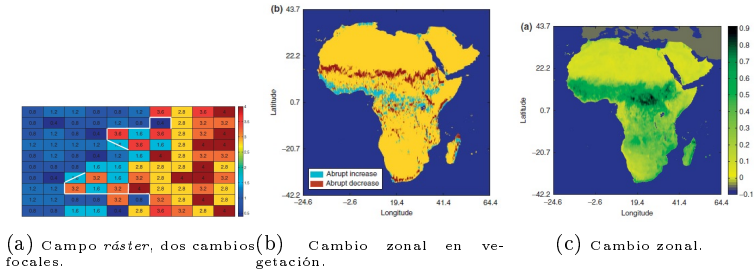


Figura 2.2: Cambios espacio-temporales, la Figura fue obtenida de Xun *et al.* (2014).

Basado en vectores, los cambios pueden ser entre: puntos, líneas, polígonos y redes, estos cambios se muestran en la Figura 2.3.

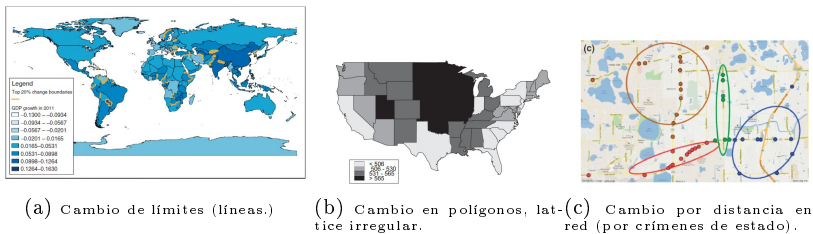


Figura 2.3: Cambios basados en vectores, la Figura fue obtenida de Xun *et al.* (2014).

También se analizan cambios basados en imágenes: los cambios en imágenes incluyen una sola imagen o un conjunto de imágenes. Para cambios en espacio-tiempo el análisis es para las diferentes combinaciones de cambios temporales y espaciales.

Además de los diferentes puntos de cambio, para realizar inferencia sobre los puntos de cambio en la estadística clásica, se requiere una estadística de prueba y contrastar un juego de hipótesis, así existen métodos paramétricos y no paramétricos. De aquí en adelante se incluye el contenido del artículo publicado, Muñiz L., *et al.* (2019).

2.4 Métodos de puntos de cambio en estadística clásica

En esta sección se presentan algunos métodos paramétricos y no paramétricos de puntos de cambio. Enseguida se describe en qué consiste cada uno

de ellos.

Métodos paramétricos

Los métodos paramétricos son: la prueba de homogeneidad normal estándar Alexandersson (1986), Von Newman Von Neumann (1941), Rangos de Buishand Buishand (1982), Movimiento t Yin *et al.* (2015), Análisis de espectro singular (Moskvina 2001, Moskvina y Zhigljavsky 2003), la Prueba de Anomalía Acumulativa Ran *et al.* (2010) y una más basada en kernel Jansenberger y Steinnocher (2014).

Prueba de homogeneidad normal estándar.

La prueba de homogeneidad normal estándar fue desarrollada por Alexandersson (1986) para comparar la media de los primeros k años del registro con el de los últimos $n - k$ años, el contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 \quad vs \quad H_a : \mu_1 \neq \mu_2$$

donde μ_1 es la media de los primeros k años y μ_2 de los últimos $n - k$ años. La estadística $T(k)$ es calculada como:

$$T(k) = k\bar{z}_1^2 + (n - k)\bar{z}_2^2 \quad k = 1, \dots, n, \quad \text{donde} \quad \bar{z}_1 = \frac{1}{k} \sum_{i=1}^k z_i,$$

$$\bar{z}_2 = \frac{1}{n - k} \sum_{i=k+1}^n z_i.$$

Si ocurre un cambio en el año k_0 , entonces $T(k)$ alcanza un máximo cerca del año $k = k_0$. La estadística de prueba T_0 se calcula como:

$$T_0 = \text{máx}(T(k)) \quad \text{para} \quad 1 \leq k \leq n.$$

La hipótesis nula se rechaza cuando T_0 es más grande que un cierto valor crítico, es decir existe un punto de cambio.

Prueba de relación Von Neumann

La prueba de relación de Von Neumann (1941) está relacionada con el coeficiente de correlación serial de primer orden. La relación de Von Neumann se define como la razón de la diferencia del cuadrado medio sucesivo (año con año) N . Su estadística de prueba para detección de punto de cambio en la serie temporal de $X_1, X_2, X_3, \dots, X_n$ se describe como:

$$N = \frac{\sum_{i=1}^n (X_i - X_{i-1})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}, \quad (2.1)$$

las X_i son distribuidas normalmente con media μ y varianza σ^2 , las hipótesis respecto al punto de cambio con respecto a la media son las siguientes:

$H_0 : E(N), \quad i = 1, \dots, m$, es decir la media es constante,

vs

$H_a : E(N) + \Delta, \quad i = m + 1, \dots, n$ existe un cambio de tamaño Δ .

Para series homogéneas, el valor esperado bajo la hipótesis nula es constante $E(N) = 2$. Las series o muestras no homogéneas con un cambio tendrán un valor de N menor que 2, cualquier otro valor implica que la serie de tiempo tiene una variación rápida en su media.

Prueba de rangos Buishand

Buishand (1982) desarrolló esta prueba estadística. Aquí, la suma parcial ajustada, que es la suma acumulada de la desviación de la media para la observación k de una serie $X_1, X_2, X_3, \dots, X_n$ con media (μ) se puede calcular como:

$$S_{0*} = 0 \text{ y } S_{k*} = \sum_{i=1}^k (X_i - \bar{X}), \quad k = 1, \dots, n.$$

Donde las X_i tienen distribución normal. Para series homogéneas, los valores de S_{k*} fluctúan alrededor de cero ya que en la serie de tiempo aleatoria la desviación de su media se distribuye generalmente en ambos lados de la media de la serie. Si la serie se rompe en el año K , entonces S_{k*} alcanza un máximo (cambio negativo) o un mínimo (desplazamiento positivo) cerca del año $k = K$.

Las sumas parciales reajustadas son obtenidas por dividir a S_{k*} por la desviación estándar muestral:

$$S_{k**} = \frac{S_{k*}}{D_x}, \quad k = 0, \dots, n, \quad \text{con } D_x^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / n.$$

La prueba de homogeneidad está basada en las sumas parciales reescaladas ajustadas S_{k**} . La estadística para desarrollos de homogeneidad es:

$$Q = \max_{0 \leq k \leq n} |S_{k**}|,$$

valores altos de Q son una indicación para un cambio Buishand (1982).

Prueba de movimiento t

El método de la prueba del movimiento-t es utilizada por Yin *et al.* (2015) ellos la usaron para detectar los posibles puntos de cambio abruptos. La prueba del movimiento t es la prueba t de dos muestras Snedecor y Cochran (1989) la cual es usada para determinar si dos medias poblacionales son iguales. Los datos pueden ser pareados o no pareados. Si son pareados significa que existe una correspondencia uno a uno entre los valores en las dos muestras, para muestras pareadas usualmente se calcula la diferencia. Para muestras no pareadas, los tamaños de muestra para dos muestras pueden no ser iguales. Las varianzas de las muestras pueden asumirse que son

iguales o diferentes.

El contraste de hipótesis respecto a la media para detectar un cambio es:

$$H_0 : \mu_1 = \mu_2 \quad vs \quad H_a : \mu_1 \neq \mu_2$$

cuando las varianzas son iguales la estadística de prueba es:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad \text{donde} \quad s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}. \quad (2.2)$$

Esta estadística de prueba es usada para detectar un punto de cambio, para una serie de tiempo dada $\{x_n\}$ la cual tiene n muestras aleatorias independientes, el punto de referencia es establecido para un tiempo dado. Las muestras de dos subsecciones antes y después de un dato puntual son n_1 y n_2 . \bar{x}_1 y \bar{x}_2 son el promedio y s_1^2 y s_2^2 son las varianzas de las respectivas subsecciones.

La prueba de movimiento t según Yin *et al.* (2015) se lleva a efecto en tres pasos para detectar el cambio abrupto. Primero, se define la misma longitud de dos subsecciones antes y después del dato puntual; normalmente, $n_1 = n_2$. Segundo, de acuerdo a la expresión (2.2), el valor estadístico de dos subsecciones es sucesivamente calculado usando el método de movimiento t para un conjunto de datos puntuales. Tercero, los valores promedio de dos muestras son comparadas en un nivel de significancia dado para detectar el cambio. Si $|t_i| < t_\alpha$ donde t_i es la estadística de prueba al considerar al i -ésimo punto como punto de referencia para construir la estadística 2.2 y t_α es el cuantil de tamaño α , entonces la variable analizada tiene cambio abrupto en el dato puntual.

Análisis de espectro singular (SSA)

El análisis de espectro singular (Moskvina 2001, Moskvina y Zhigljavsky 2003) es una técnica potente del análisis de series de tiempo que se utiliza para detectar puntos de cambio.

Algoritmo para detectar puntos de cambio

Sea x_1, x_2, \dots, x_T una serie de tiempo con $T < \infty$. Se seleccionan: el ancho de la ventana $N (N \leq T)$, el parámetro lag $M (M \leq \frac{N}{2})$, $k = N - M + 1$ y $0 \leq p^* < q$, $I = (1, \dots, l)$ donde l denota el número de eigenvectores que forman la base del subespacio $L_{n,I}$, $l < M$.

Para cada $n = 0, 1, \dots$ se calcula:

1. La matriz de base $X^{(n)}$, llamada la matriz de trayectoria, $X^{(n)}$ tiene datos multivariados con M características y k observaciones. Las columnas

$X_j^{(n)}$ con $j = 1, \dots, k$ de $X^{(n)}$ son consideradas como vectores que caen en el espacio R^M dimensional.

2. La matriz de covarianza lag. $R_n = X^{(n)}(X^{(n)})^T$.

3. Los M eigenvalores y eigenvectores de R_n .

4. La matriz de prueba $D_{n,I,p^*,q}$ es la matriz de la suma de las diferencias, sus vectores $X_j^{(n)}$ ($j = p^*+1, \dots, q$) forman las columnas de la matriz.

$$D_{n,I,p^*,q} = \sum_{j=p^*+1}^q ((X_j^{(n)})^T X_j^{(n)} - (X_j^{(n)})^T U_{ij} U_{ij}^T X_j^{(n)}),$$

y U_{i1}, \dots, U_{il} denota los l eigenvectores que forman la base del subespacio $L_{n,I}$, $l < M$.

5. La distancia cuadrada normalizada S_n , donde

$S_n = \frac{\tilde{D}_{n,I,p^*,q}}{v_n}$ y $\tilde{D}_{n,I,p^*,q} = \frac{1}{M(q-p^*)} D_{n,I,p^*,q}$ la cual es la suma normalizada de distancias cuadradas (la normalización es hecha con respecto al número de elementos en la matriz de prueba). Un estimador de la suma normalizada de distancias cuadradas $\tilde{D}_{n,I,p^*,q}$ es v_j en los intervalos de tiempo $[j+1, j+N]$ donde la hipótesis de no cambio puede ser aceptada. Se sugiere usar $\tilde{D}_{\tilde{n},I,0,k}$, donde \tilde{n} es el valor más grande de $j < n$ tal que la hipótesis nula de no cambio en el intervalo $[j+1, j+N]$ ha sido aceptada.

La regla de decisión en el algoritmo que se propone es declarar un cambio si para algún n , $S_n \geq H$. Donde H es un umbral fijo, es decir valores grandes de $D_{n,I,p^*,q}$ y S_n indican que existe un cambio en la estructura de la serie.

Nota: Una recomendación general es seleccionar $p^* \geq k$; esto hace que las columnas de la base y la matriz de prueba consistan de elementos diferentes. En este caso, el algoritmo es más sensible a cambios que su versión económica (en el sentido del número de x_t involucrado en cada iteración n) cuando $p^* < k$ y así algunas de las columnas de la base y las matrices de prueba coinciden.

Para obtener un comportamiento más suave de la estadística de prueba $D_{n,I,p^*,q}$, necesita seleccionarse q ligeramente más grande que p^* . Si la diferencia $q - p^*$ es también grande, entonces el comportamiento de $D_{n,I,p^*,q}$ llega a ser también más grande; esto quizá por ejemplo cuando $p^* = 0$ y $q = p^*$ (esto es, la base y la matriz de prueba coinciden).

Prueba de Anomalía Acumulativa

Ran *et al.* (2010), utilizan el método de anomalía acumulativa para revelar cambios abruptos de descargas del agua y concentración de sedimentos suspendidos (SSC). Según Ran *et al.* la anomalía acumulada es un método para distinguir un cambio de tendencia de datos discretos y es ampliamente usado en meteorología para analizar variaciones en precipitación y tempe-

ratura. Para una serie discreta $\{x_i\}$, la anomalía acumulativa (X_t) para datos puntuales x_t puede ser expresada como:

$$\hat{X}_t = \sum_{i=1}^t (x_i - \bar{X}), \quad t = 1, 2, \dots, n,$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i,$$

donde \bar{X} es el valor medio de la serie $\{x_i\}$, y n es el número de puntos discretos. Como lo sugiere la ecuación, la anomalía acumulativa puede ser usada para analizar la magnitud de la fluctuación de una serie de datos discretos, normalmente, el incremento del valor de anomalía acumulativa indica los datos puntuales involucrados que son más grandes que el promedio, de otra manera más bajos que el promedio. En su estudio la variable x representa la descarga promedio anual y el promedio anual de concentración de sedimentos suspendidos (SSC), respectivamente.

Ésta es una prueba que ha sido utilizada por algunos otros investigadores además de Ran *et al.*, para analizar cambios, ellos son: Wang *et al.* (2012) y para cambios en espacio-tiempo Xiujing *et al.* (2013).

Otra prueba de puntos de cambio es con kernels y es la siguiente.

Jansenberger y Steinnocher (2014) realizaron una contribución la cual se enfoca en cambios espacio-temporales. Para la cuantificación espacial de tales cambios fue usado el método de estimación de densidad de kernel dual (KDE). Para tal método, dos conjuntos de datos diferentes fueron relacionados uno con otro. Se analizaron los cambios en concentración espacial de tiendas de comestibles de dos grupos minoristas en la provincia de Austria y de Austria Superior para 1998 y 2001.

Las investigadoras utilizan una función kernel cuártica. Sin embargo, dado que el resultado de un análisis no es fuertemente influenciado por la función seleccionada, no existen reglas concernientes con la selección de una función apropiada Jansenberger y Steinnocher (2014). Existe un número de diferentes funciones kernel, la función kernel más común es la función de distribución normal:

$$k\left(\frac{d_i}{u}\right) = \frac{1}{2\pi} \exp\left[-\frac{1}{2}\left(\frac{d_i}{u}\right)^2\right], \quad (2.3)$$

con $u = \left(\frac{s-s_i}{b}\right)$, basándose en esta función la estimación de la densidad se expresa como:

$$\hat{f}(s) = \frac{1}{b^2} \sum_{i=1}^n K\left(\frac{d_i}{b}\right), \quad (2.4)$$

donde d_i es la distancia entre los puntos s y la localización del punto observado s_i . Debido a que el ancho de banda b es la desviación estándar de la distribución normal, esta función se extiende al infinito en todas las

direcciones, es decir, ésta fue aplicada a cada uno de los puntos en la región.

En este análisis se utiliza la función kernel cuártica que tiene un radio circunscrito b , el cual es también el ancho de banda. La función kernel cuártica es aplicada a un área limitada alrededor de cada localización y tiene la siguiente forma funcional

$$K(u) = \begin{cases} \frac{3}{4}(1 - u^3)^2, & \text{para } u^3 \leq 1 \\ 0, & \text{d.of.} \end{cases}$$

Se estima también la densidad kernel, la cual es una técnica de interpolación que relaciona localizaciones puntuales individuales o puntos s_i para un área entera y provee estimaciones de la densidad, $\lambda(s)$, en una localización dentro de la región de estudio R .

$$\hat{\lambda}(s) = \frac{1}{b^2} \sum_{i=1}^n k\left(\frac{d_i}{b}\right). \quad (2.5)$$

Cuando el KDE es aplicado a una variable el resultado es una estimación de la densidad singular, esta es llamada una estimación de la densidad dual. En el último caso, una densidad kernel es estimada para cada una de las variables individualmente y entonces las dos estimaciones de la densidad son relacionadas una con la otra mediante una simple operación algebraica tal como la suma, la diferencia y el cociente. La operación más usada es el cociente. En este estudio, sin embargo, fue usada la diferencia en valor absoluto de las densidades y con ella se visualizaron los cambios espaciotemporales en la región de estudio.

Métodos no paramétricos

Dentro de las pruebas no paramétricas se tiene: la de Pettitt (Pettitt, 1979), Mann-Kendall Liu Liu *et al.* (2012), Mann-Whitney Jean y Subhabrata (2003) y Lepage Lepage (1971):

Prueba de Pettitt

La prueba no paramétrica Pettitt (Pettitt, 1979), detecta un cambio en un tiempo desconocido, se basa en una versión de la prueba de dos muestras Mann-Whitney, y calcula su significancia estadística. Considera una serie de tiempo $\{x_i\}(1 \leq i \leq n)$ y usa una estadística $U_{t,n}$, que es equivalente a una estadística de Mann-Whitney, la cual se usa para probar que las dos muestras x_1, \dots, x_t y x_{t+1}, \dots, x_n son de la misma población. La hipótesis nula H_0 de la prueba Pettitt es la ausencia de puntos de cambio, mientras la hipótesis alternativa es la existencia de un punto de cambio. Se considera

a $D_{ij} = \text{sgn}(x_i - x_j)$, donde

$$\text{sgn}(x) = \begin{cases} 1, & x > 0; \\ 0, & x = 0; \\ -1, & x < 0. \end{cases}$$

La estadística de prueba es:

$$U_{t,n} = \sum_{i=1}^t \sum_{j=i+1}^n D_{ij}, \quad t = 2, \dots, n.$$

La prueba Pettitt usa la estadística $K_t = \max_{1 \leq t \leq n} |U_{t,n}|$.

Para la prueba de una cola y para el cambio en direcciones, se usan las estadísticas:

$$K_t^+ = \max_{1 \leq t < n} |U_{t,n}|, \quad K_t^- = \min_{1 \leq t < n} |U_{t,n}|.$$

Debe notarse que en la hipótesis nula, $E(D_{ij}) = 0$ y la distribución de $U_{t,n}$ es simétrica alrededor de cero para cada t y la hipótesis alternativa $E(D_{ij}) \neq 0$. Así K_t^+ y K_t^- tiene la misma distribución nula.

Las estadísticas K_t^+ y K_t^- son de una cola, y usan la teoría de Mann-Whitney, se puede esperar que K_t^+ sea grande si ha habido un cambio hacia abajo en el nivel del comienzo de la serie. K_t^+ puede ser grande si $F_1(x) \leq F_2(x)$, con la desigualdad estricta para al menos algún x . Similarmente se puede esperar que K_t^- sea grande si ha habido un cambio hacia arriba o $F_1(x) \geq F_2(x)$.

El punto de cambio significativo se encuentra en el valor $|U_{t,n}|$ máximo y el nivel de significancia asociado con k_t^+ , k_t^- se determina aproximadamente por la probabilidad $p = 2 \exp\{-6k_t^2/(n^3 + n^2)\}$, si p es más pequeño que el nivel de significancia específico, por ejemplo 0.05, la hipótesis nula es rechazada. En otras palabras, si existe un punto de cambio significativo, la serie de tiempo es dividida en 2 partes en la localización del punto de cambio t . La probabilidad de significancia aproximada para un punto de cambio es definida como $p_0 = 1 - p$.

Prueba no paramétrica de Mann-Kendall

1.- La prueba Mann-Kendall, es un método no paramétrico para la detección de tendencia y puntos de cambio debido a su robustez y simplicidad. La prueba Mann-Kendall ha sido ampliamente usada para evaluar la significancia de tendencias monotónicas de variables hidrometeorológicas.

Liu Liu *et al.*(2012) citan esta prueba en su análisis y hacen mención de la prueba Mann-Kendall- Sneyers, por su parte Sneyers (1990) a esta estadística la llama estadística de Mann.

Para las series de tiempo, las magnitudes $x_i (i = 1, 2, \dots, n)$ significan series de tiempo que son comparadas con $x_j (j = 1, 2, \dots, i-1)$. Para cada comparación, se cuenta el número de casos $x_j > x_i$ el cual es denotado por r_i . La hipótesis nula H_0 indica la existencia de no tendencia en la serie de tiempo,

mientras la hipótesis alternativa H_1 establece que existe una tendencia en el conjunto de datos. Bajo la hipótesis nula (no tendencia), la serie de rangos es:

$$S_k = \sum_{i=1}^k r_i,$$

donde

$$r_i = \begin{cases} +1, & \text{si } x_i > x_j, \\ 0, & \text{si } x_i \leq x_j, \end{cases} \quad ; \quad j = 1, 2, 3, \dots, i,$$

tiene distribución aproximadamente normal con media y varianza dadas por: $E(S_k) = \frac{k(k+1)}{4}$, $var(S_k) = \frac{k(k-1)(2k+5)}{72}$.

La estadística secuencial forward

$$UF_k = \frac{[S_k - E(S_k)]}{\sqrt{var(S_k)}}, \quad k = 1, 2, \dots, n$$

es una variable normal estandarizada. La sucesión backward UB_k es estimada usando la misma ecuación pero con una serie invertida de datos. En una prueba de tendencia de 2 colas, la hipótesis nula se acepta con un nivel de significancia α si $|UF_k| \leq (UF_k)_{1-\frac{\alpha}{2}}$ donde $(UF_k)_{1-\frac{\alpha}{2}}$ es el valor crítico de la distribución normal estándar con una probabilidad $0 \leq \alpha \leq 1$. $UF_\alpha > 0$ denota una tendencia hacia arriba, mientras lo contrario denota una tendencia a la baja (i.e UB_k es similar a UF_k). La versión secuencial de la prueba usada permite la detección del tiempo aproximado de ocurrencia del cambio de tendencia localizando la intersección de las curvas forward y backward (Lapiedra (2004)) de la estadística de prueba. Un punto de intersección dentro del intervalo de confianza indica un punto de cambio.

2.- Otra forma de la estadística de Mann-Kendall utilizada por Xiaomeng *et al.* (2014) es considerada de la siguiente forma: para una serie de tiempo dada x_1, x_2, \dots, x_n , la estadística de Mann-Kendall S es definida como

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i).$$

Donde n es la longitud del dato registrado, x_i y x_j son los valores de los datos secuenciales.

La estadística S está distribuida aproximadamente normal con media y varianza:

$$\frac{n(n-1)(2n+5) - \sum_{i=1}^n t_i(i-1)(2i+5)}{18},$$

donde t_i es considerada como el número de uniones en la muestra i . La estadística de prueba es una normal estándar Z utilizada de la siguiente manera

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}}, & S > 0, \\ 0, & S = 0, \\ \frac{S+1}{\sqrt{VarS}}, & S < 0. \end{cases}$$

La hipótesis nula de que no existe tendencia en los datos es rechazada (no rechazada) si la estadística Z es más grande (menor) que el valor crítico $Z_{\alpha/2}$ obtenido en el nivel de significancia α . Un valor positivo (negativo) de Z significa un valor de tendencia hacia adelante (hacia atrás).

Además, la prueba Mann-Kendall no paramétrica puede también ser usada para detectar los puntos de cambio de series de tiempo. Similarmente, los valores de UB se calculan hacia atrás desde el final de la serie de tiempo. Si las curvas UF y UB se intersectan, entonces divergen y adquieren valores de umbral específicos, entonces existe una tendencia estadísticamente significativa. El punto de intersección muestra el punto de cambio aproximado en el cual empieza la tendencia.

Prueba Mann-Whitney

La prueba Mann-Whitney, es una prueba no paramétrica de comparación de dos muestras independientes. La estadística de prueba U de Mann-Whitney es definida como el número de veces que y precede a x en el arreglo ordenado combinando observaciones de las 2 muestras aleatorias independientes x_1, x_2, \dots, x_m y y_1, y_2, \dots, y_n , cuando hay empates se saca una especie de promedio y se reparten esos valores

en una sucesión singular de $m+n = N$ variables crecientes en magnitud. Se asume que las 2 muestras provienen de distribuciones continuas, así que la posibilidad $x_i = y_j$ para algunos i y j no son necesariamente considerados.

Si las variables aleatorias indicadoras m_n son definidas como

$$D_{ij} = \begin{cases} 1, & \text{si } y_j < x_i \\ 0, & \text{si } y_j > x_i \end{cases} \quad \text{para } i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n$$

entonces una representación simbólica de la estadística U Mann Whitney es:

$$U = \sum_{i=1}^m \sum_{j=1}^n D_{ij}.$$

En el contraste de hipótesis, la hipótesis nula considera que las muestras provienen de una misma población, esto es:

$$H_0 : F_y(x) = F_x(x) \quad \text{vs} \quad H_a : F_y(x) \neq F_x(x).$$

La región de rechazo es:

$$U - mn = 2 > k.$$

Prueba de punto de cambio de Lepage

La prueba Lepage es una prueba no paramétrica de dos muestras para ubicación y dispersión Lepage (1971), que ha sido ampliamente utilizada para detectar cambios tales como tendencias a largo plazo, variaciones cíclicas y cambios escalonados para la precipitación. Lepage, supone que el tamaño de la serie de estudio es igual o mayor a diez y la estadística Lepage (HK) sigue la distribución Chi-cuadrada (χ^2) con dos grados de libertad. Se supone que las muestras provienen de distribuciones continuas y que son independientes. La estadística de Lepage (HK) es la siguiente

$$HK = \frac{[W - E(W)]^2}{V(W)} + \frac{[A - E(A)]^2}{V(A)}, \quad (2.6)$$

donde A y W se escriben más adelante en este mismo apartado. Sea x_1, x_2, \dots, x_{n_1} y y_1, y_2, \dots, y_{n_2} dos muestras independientes de tamaño n_1 y n_2 . Se asume que $u_i = 1$ si la i -ésima observación más pequeña en una muestra combinada del tamaño (n_1, n_2) pertenece a x y $u_i = 0$ si pertenece a y .

La hipótesis nula H_0 de la prueba Lepage supone que las distribuciones de donde provienen las dos muestras son iguales, se contrasta contra la alternativa H_a en la cual se considera que son diferentes. Si la estadística de prueba HK excede 5.99 la diferencia entre dos muestras es juzgada como significativa en el nivel de confianza del 95 por ciento (nivel de significancia del 5 por ciento), es decir se rechaza la hipótesis nula de que las distribuciones son iguales, por lo tanto existe un punto de cambio.

Los términos en ecuación (2.6) se pueden derivar en base a las siguientes ecuaciones:

$$W = \sum_{i=1}^{n_1+n_2} i u_i, \quad E[W] = \frac{n_1(n_1+n_2+1)}{2}, \quad V[W] = \frac{n_2 n_1 (n_1+n_2+1)}{2}.$$

Además

$$A = \frac{1}{2} n_1 (n_1 + n_2 + 1) + \sum_{i=1}^{n_1+n_2} |i - \frac{1}{2} (n_1 + n_2 + 1)| u_i.$$

Si $n_1 + n_2$ es par, entonces $E[A]$ y $V[A]$ serán estimados como: $E[A] = \frac{n_1(n_1+n_2+2)}{4}$ y $V[A] = \frac{n_1 n_2 (n_1+n_2-2)(n_1+n_2+2)}{48(n_1+n_2-1)}$.

Si $n_1 + n_2$ es impar, $E[A]$ y $V[A]$ serán estimados como: $E[A] = \frac{n_1(n_1+n_2+1)^2}{4(n_1+n_2)}$ y $V(A) = \frac{n_1 n_2 (n_1+n_2+1)[(n_1+n_2)^2+3]}{48(n_1+n_2)^2}$.

Las características estadísticas de los segmentos divididos por los puntos de cambio se detectan por la media y el coeficiente de variación (C_v). Así, $\mu_x = E[x] = \mu'_1$ y $C_v = \frac{S_x}{\bar{x}}$, donde

$$S_x = \sqrt{\frac{\sum_{i=1}^k (x_i - \bar{x})^2}{(n-1)}}.$$

En la siguiente sección se presentan tablas con aplicaciones de la metodología de puntos de cambio.

2.5 Aplicaciones

Enseguida se presentan tres tablas 2.1, 2.2 y 2.3, las cuales resumen los diferentes problemas de aplicación de puntos de cambio estudiados y la metodología utilizada para su detección, así como los investigadores que han utilizado dichas metodologías. Las tablas constan de tres columnas las cuales contienen el problema o aplicación, el autor o autores y el modelo o metodología utilizada para detectar el punto o puntos de cambio.

Se puede observar en las tablas que los estudios de variables hidrometeorológicas son de principal interés. Como pudo observarse en las metodologías introducidas tanto paramétricas como no paramétricas, lo que se hizo fue encontrar la estadística de prueba y es por medio de ésta que al contrastar las hipótesis se decide si existe o no puntos de cambio. Por lo que si se quiere crear otra metodología el objetivo es encontrar la estadística de prueba. En cuanto a aplicaciones, por ejemplo en los modelos Bayesianos, para encontrar la distribución a posteriori se pueden trabajar tanto distribuciones a priori informativas como no informativas.

Tabla 2.1: Regresión, Kernel, Análisis de espectro singular, Pettitt, Movimiento t, Rangos de Buishand, Homogeneidad normal estándar y Von Newman.

Problema	Autores	Modelos
Indice de sequedad Temperatura en lago Precipitación, temperatura Incendios Concentración de contaminante Precipitación, temperatura Cambio en aerosoles en océano Crecimiento de abetos chinos Cambio de uso de la tierra Temperatura en superficie del mar Cambio de grano y extensión del paisaje Indice de vegetación	Abdol <i>et al.</i> (2016) Yankova <i>et al.</i> (2017) Skiriris <i>et al.</i> (2014) Anabela <i>et al.</i> (2010) Abdel H. <i>et al.</i> (2011) Chengjing Nie <i>et al.</i> (2012) Cermak <i>et al.</i> (2010) Ma. <i>et al.</i> (2006) Bollinger <i>et al.</i> (2007) González y Andón (2012) Monica G. <i>et al.</i> (1989) Luan <i>et al.</i> (2018)	Regresión
Concentración espacial de tiendas de comestible	Jansenberger, E. y Steinnocher, S. (2014)	Kernel
Hidrología (flujo)	Yang <i>et al.</i> (2009)	Análisis de espectro singular
Clima (Precipitación)	Xiujing <i>et al.</i> (2013)	Pettitt Rangos de Buishand
Clima (Temperatura)	Yin <i>et al.</i> (2015)	Movimiento t
Clima (Temperatura)	Malekian, A. y Kazemzadeh, M. (2015)	pettitt Movimiento t pettitt Rangos de Buishand
Cambio en temperatura del aire	Chakraborty <i>et al.</i> (2017)	Von Neumann Pettitt Rangos de Buishand Homogeneidad normal estándar
Precipitación y temperatura	Depesh y Madan (2016)	Pettitt
Concentración de un contaminante	Abdel H. <i>et al.</i> (2011)	Weibull lognormal
Contaminación por zink	Tlili <i>et al.</i> (2011)	Estadística Multivariada
Crecimiento de abetos chinos	Ma. <i>et al.</i> (2006)	Coefficiente de Gini Coefficiente de Lorentz
Comunidad de bacterias	Mahaffee y Kloeppe (1996)	Indice de Shannon Indice de Hill Indice de Simpson

Tabla 2.2: MANOVA, ARIMA, Bayesiano especial, Mann-Kendall, Mediana y ANOVA.

Problema	Autores	Modelos
Vegetación exotica	Tierney y Cushman (2005)	MANOVA ANOVA
Cambio de uso de la tierra	Bollinger <i>et al.</i> (2007)	Regresión Logística
Respiración del suelo	Akburak y Makineci, (2012)	ANOVA
Campo solar magnetico	Mordvinov y Plyusnina (2000)	MANOVA
Precipitación	Won-Ho <i>et al.</i> (2015)	ANOVA
Detección de fraude en telecomunicaciones	Hilas <i>et al.</i> (2013)	ARIMA
Clima (Temperatura)	Luo <i>et al.</i> (2005)	Mediana
Contenido de agua del suelo	Cubera y Moreno(2007)	ANCOVA
Contenido de agua del suelo	Jian <i>et al.</i> (2009)	I de Moran
Tasas de precipitación	Chavallaz <i>et al.</i> (2015)	Mann-Kendall Movimiento t
Diversidad de especies de vegetación	Horward (1975)	Ecuación McIntosh
Clima(Precipitación, Temperatura)	Zhao <i>et al.</i> (2016)	Método de Yamamoto Método de Wavelet Tasa de tendencia
Cambios en dos tipos de plantas del desierto	Zuo <i>et al.</i> (2005)	Binomial negativa
Precipitación, Temperatura Cambio de uso de la tierra	Chengjing <i>et al.</i> (2012) Yu (2013)	Spline Autómata celular Cadenas de Markov
Cambio en serie de ciclones tropicales	Xin y Pao (2005)	Bayesiano temporal
Coefficiente intelectual Clima(precipitación) Tasa de incidencia de cáncer	Cai <i>et al.</i> (2016) Chen <i>et al.</i> (2016) Sarate C. <i>et al.</i> (2011)	Bayesiano espacial
Violencia en ventas de alcohol	Yanjun Xu <i>et al.</i> (2012)	Bayesiano espacio tiempo

Tabla 2.3: Mann-Kendall, Pettitt, Lepage, Bayesiano, spline, comparativo de datos e imágenes.

Problema	Autores	Modelos
Hidrología (flujo)	Yang <i>et al.</i> (2009) ¹	1. Lepage
Clima (precipitación)	Liu Liu <i>et al.</i> (2012) ^{2,3}	2. Mann-Kendall
Clima (Precipitación, flujo)	Bing <i>et al.</i> (2013) ^{2,3}	3. Pettitt
Clima (Precipitación)	Xiujing <i>et al.</i> (2013) ³	4. Wilcoxon
Clima (Precipitación)	Xiaomeng <i>et al.</i> (2014) ²	5. Variables ordinales
Clima (Temperatura)	Yanyu <i>et al.</i> (2015) ²	y nominales
Clima (Temperatura)	Malekian y Kazemzadeh ^{2,3} (2015)	
Clima (Precipitación, flujo)	Tesfay <i>et al.</i> (2017) ^{2,3}	
Clima (Precipitación)	Biana <i>et al.</i> (2017) ^{2,3}	
Temperatura del agua	Peter A. (2017) ³	
Precipitación y temperatura	Deepesh y Madan (2016) ³	
Temperatura en el lago	Yankova <i>et al.</i> (2017) ²	
Incendios	Anabela <i>et al.</i> (2010) ⁴	
Incendios forestales	Inga <i>et al.</i> (2007) ⁵	
Precipitación	Adeyeri <i>et al.</i> (2017) ³	
Precipitación	Tomozeiu <i>et al.</i> (2013) ^{2,3}	
Patrones de vegetación	James y Carl (1999) ⁴	
Indices de vegetación	Marcoe y Pilson (2017) ⁷	7. Comparativo de datos
Cambio de hábitat en estuario.		
Cubierta de bosques	Madhushree <i>et al.</i> (2012) ⁸	8. Comparación de mapas
Cambio en Isla calurosa	Shurtilipi <i>et al.</i> (2007) ⁹	9. Fotografías aéreas
Cambio de la playa	Michalowska <i>et al.</i> (2016) ¹⁰	10. Imágenes
Cambio geográfico	Yang <i>et al.</i> (2009) ⁷	

Capítulo 3

Detección de puntos de cambio

En este Capítulo se incluye la programación de algunos procedimientos para determinar los puntos de cambio, para ello se define el factor de Bayes el cual fue utilizado para seleccionar puntos de cambio temporales y espacio temporales. Se define también un proceso Poisson homogéneo, ya que se simulan datos con este proceso. Se detectaron con el factor de Bayes múltiples puntos de cambio temporales y puntos de cambio espacio tiempo. Se detectaron también cambios con el método de umbrales y con el método de suma acumulativa (CUSUM). Finalmente se aplicaron algunas pruebas como pruebas de rangos de Buishand, Pettitt y prueba de homogeneidad normal estándar, las cuales confirmaron el punto de cambio con el método de suma acumulativa. Para el método con el factor de Bayes y el método de umbrales se aplicaron métodos propuestos en la tesis de Altieri (2015), para el método de suma acumulativa se utilizó el artículo de Taylor (2000). Las pruebas de rangos de Buishand, Pettitt y prueba de homogeneidad normal estándar ya fueron revisadas con anterioridad.

Para todo el trabajo se elaboraron algoritmos y programas, estos últimos se hicieron con el objetivo de detectar los puntos de cambio, se analizan y comparan sus resultados, fueron creados con algunas instrucciones del paquete INLA de R presentadas en los libros Gómez (2020) y Blangiardo y Cameletti (2015) para ello se usaron las librerías INLA y Poisson, la única excepción fue para las pruebas de homogeneidad normal estándar, Pettitt y rangos de Buishand para las cuales se utilizó el paquete *trend de R* propuesto en Thorsten (2020).

Las bases de datos trabajadas en este Capítulo son la *cement* que se tomó del libro de Gómez(2020), es una base de datos de un ejemplo de regresión y se utilizó para trabajar el método de umbrales, la de datos de accidentes de tráfico tomada de la página de INEGI y la de coronavirus tomada de la página *coronavirus.gob.mx*, fueron utilizadas para detectar los puntos de cambio con la estadística CUSUM . También se utilizó una base de datos

simulada de un proceso Poisson, con la cual se detectaron los cambios con el factor de Bayes en tiempo y espacio-tiempo.

3.1 Factor de Bayes

Si se presenta un problema de selección de modelos en el que se debe elegir entre dos posibles modelos, en base a un conjunto de datos observados D , la plausibilidad de los dos diferentes modelos M_1 y M_2 , parametrizados por vectores de parámetros θ_1 y θ_2 se puede medir mediante el factor Bayes.

El factor de Bayes se define como:

$$B = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\int_{\theta_1} P(D|\theta_1, M_1)\Pi(\theta_1|M_1)d\theta_1}{\int_{\theta_2} P(D|\theta_2, M_2)\Pi(\theta_2|M_2)d\theta_2},$$

donde $P(D|M_1)$ se denomina verosimilitud marginal o verosimilitud integrada. Esto es similar a lo que se hace en las pruebas de la razón de verosimilitudes pero ahora, en lugar de maximizar la verosimilitud, el factor Bayes realiza un promedio ponderado mediante la distribución de los parámetros.

Un valor de $B > 1$ significa que M_1 es apoyado por los datos más que M_2 .

En el caso del factor de Bayes, Jeffreys estableció una escala de interpretación de B , la cual se muestra en la Tabla 3.1.

Tabla 3.1: Escala de interpretación de B , según Jeffreys.

B	Fuerza de la evidencia a favor de M_1
$B \leq 1$	Negativa apoya M_2
$1 < B \leq 3$	Muy escasa
$3 < B \leq 10$	Sustancial
$10 < B \leq 30$	Fuerte
$30 < B \leq 100$	Muy fuerte
> 100	Decisiva

Otra forma de considerar el factor de Bayes es la siguiente:

Supongamos dos hipótesis H_0 y H_1 tales que las probabilidades de las distribuciones a priori son $f_0 = P(H_0)$ y $f_1 = P(H_1)$. Después de observar una muestra, las probabilidades a posteriori de ambas hipótesis son $\alpha_0 = P(H_0|x)$ y $\alpha_1 = P(H_1|x)$. Se define el factor de Bayes a favor de H_0 como

$$B = \frac{\frac{\alpha_0}{\alpha_1}}{\frac{f_0}{f_1}} = \frac{\alpha_0 f_1}{\alpha_1 f_0}.$$

Así, el factor de Bayes representa la plausibilidad a posteriori dividida entre la plausibilidad a priori. Nos informa de los cambios en nuestras creencias introducidas por los datos. Tiene la propiedad de que es casi objetivo y elimina parcialmente la influencia de la distribución a priori.

Como ejemplo supóngase el contraste simple:

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1.$$

Tenemos que las distribuciones a posteriori son:

$$\alpha_0 = P(H_0|x) = \frac{f_0 L(\theta_0|x)}{f_0 L(\theta_0|x) + f_1 L(\theta_1|x)} \quad y \quad \alpha_1 = P(H_1|x) = \frac{f_1 L(\theta_1|x)}{f_0 L(\theta_0|x) + f_1 L(\theta_1|x)}$$

Entonces el factor de Bayes es:

$$B = \frac{\alpha_0 f_1}{\alpha_1 f_0} = \frac{f_0 L(\theta_0|x) f_1}{f_1 L(\theta_1|x) f_0} = \frac{L(\theta_0|x)}{L(\theta_1|x)},$$

que coincide con la razón de verosimilitudes, de modo que la distribución a priori no influiría en este caso en el factor de Bayes.

Así el factor de Bayes para el punto de cambio cuando se divide en dos segmentos esta dado por la razón de verosimilitudes:

$$\frac{L_1}{L_0} = \frac{Q_1 Q_2}{L_0},$$

donde Q_1 es la verosimilitud del segmento 1 y Q_2 es la verosimilitud del segmento 2 bajo la hipótesis alternativa y L_0 es la verosimilitud bajo la hipótesis nula.

Aplicando el logaritmo se tiene

$$B = \log\{Q_1\} + \log\{Q_2\} - \log\{L_0\}.$$

3.2 Proceso de Poisson homogéneo

Definición 3.1: Una colección de variables aleatorias $\{N(t) : t \geq 0\}$ definidas en un espacio de probabilidad (ω, F, P) se llama proceso de Poisson (homogéneo) con intensidad $\lambda > 0$ si satisface las siguientes propiedades:

i) $P(N(0) = 0) = 1$.

ii) Para todo $0 < s < t$, $N(t) - N(s)$ tiene distribución de Poisson de parámetro $\lambda(t - s)$.

iii) Para todo $0 \leq t_1 < \dots < t_n$, $n \geq 1$ (es decir para todo conjunto finito de tiempos), las variables aleatorias $N(t_n) - N(t_{n-1}), \dots, N(t_2) - N(t_1), N(t_1) - N(0), N(0)$, son independientes. Esta propiedad se conoce como propiedad de incrementos independientes.

3.3 Método de bisección

Para determinar los puntos de cambio con el factor de Bayes se utilizó el método de bisección. Éste consiste en dividir los datos a la mitad y se busca un punto de cambio con el factor de Bayes y las verosimilitudes correspondientes a la hipótesis nula y la alternativa, enseguida se va a la izquierda y se divide también a la mitad, posteriormente se divide el lado derecho y así se sigue sucesivamente yendo a la izquierda y a la derecha dividiendo a la mitad y buscando puntos de cambio con el factor de Bayes.

3.4 Múltiples puntos de cambio temporales

Con el objetivo de detectar, analizar y comparar resultados de puntos de cambio se programó para detectar 5 puntos de cambio, los datos utilizados son simulados de un proceso Poisson del cual se obtuvo un conjunto de 60 datos con 6 valores diferentes para el parámetro λ . Se pretende en este problema detectar los cambios que se generan con los distintos valores de λ , se hizo además un comparativo de los resultados al utilizar a prioris uniforme, log-gamma y una Gaussiana.

Se corrió un programa en R, para detectar múltiples puntos de cambio temporales, en total para 5, se utilizaron 60 datos y el segmento más pequeño fue de 4 puntos. Se simuló un proceso Poisson uniforme con diferentes parámetros λ y se aproximó a la distribución a posteriori con el paquete de R INLA, el cual aproxima con series de Taylor. Para detectar los puntos de cambio se utilizó el factor de Bayes y el método de bisección.

Se utilizó una distribución a priori uniforme, log-gamma y una Gaussiana. La Figura 3.1, muestra 5 puntos de cambio.

Para el siguiente problema se usan 60 datos y se llega a la división más pequeña que fue de cuatro datos, es decir tenemos quince divisiones (Tabla 3.2).

Los puntos de cambio para una distribución a priori uniforme con $\lambda = 2, 1, 4, 7, 6, 1$ y en las divisiones de datos o segmentos 8, 7, 15, 15, 8, 7, esto debido a que en los segmentos dividiendo de dos en dos para 30 datos, quedo en un segmento 15 y en el otro 15, al dividir 15 quedo en un segmento 8 datos y en el otro 7. El resultado para la detección de los puntos de cambio fueron como se muestra en la Tabla 3.2. Los puntos de cambio son los primeros cuatro y el séptimo.

Con los mismos puntos de cambio y con la distribución a priori log-gamma con parámetros 0.01 y 0.01, el resultado se muestra en la Tabla 3.2. Se puede observar que en comparación con la distribución uniforme las log-verosimilitudes de la log-gamma son más pequeñas, sin embargo también detectan los puntos de cambio.

Con la distribución a priori Gaussiana con media cero y parámetro de precisión 0.001 para el mismo número de puntos de cambio como la distribución uniforme y la log-gamma se detectaron los cinco puntos de cambio, se puede observar en los datos a continuación en la Tabla 3.2, son los primeros cuatro y el siete, aunque el valor de la log-verosimilitud del séptimo punto de cambio es menor en comparación con las siguientes posiciones como se puede observar.

Los 5 puntos de cambio que se observan en la tabla 3.2, fueron detectados con el factor de Bayes B mencionado anteriormente, los valores para B indican la fuerza de la evidencia a favor de la hipótesis del punto de cambio, valores menores a 1 indican que no existe punto de cambio. Los datos en negrita que son los puntos de cambio se detectaron de la siguiente forma: el valor de 1 fue detectado a la mitad de los datos, el valor 3 en el dato número 15, el valor 3 en el dato 45, el número 4 en el dato número 8 y el 7 en el dato 53, como se puede observar los valores en negrita de los puntos de cambio no son iguales entre si ni con el resto, en cambio en donde no sucedio un punto de cambio los valores son iguales, por ejemplo los valores 5 y 6 son iguales; el 8,10 y 14 también son iguales; el 9, 11, 13 y 15 son iguales, por lo que todos estos valores estan indicando que no existen puntos de cambio. La fuerza de la evidencia a favor de la hipótesis alternativa indica según los puntos de cambio que al aplicar el método de bisección e ir detectando los puntos de cambio, la evidencia es fuerte en el primer punto de cambio y va disminuyendo en los siguientes, obsérvese ésto en la tabla 3.2

Los puntos de cambio se muestran en la Figura 3.1 y son en 8, 15, 30, 45 y 53.

Tabla 3.2: Puntos de cambio.

Num	uniforme	<i>log gamma</i>	Gaussiana
1	49.565011	43.644575	42.972975
2	24.392282	25.829232	26.592785
3	13.940830	26.041694	28.329433
4	16.000633	9.990938	11.641769
5	12.478724	2.387813	4.477951
6	12.478724	2.387813	4.477951
7	12.572475	4.498905	5.523642
8	5.701123	1.835601	5.801758
9	4.732525	1.384310	4.912561
10	5.701123	1.835601	5.801758
11	4.732525	1.384310	4.912561
12	5.701123	1.835601	5.801758
13	4.732525	1.384310	4.912561
14	5.701123	1.835601	5.801758
15	4.732525	1.384310	4.912561

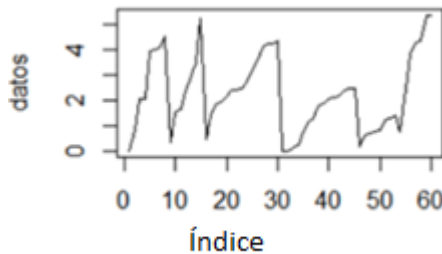


Figura 3.1: Gráfica de 5 puntos de cambio.

3.5 Para 6 puntos de cambio

Con el mismo modelo uniforme se detectaron 6 puntos de cambio para $\lambda = 2, 1, 4, 7, 6, 1, 2$ y en las divisiones de datos o segmentos de 8, 7, 15, 15, 8, 4, 3.

Los puntos de cambio que se detectaron son los primeros cuatro, el séptimo y el último. Se muestran estos en las Tablas 3.3 y 3.4.

Con la distribución a priori el logaritmo de gamma, para los 6 puntos de cambio que son los 4 primeros, el siete y el último, también se detectaron los 6 puntos de cambio, la diferencia con la uniforme es que la log verosimilitud es más pequeña, como se puede observar en la Tabla 3.3.

Para la distribución a priori Gaussiana también se detectaron los 6 puntos de cambio, se muestran enseguida en valores de log verosimilitud, es un poco

parecida a la uniforme. Estos son los puntos de cambio, los cuatro primeros, el siete y el último, se muestran enseguida en la Tabla 3.3. La diferencia entre la Gaussiana y la uniforme es que el último punto de cambio es menor a los valores de log verosimilitud anteriores, en donde no hay punto de cambio.

Con la distribución logit-beta con parámetros 2 y 2 se detectaron 6 puntos de cambio, y son los mismos que para la distribución uniforme y la distribución log-gamma, los puntos de cambio se detectan bien, son los cuatro primeros, el séptimo y último. No se visualizan grandes diferencias con respecto a la uniforme.

Para la distribución a priori normal truncada con parámetros media 0 y precisión 0.001, la verosimilitud es más pequeña comparada con la uniforme, esta es muy parecida a la log-gamma, sin embargo también se detectaron 6 puntos de cambio, los cuatro primeros, el siete y el último, el resultado está en las Tablas 3.3 y 3.4.

De forma similar como en el problema anterior se presentan los puntos de cambio que se observan en las tablas 3.3 y 3.4, éstos valores no se repiten entre ellos ni entre el resto. En cambio el resto de los valores si se repiten se puede observar por ejemplo que los valores 5 y 6 son iguales; el 8, 10, 12 y 14 también son iguales ; el 9, 11 y 13 son iguales, por lo tanto todos estos valores aún cuando sean mayores que 1 no indican punto de cambio, pues se estan repitiendo.

Tabla 3.3: 6 puntos de cambio.

Num	uniforme	log-gamma	Gaussiana
1	37.382628	44.717137	38.258157
2	23.107912	28.913661	26.506732
3	22.987741	27.851420	26.529322
4	16.466681	5.045601	11.798824
5	12.478724	2.387813	4.477951
6	12.478724	2.387813	4.477951
7	16.555598	3.062915	12.572375
8	5.701123	1.835601	5.801758
9	4.732525	1.384310	4.912561
10	5.701123	1.835601	5.801758
11	4.732525	1.384310	4.912561
12	5.701123	1.835601	5.801758
13	4.732525	1.384310	4.912561
14	5.701123	1.835601	5.801758
15	4.733385	1.720949	4.280097

Tabla 3.4: 6 puntos de cambio.

Num	logit-beta	normal truncada
1	49.664317	51.841263
2	30.567246	16.016904
3	18.174484	33.658036
4	9.944739	7.969972
5	4.633706	2.186421
6	4.633706	2.186421
7	11.248933	1.984581
8	4.076750	1.881876
9	3.794214	1.783849
10	4.076750	1.881876
11	3.794214	1.783849
12	4.076750	1.881876
13	3.794214	1.783849
14	4.076750	1.881876
15	4.405551	2.260447

Los puntos de cambio se muestran en la Figura 3.2, estan dados en 8, 15, 30, 45, 53 y 57.

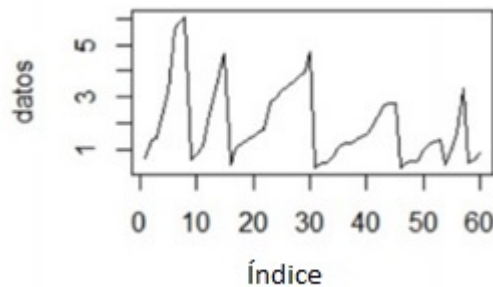


Figura 3.2: Gráfica de 6 puntos de cambio.

Se programó en R, el programa se encuentra en el apéndice y es etiquetado como programa 1.

Se creó un algoritmo para el problema de detectar 6 puntos de cambio con el factor de Bayes el cual se presenta a continuación:

Algoritmo 1: Detecta 6 puntos de cambio

Entrada La función INLA regresa variables **mp1**, **mp2**, y **mp** las cuales contienen los datos simulados con un modelo autoregresivo AR1 y de una familia Poisson para **datos A**, **datos B** y **datos**.

Salida El vector **respuesta** guarda las verosimilitudes de 15 divisiones.

Inicializar se inicializan los vectores **mp1**, **mp2**, y **mp** los cuales contienen los datos simulados con un modelo autoregresivo AR1 y de una familia Poisson para **datos A**, **datos B** y **datos**.

Regresa respuesta vector que guarda las verosimilitudes de las divisiones.

El vector **datos** se divide de 1 a 30 y de 31 a 60 y se asigna a **datos A** y **datos B**.

- i) Se calcula la verosimilitud de **datos A** y **datos B** y se guarda en vector **respuesta**,
- ii) Se dividen a la mitad los vectores **datos A** y **datos B** y se asignan los datos a **datos C** y **datos D**,
- iii) Se calcula la verosimilitud para **datos C** y **datos D**.
- iv) Se repiten los pasos i) y ii) con las nuevas divisiones de los datos hasta llegar a la división de solo dos datos.
- v) se registran los resultados obtenidos.

3.6 Múltiples puntos de cambio en espacio tiempo

Para múltiples puntos de cambio en espacio tiempo, lo que se hizo fue asignar posiciones, la primera en el centro y el resto alrededor de ésta en forma ascendente y en contra de las manecillas del reloj. Se aproximó la distribución a posteriori con el paquete R-INLA, se utilizó la distribución a priori uniforme, el factor de Bayes y el método de bisección, así como la simulación de valores se hizo para un proceso Poisson homogéneo, se utilizaron 4 valores de λ aleatorios del 1 al 15 para cada posición, es decir se estimaron 3 puntos de cambio, estos valores de λ son dados en cada fila de la Tabla 3.5.

Y el resultado de los puntos de cambio son dados en las Tablas 3.6, 3.7, 3.8 y 3.9, son los primeros tres, es decir, se encuentran en 15, 30 y 45.

El comportamiento de los valores es de forma similar como en el primer problema presentado en este Capítulo, los puntos de cambio que se indican en negrita en las tablas 3.6, 3.7, 3.8 y 3.9 no se repiten, a diferencia del resto en los que si existe repetición. De acuerdo a la fuerza de la evidencia del factor de Bayes B los valores menores que 1 indican que no existe punto de cambio, éstos valores se observan en la tabla 3.7, posición 6, valor 3 y en la tabla 3.8, posición 9, valor 2.

3.6. MÚLTIPLES PUNTOS DE CAMBIO EN ESPACIO TIEMPO59

Tabla 3.5: Valores de λ .

	[1]	[2]	[3]	[4]
[1]	10	3	5	13
[2]	8	12	6	5
[3]	14	8	13	5
[4]	5	4	7	12
[5]	12	7	6	8
[6]	12	9	1	6
[7]	1	13	5	3
[8]	11	14	1	3
[9]	1	13	5	15
[10]	13	1	7	4
[11]	5	11	8	9
[12]	10	3	7	15
[13]	1	13	12	6
[14]	7	1	11	12
[15]	13	12	4	1
[16]	5	6	12	7

Tabla 3.6: Puntos de cambio espacio tiempo.

	[1]	[2]	[3]	[4]
1	33.693643	44.955163	35.315124	41.732270
2	27.795301	25.226235	28.387088	19.482977
3	28.727677	21.122882	27.775696	28.697886
4	12.478724	12.478724	12.478724	12.478724
5	12.478724	12.478724	12.478724	12.478724
6	12.478724	12.478724	12.478724	12.478724
7	12.478724	12.478724	12.478724	12.478724
8	5.701123	5.701123	5.701123	5.701123
9	4.732525	4.732525	4.732525	4.732525
10	5.701123	5.701123	5.701123	5.701123
11	4.732525	4.732525	4.732525	4.732525
12	5.701123	5.701123	5.701123	5.701123
13	4.732525	4.732525	4.732525	4.732525
14	5.701123	5.701123	5.701123	5.701123
15	4.732525	4.732525	4.732525	4.732525

Tabla 3.7: Puntos de cambio espacio tiempo.

	[5]	[6]	[7]	[8]
1	33.574630	61.47789451	52.546434	44.358894
2	27.549217	27.37020690	18.324691	21.142981
3	28.777261	-0.08807506	20.678559	1.851282
4	12.478724	12.47872437	12.478724	12.478724
5	12.478724	12.47872437	12.478724	12.478724
6	12.478724	12.47872437	12.478724	12.478724
7	12.478724	12.47872437	12.478724	12.478724
8	5.701123	5.70112304	5.701123	5.701123
9	4.732525	4.73252533	4.732525	4.732525
10	5.701123	5.70112304	5.701123	5.701123
11	4.732525	4.73252533	4.732525	4.732525
12	5.701123	5.70112304	5.701123	5.701123
13	4.732525	4.73252533	4.732525	4.732525
14	5.701123	5.70112304	5.701123	5.701123
15	4.732525	4.73252533	4.732525	4.732525

Tabla 3.8: Puntos de cambio espacio tiempo.

	[9]	[10]	[11]	[12]
1	71.06948660	35.332035	34.663739	28.888859
2	-0.08807506	28.183967	26.321000	25.185406
3	19.83004426	27.173106	28.558468	28.850767
4	12.47872437	12.478724	12.478724	12.478724
5	12.47872437	12.478724	12.478724	12.478724
6	12.47872437	12.478724	12.478724	12.478724
7	12.47872437	12.478724	12.478724	12.478724
8	5.70112304	5.701123	5.701123	5.701123
9	4.73252533	4.732525	4.732525	4.732525
10	5.70112304	5.701123	5.701123	5.701123
11	4.73252533	4.732525	4.732525	4.732525
12	5.70112304	5.701123	5.701123	5.701123
13	4.73252533	4.732525	4.732525	4.732525
14	5.70112304	5.701123	5.701123	5.701123
15	4.73252533	4.732525	4.732525	4.732525

3.6. MÚLTIPLES PUNTOS DE CAMBIO EN ESPACIO TIEMPO61

Tabla 3.9: Puntos de cambio espacio tiempo.

	[13]	[14]	[15]	[16]
1	55.319548	36.493439	53.350957	37.905475
2	7.263903	27.399136	14.610914	25.226621
3	28.683089	26.696299	17.792607	28.209964
4	12.478724	12.478724	12.478724	12.478724
5	12.478724	12.478724	12.478724	12.478724
6	12.478724	12.478724	12.478724	12.478724
7	12.478724	12.478724	12.478724	12.478724
8	5.701123	5.701123	5.701123	5.701123
9	4.732525	4.732525	4.732525	4.732525
10	5.701123	5.701123	5.701123	5.701123
11	4.732525	4.732525	4.732525	4.732525
12	5.701123	5.701123	5.701123	5.701123
13	4.732525	4.732525	4.732525	4.732525
14	5.701123	5.701123	5.701123	5.701123
15	4.732525	4.732525	4.732525	4.732525

Fue programado en R, el programa se encuentra en el apéndice y está etiquetado como programa 2.

Se construyó un algoritmo para detectar 3 puntos de cambio en espacio tiempo en 16 posiciones, éste es dado a continuación:

Algoritmo 2: Detecta 3 puntos de cambio en espacio tiempo

Entrada La función INLA regresa variables **mp1**, **mp2**, y **mp** las cuales contienen los datos simulados con un modelo autoregresivo AR1 y de una familia Poisson para **datos A**, **datos B** y **datos**.

Salida La **matriz respuesta** guarda las verosimilitudes de 15 divisiones en 16 posiciones.

Inicializar Se inicializan los vectores **mp1**, **mp2**, y **mp** los cuales contienen los datos simulados con autoregresivo AR1 y de una familia Poisson para **datos A**, **datos B** y **datos**.

Para i en 1:n

Se calculan 4 parámetros **lambda** de un proceso Poisson aleatoriamente de 1 a 15 y se incluyen en el vector **vlambda**.

Se incluye en el vector **datos** que tiene espacio para 60 datos los 4 vectores simulados con puntos de cambio para los 4 parámetros lambda.

Fin Para se termina el ciclo **Para**

Para i en 1:n

El vector **datos** se divide de 1 a 30 y de 31 a 60 y se asigna a **datos A** y **datos B**.

i) Se calcula la verosimilitud de cada vector en **datos A** y **datos B** y se guarda en la **matriz respuesta**,

ii) Se dividen a la mitad los vectores **datos A** y **datos B** y se asignan los datos a **datos C** y **datos D**,

iii) Se calcula la verosimilitud para **datos C** y **datos D**.

iv) Se repiten los pasos i) y ii) con las nuevas divisiones de los datos hasta llegar a la división de solo dos datos.

v) se registran los resultados obtenidos.

Fin, se detiene

Regresa la matriz respuesta, que contiene las verosimilitudes de 15 divisiones en 16 posiciones.

3.7 Método de umbrales

El siguiente es un ejemplo de puntos de cambio con el método de umbrales, cuyo método es explicado en Altieri(2015), el cual consiste en asociar a los datos una probabilidad a posteriori y definir un umbral, entonces si la probabilidad es menor que ese umbral, existe un punto de cambio.

Para asociar probabilidades se da un ejemplo con regresión lineal múltiple:

$$Y_i = \beta_0 + \sum_{j=1}^4 \beta_j X_{j,i} + \epsilon_i.$$

Para este ejemplo se consideró la base de datos *cement* de Gómez(2020). En el modelo, Y_i representa el calor evolucionado de la observación i y $X_{j,i}$ es la proporción de la componente j en la observación i . El parámetro β_0 es un intercepto y β_j , $j = 1, \dots, 4$ son coeficientes asociados con las covariables. Finalmente, ϵ_i , $i = 1, \dots, n$ es un término de error con una distribución Gaussiana con media cero y precisión τ .

La base de datos *cement* se muestra en la Tabla 3.10.

Tabla 3.10: Datos de regresión.

X_1	X_2	X_3	X_4	Y
7	26	6	60	78.5
1	29	15	52	74.3
11	56	8	20	104.3
11	31	8	47	87.6
7	52	6	33	95.9
11	55	9	22	109.2
3	71	17	6	102.7
1	31	22	44	72.5
2	54	18	22	93.1
21	47	4	26	115.9
1	40	23	34	83.8
11	66	9	12	113.3
10	68	8	12	109.4

Para el ejemplo del punto de cambio se utilizó la variable X_1 , se le asignaron probabilidades de la distribución a posteriori a estos datos y posteriormente se fijó un umbral de $3.041631e - 07$, aquellos datos que tuvieran una probabilidad menor que este umbral son los puntos de cambio.

Se elaboró un programa en R (apéndice), y es etiquetado como programa 3.

El resultado es:

```
0 0 1 1 0 1 0 0 0 1 0 1 1
```

El cambio es donde empiezan los unos a ser más constantes, es decir, es donde al asignarle a los datos una probabilidad a posteriori y definir un umbral, entonces la probabilidad fue menor que ese umbral, por lo tanto existe un punto de cambio, esto sucede en la posición diez y tres, con un

umbral de 3.041631e-07.

Se creó un algoritmo el cual detecta múltiples puntos de cambio con el método de umbrales, éste se describe a continuación:

Algoritmo 3: Detecta múltiples puntos de cambio con el método de umbrales

Entrada Matriz cement, matriz datos, umbral, regresión en ml.

Salida Vector menores incluye ceros y unos, 1 significa que la probabilidad a posteriori es menor que el umbral y 0 que es mayor que el umbral.

Inicializar n=13 es el número de datos.

Para h en 1:n

Se calcula la probabilidad a posteriori de cada dato.

Fin, se detiene

Para h en 1:n

Si la probabilidad a posteriori es menor que el umbral se guarda un 1 en el vector **menores de lo contrario** si es mayor se guarda un cero.

Fin, se detiene

Imprime menores

3.8 CUSUM con datos de accidentes de tráfico

El método consiste en lo siguiente:

- 1.- Se calcula el promedio $\frac{x_1+x_2+\dots+x_{32}}{32}$,
2. Empieza la suma acumulativa en cero $S_0 = 0$,
3. Se calculan las otras sumas acumulativas de la forma $S_i = S_{i-1} + (X_i - \bar{X})$ para $i = 2, \dots, 32$,
 $S_0 = 0$,
 $S_1 = S_0 + (X_1 - \bar{X})$,
 $S_2 = S_1 + (X_2 - \bar{X})$,
 \dots

$$S_{32} = S_{31} + (X_{32} - \bar{X}).$$

Se realiza también un bootstrap, pero antes de ello se requiere un estimador de la magnitud del cambio. Una opción que funciona bien independientemente de la distribución y a pesar de los múltiples cambios es *Sdiff*.

$$Sdiff = Smax - Smin, \text{ donde, } Smax = \max_{i=0, \dots, 33} S_i,$$

$$Smin = \min_{i=0, \dots, 33} S_i.$$

Un análisis bootstrap consiste en realizar una gran cantidad de bootstrap y contar la cantidad de bootstrap para los cuales *S^odiff* es menor que *Sdiff*. Sea *N* el número de muestras de bootstrap realizado y sea *X* el bootstrap para los cuales *S^odiff* < *Sdiff*. Luego el nivel de confianza de que se produjo un cambio como porcentaje se calcula de la siguiente forma:

$$\text{Nivel de confianza} = 100 \frac{X}{N} \%.$$

La magnitud del cambio es *Sdiff*, nos da un valor positivo que indica que hubo un cambio de menor a mayor, es decir cambiaron los accidentes de tráfico que es el tema de este problema de menor a mayor. Enseguida se realiza un análisis de bootstrap, una sola rutina. El procedimiento es el siguiente:

1. Se generó una muestra de bootstrap de 32 unidades, denotada por $X_1^o, X_2^o, \dots, X_{32}^o$, reordenando aleatoriamente, los 32 valores originales. Esto se llama muestreo sin reemplazo.

2. Basado en las muestras bootstrap, se calcula el bootstrap CUSUM, denotado por $S_0^o, S_1^o, \dots, S_{32}^o$.

3. Se calcula el máximo y el mínimo y la diferencia de bootstrap CUSUM denotado por *S^omax*, *S^omin*, y *S^odiff*.

4. Se determina si la diferencia bootstrap *S^odiff* es menor que la diferencia *Sdiff*.

Esto se aplicó a datos de accidentes de tráfico en 32 ciudades, los datos fueron obtenidos de la página de INEGI y se presentan en la Tabla 3.11:

De acuerdo a la magnitud del cambio, la suma mínima es de -439.6875, y esta suma se encuentra en el dato con posición 7. Enseguida las estadísticas son:

$$\begin{aligned} Smax &= 183.875, \\ Smin &= -439.6875, \\ Sdiff &= 623.5625. \end{aligned}$$

Para 100 muestreos bootstrap el Nivel de confianza = $100 \frac{X}{N} \% = 83 \%$.

Tabla 3.11: Datos de accidentes de tráfico.

Datos	Datos	Datos	Datos
74	188	206	37
60	56	146	83
26	36	69	
24	351	113	
94	49	278	
22	171	208	
78	68	41	
270	42	136	
223	229	80	
104	36	140	

Se puede observar en la Figura 3.3 que el punto de cambio se encuentra en el valor 7, donde cambia de menor número de accidentes de tráfico a mayor número.

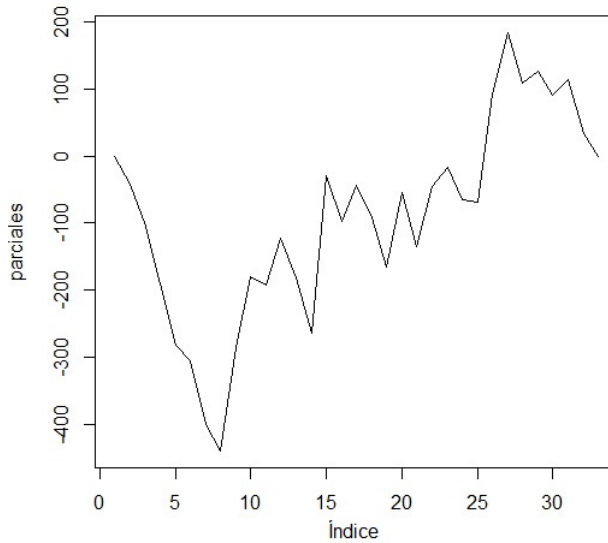


Figura 3.3: Gráfica con punto de cambio en 7.

Se aplicó lo que dice el algoritmo y se elaboró un programa, el cual se encuentra en el apéndice, etiquetado como programa 4.

Se creó un algoritmo con el método de suma acumulativa, éste detecta un solo punto de cambio, el algoritmo es mostrado a continuación:

Algoritmo 4: Detecta un solo punto de cambio

Entrada matriz datos

Salida **parciales** son las sumas acumulativas de los datos, **rparciales** son sumas acumulativas con bootstrap, **rSdif** son las diferencias de remuestreo.

Inicializar $n=33$ son 33 datos

$m=100$ es remuestreo de tamaño 100 con 33 datos

Para i en 2:n

Se calcula la suma acumulativa se guarda en **parciales**

Fin, se detiene

se calculan el máximo de los datos **Smax**=**max**(**parciales**), el mínimo **Smin**=**min**(**parciales**) y la diferencia **Sdif**=**Smax**-**Smin**

Para k en 2:n

Se calcula la suma acumulativa del remuestreo y se guarda en **rparciales**
Fin, se detiene

Se calcula el máximo de **rparciales** **rSmax**=**max**(**rparciales**)

el mínimo **rSmin**=**min**(**rparciales**), y la **rSdif**=**rSmax**-**rSmin**

Regresa rSdif regresa el vector de diferencias de remuestreo

Para j en 1:m se hace un remuestreo en bootstrap de tamaño 100 con 33 datos

Se calculan las diferencias de cada remuestreo

diferencias[m] = remuestreo(sample(datos, 32))

Fin, se detiene

Para j en 1:m

Si las diferencias del remuestreo son menores que **Sdif** la diferencia normal se guardan en el vector respuesta.

Fin, se detiene

Se calcula el intervalo de confianza

Nconfianza=100*sum(respuestas)/length(respuestas)

3.9 Otras Pruebas de datos de accidentes de tráfico

Se realizaron las pruebas de rangos de Buishand, Pettitt y prueba de homogeneidad normal estándar a los mismos datos de accidentes de tráfico, y para todas las pruebas el resultado fue el punto de cambio en la posición 7. Se utilizó la librería Trend de R, el código utilizado es el siguiente:

```
library(trend)
datos <- -c(74, 60, 26, 24, 94, 22, 78, 270, 223, 104, 188, 56, 36, 351, 49,
171, 68, 42, 229, 36, 206, 146, 69,
113, 278, 208, 41, 136, 80, 140, 37, 83)
y <- -ts(datos, start = c(1900), freq = 1)
x <- -ts(datos, start = c(1900), freq = 1)
prueba de rangos de bushand
br.test(y, m = 20000)
prueba petitt
pettitt.test(x)
prueba de homogeneidad normal estandar
snh.test(x, m = 20000)
```

3.10 CUSUM con datos de coronavirus

Se utilizaron datos de coronavirus de la página coronavirus.gob.mx con el método de suma acumulativa para detectar los puntos de cambio, los grupos de datos se tienen en la Tabla 3.12, se formaron tres grupos, uno con 8, el otro con 7 y el último con 5, los datos corresponden a mayo del 2020. En cada base de datos se detectó donde ocurrió el cambio de menor a mayor, es decir en donde empezaron a ocurrir un mayor número de casos de contagios, esto en todo el país. En la base de datos están divididos en contagios de hombres y mujeres, para analizar los cambios se sumaron y se presentan los totales de hombres más mujeres. La Tabla 3.12 muestra los datos utilizados.

Los datos de contagios se muestran en una gráfica de barras, en la Figura 3.4. Dicha gráfica muestra como disminuyen el número de contagios y vuelven a incrementar.

Los incrementos o cambios sucedieron en el primer grupo en el dato del 4 de mayo, en el segundo en el dato del 11 de mayo y en el último grupo en el dato 18 de mayo, esto se observa en la gráfica de sumas acumulativas CUSUM que se muestran en la Figura 3.5.

Tabla 3.12: Datos de contagios de coronavirus.

Fecha	datos	Fecha	datos	Fecha	datos
1	1225	9	1270	16	1483
2	1241	10	1161	17	1145
3	1121	11	2512	18	2747
4	2298	12	2428	19	2261
5	1840	13	2516	20	2083
6	2024	14	2353		
7	2085	15	2564		
8	2071				



Figura 3.4: Gráfica de contagios.

Para cada suma acumulativa las estadísticas y el porcentaje son los mostrados en la Tabla 3.13.

Tabla 3.13: Estadísticas y porcentaje.

Grupo 1	Grupo 2	Grupo 3
$Sdif = 1627.375$	$Sdif = 1798.714$	$Sdif = 1259.6$
$Smax = 0$	$Smax = 0$	$Smax = 0$
$Smin = -1627.375$	$Smin = -1798.714$	$Smin = -1259.6$
Nconfianza=89 %	83 %	61 %

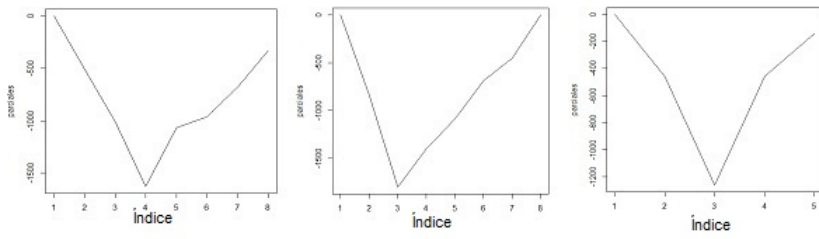


Figura 3.5: Datos grupos 1, 2 y 3

Capítulo 4

Conclusiones

- Se investigaron y analizaron metodologías desarrolladas en el estudio de puntos de cambio y sus diferentes aplicaciones.
- Muchos de los trabajos en puntos de cambio tanto paramétricos como no paramétricos aplican la metodología de puntos de cambio en precipitación, temperatura, cambio en la temperatura del aire, temperatura y flujo del agua, por mencionar algunos. Aunque estos problemas de aplicación se presentan con frecuencia, también existen algunos otros que han sido resumidos en las tablas del Capítulo 2, en la sección de aplicaciones, sin embargo, las variables hidrometeorológicas como se puede observar son de principal interés.
- Es de interés cuando se hace inferencia identificar los puntos de cambio ya que eso daría resultados más precisos o exactos.
- Se elaboraron programas para los métodos factor de Bayes, umbrales y CUSUM, la finalidad fue detectar numéricamente los puntos de cambio y comparar los resultados de las distribuciones a priori utilizadas, esto en el caso del factor de Bayes. Para los métodos umbrales y CUSUM el objetivo fue detectar los puntos de cambio. Se utilizaron bases de datos de coronavirus, tráfico de autos, de regresión (*cement*) y una base de datos simulada con un proceso Poisson.
- Se usaron los métodos Pettitt, homogeneidad normal estándar y rangos de Buishand, se aplicó el paquete *Trend de R* el cual contiene éstos métodos para detectar los puntos de cambio.
- Se puede concluir que los métodos de umbrales y de CUSUM detectan bien los puntos de cambio, la diferencia es que con el método de umbrales se detectan múltiples puntos de cambio en una regresión.
- Con el método CUSUM sólo se detecta un punto de cambio, esto también sucede con los métodos Pettitt, homogeneidad normal estándar y rangos de Buishand. Con el método CUSUM se detectan los cambios aún cuando la base de datos sea pequeña, que es lo que

sucedió con los datos de coronavirus.

- El método del factor de Bayes detectó bien los puntos de cambio, la debilidad del mismo es que detecta cambios hasta una división de cuatro, a menos datos en la división ya no se detectan los puntos de cambio.

Trabajo a futuro

- Implementar una metodología en espacio-tiempo para analizar puntos de cambio usando la convergencia a puentes Brownianos para obtener la distribución asintótica del estadístico de prueba desde el enfoque de máxima verosimilitud.
- Obtener la estadística de prueba por máxima verosimilitud para uno o varios puntos de cambio en problemas que involucran espacio y tiempo.
- Determinar la distribución límite de la estadística de prueba para puntos de cambio en problemas que involucran espacio y tiempo.
- Aplicar la distribución asintótica de la estadística de prueba para contrastar juegos de hipótesis de uno o varios puntos de cambio en espacio-tiempo.

Bibliografía

Abdel, H., El, S., Sean, B., Rong, Z., and Yalin, C. Modelling temporal and spatial changes of PCBs in fish tissue from Lake Huron. *Environ Monit Assess*, 173, 611-623. 2011.

Abdol Rassoul Zarei, Mohammad Mehdi Moghimi, Mohammad Reza Mahmoudi. Analysis of Changes in Spatial Pattern of Drought Using RDI Index in south of Iran. *Water Resour Manage* 30:3723–3743. 2016.

Adeyeri, Lamptey, Lawin, and Sanda. Spatio-Temporal Precipitation trend and homogeneity analysis in komadugu yobe Basin, Lake Chad region graduate research programme on climate change and water resources. *Climatol Weather Forecasting*, 5(3). 2017.

Akburak, S., and Makineci, E. Temporal changes of soil respiration under different tree species. *Environ. Monit Assess*, 185, 3349-3358. 2012.

Alexandersson, H. A homogeneity test applied to precipitation. *Journal of Climatology*, 6, 661-675. 1986.

Altieri Linda. A Bayesian changepoint analysis on spatio-temporal point processes. Thesis, Università di Bologna. 2015.

Anabela Carvalho, Mike D. Flannigan, Kim A. Logan, Lynn M. Gowman, Ana Isabel Miranda, Carlos Borrego. The impact of spatial resolution on area burned and fire occurrence projections in Portugal under climate change. *Climatic Change*.98:177–197. 2010.

Biana, G., Dua, J., Songa, M., Xua, Y., Xiea, S., Zhenga, W., and Xuc, C. A procedure for quantifying runoff response to spatial and temporal changes of impervious surface in Qinhuai River basin of southeastern China. *Catena*, 157, 268-278. 2017.

Bin Yong, Liliang Ren, Yang Hong, Jonathan J. Gourley, Xi Chen, Jinwei Dong, Weiguang Wang, Yan Shen, and Jill Hardy. Spatial–Temporal Changes of Water Resources in a Typical Semiarid Basin of North China over the Past 50 Years and Assessment of Possible Natural and Socioeconomic Causes. *Journal of Hydrometeorology*, vol.14. 2013.

Blangiardo Marta y Cameletti Michela. Spatial and Spatio-temporal Bayesian Models with R-INLA, John Wiley and Sons. 2015.

Bollinger, J., Kienast, F., Soliva, R., and Rhuterford, G. Spatial sensitivity of species habitat patterns to scenarios of land use change (Switzerland). *Landscape Ecol.*, 22, 773-789. 2007.

Brodsky, B. E., and Darkhovsky, B. S. Mathematics and its Applications Noparametrics Methods in Change-Point Problems (Vol. 243). Kluwer Academic Publishers. 1993.

Brodsky, B. E., and Darkhovsky, B. S. Non Parametric Diagnostic Problems and Methods. Mathematics and Its Applications. Kluwer Academic, Dordrecht. 2000.

Buishand, T. A. Some methods for testing the homogeneity of rainfall records. *Journal of Hidrology*, 58, 11-27. 1982.

Cai, B., Lawson, A., McDermott, S., and Aelion, C. A Bayesian semiparametric approach with change points for spatial ordinal data. *Stat. Methods Med. Res.*, 25(2), 644-658. 2016.

Cermak, J., Wild, M., Knutti, R., Mishchenko, M., and Heidinger, K. Consistency of global derived aerosol and cloud data sets with recent brightening observations. *Geophysical research Letters*, 37. 2010.

Chakraborty, D., Saha, S., Sink, R., Sethy, B., Kumar, A., Saikia, U., and Daschauthuri, D. Trend Analysis and Change Point Detection of North Eastern India. *Environ. Process*. 2017.

Chavaillaz, Y., Joussaume, S., Bony, S., and Braconnot, P. Spatial stabilization and intensification of moistening and drying rate patterns under future climate change. *Clim Dyn*, 47, 951-965. 2015.

Chen, J., and Gupta, A. Parametric Statistical Changepoint Analysis. Boga: Birkhauser. 2012

Chen, S., Kimb, S., and Kimc, T.-W. Spatial change Assessment of the observed and proyected extreme precipitation via change point detection. *Procedia Engineering*, 154, 1359-1367. 2016.

Chengjing, N., Hairong, L., Linsheng, Y., Ye, C., Erfu, D., Shaohong, W., and Yongfeng, L. Spatial and temporal changes in extreme temperature and extreme precipitation in Guangxi. *Quaternary International*, 263, 162-171. 2012.

Cubera, E., and Moreno, G. Effect of single Quercus ilex trees upon spatial and seasonal changes in soil water content in deshesas of central western Spain. *Ann. For. Sci.*, 64, 355-364. 2007.

Deepesh Machiwa, Madan Kumar Jha. Evaluating persistence and identifying trends and abrupt changes in monthly and annual rainfalls of a semiarid region in Western India. *Theor Appl Climatol* 128:689–708. 2016.

Gebremicael, T., Mohamed, Y., Zaag, P., and Hagos, E. Temporal and spatial changes of rainfall and streamflow in the Upper Tekeze-Atbara river basin Ethiopia. *Hydrol. Earth Syst Sci.*, 21, 2127-2142. 2017.

González Taboada, F., and Andón, R. Patterns of change in sea surface temperature in the North Atlantic during the last three decades: beyond mean trends. *Climatic Changes*, 115, 419-431. 2012.

Gómez Rubio Virgilio. *Bayesian inference with INLA*, Chapman and Hall. 2020.

Hall, P. On the rate of convergence of normal extremes. *J. Appl. Probab.* 16, 433-439. 1979.

Hawkins, D. M. Testing a sequence of observations for a shift in location. *Journal of the American Statistical Association*, 72, 180-186. 1977.

Hilas, C., Rekanos, L., and Mastorocostas, P. Change Point Detection in Time Series Using Higher-Order Statistics: A Heuristic Approach. *Mathematical Problems in Engineering*. 2013.

Hinkley, D. V. Inference about the point in a sequence of random variables, *Biometrika*, 57 1, 1-17. 1970.

Horvath, L. The maximum likelihood method for testing changes in the parameters of normal observations. *The Annals of Statistics*, Vol. 21, No. 2, pp. 671-680. 1993.

Howard Williams, C. Seasonal and spatial changes in the composition of the aquatic and semiaquatic vegetation of lake Chilima, Malawi. *Vegetatio* 30, 1: 33-39. 1975.

Huiying, Z., Lijuan, G., Huihui, Q., Haixia, Z., Xiufen, L., and Heilongjiang, Z. The climate change variations in the northern Greater Kuningan Mountains during the past centuries, Institute of Meteorological Science, Harbin 150030, China. *J. Geogr. Sci.*, 26(5), 585-602. 2016.

Inga Mau, Kathleen Stewart Hornsby, and Ian D. Bishop. Modeling Geospatial Events and Impact Through Qualitative Change, Department of Geomatics, University of Melbourne, Parkville, Victoria 3010, Australia. 2007.

James E. Perry and Carl H. Hershner. Temporal changes in the vegetation pattern in a tidal freshwater marsh. *Plant Ecology* 178:189–199. 1999.

Jansenberger, E., and Steinnocher, S. Dual Kernel Density Estimation

as a Method Describing Spatio Temporal Changes in the Upper Austrian Food Retailing Market. 7th AGILE Conference on Geographic Information Science 29 April- 1 Mayo 2004. 2014.

Jean D. G. and Subhabrata C. Nonparametric Statistical Inference. Fourth Edition, Marcel Dekker, Inc. 2003.

Jiang Xueding, Xia Beicheng, Lin Guangfa and Lin Wenshi. Daily Changes of Spatial Patterns of Meteorological Elements over, Pearl River Delta Based on GIS and MM5. *Chin. Geogra. Sci.* 19(1) 069–076. 2009.

Lapedra, I. Desarrollo de algoritmos de selección de variables para clasificadores neuronales: métodos no estocásticos. Tesis, Universitat Rovira Virgili. 2004.

Lee, Peter M. Bayesian Statistics an Introduction. Second Edition, Arnold. 2012.

Lepage, Y. A Combination of Wilcoxon's and Ansari Bradley's Statistics. *Biometrika*, 58(1), 213-217. 1971.

Lishan, R., Suiji, W., and Xiaoli, F. Channel change at Toudaoguai Station and its responses to the operation of upstream reservoirs in the upper Yellow River. *J. Geogr. Sci.*, 231-247. 2010.

Liu, Liu., Xu, Z., and Huang, J.-X. Spatio temporal variation and abrupt changes for major climate variables in the Taihu Basin, China. *Stch Environ Res Risk Assess*, 26, 777-791. 2012.

Luan, J., Liu, D., Zhang, L., Huang, Q., Feng, J., Lin, M., and Li, g. Analysis of the spatial-temporal change of the vegetation index in the upper reach of Han River Basin in 2000-2016. *Proc. IAHS*, 379, 287-292. 2018.

Luo, Q., Bryan, B., Bellotti, W., and Williams, M. Spatial analysis of environmental change impacts on wheat production in mid-lower north, south Australian. *Climatic Change*, 72, 213-228. 2005.

Ma, Z., Shi, J., Wang, G., and He, Z. Temporal changes in the inequality of growth of *Cunninghamia lanceolata* (Lamb) Hook: a novel application of the Gini coefficient and Lorenz asymmetry. *Genetica*, 126, 343-351. 2006.

Madhushree Munsli, G. Areendran, P. K. Joshi. Modeling spatio-temporal change patterns of forest cover: a case study from the Himalayan foothills (India). *Reg Environ Change*, 12:619–632. 2012.

Mahaffee, W. F. and Kloeppe J. W. Temporal Changes in the Bacterial Communities of Soil, Rhizosphere, and Endorhiza Associated with Field-Grown Cucumber (*Cucumis sativus* L.). *Microb Ecol* 34:210–223. 1996.

Malekian, A., and Kazemzadeh, M. Spatio-Temporal Analysis of Regional Trends and Shift Changes of Autocorrelated Temperature Series in Urmia

Lake Basin. *Water Recour Manage*, 30, 785-803. 2015.

Marcoe, K., and Pilson, S. Habitat change in the lower Columbia River estuary 1870-2009. *J. Coast Conserv*, 21, 505-525. 2017.

Meng Xiujing, Zhang Shifeng, Zhang Yongyong and Wang Cuicui. Temporal and spatial changes of temperature and precipitation in Hexi Corridor during 1955–2011. *J. Geogr. Sci.*, 23(4): 653-667. 2013.

Michalowska, K., Glowienka, E., and Pekala, A. Spatial-temporal detection of changes southern coast of the baltic sea based on multitemporal aerial photographs. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciencies*, XLI-B2. 2016.

Monica G. Turner, Robert V. O'Neill, Robert H. Gardner and Bruce T. Milne. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* vol. 3 nos. 3/4 pp 153-162. 1989.

Mordvinov, A. V. and Plyusnina, L. A. Cyclic changes in solar rotation inferred from temporal changes in the mean magnetic field. *Institute of Solar-Terrestrial Physics*. 2000.

Morkvina, V. Application of the singular spectrum analysis for change-point detection in time series. Ph. D. Thesis, Cardi University. 2001.

Morskvina, V., and Zhigljavsky, A. An on singular based on singular spectrum analysis for change-point detection, communication in statistics. *Stat. Simul.*, 32, 319-352. 2003.

Muñiz, L., Juárez, B. and Cruz, H. Change points in space-time, methodology and applications. *Journal of Quantitative and Statistical Analysis*. Vol.6 No.18, 17-28. 2019.

Otto, P. and Wolfgang, S. Detection of spatial change points in the mean and covariances of multivariate simultaneous autoregressive models. *J. Biometrical*, 58, 1113-1137. 2016.

Page, E. S. On problem in which a change in a parameter occurring at an unknown point. *Biometrika*, 44, 248-252. 1957.

Peter, A. Long-term temporal and spatial changes in the richness and relative abundance of the inshore fish community of the British North Sea Coast. *Journal of Sea Reseach*. 2017.

Pettitt, A. A Non-Parametric Approach to the Change-Point Problem. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 28(2), 126-135. 1979.

Quant, R. E. The estimation of parameters of a linear regression system obeying two separate regimes. *J. Amer. Statist. Assoc.*, 50, 853-880. 1958.

Ran L., Suiji W., and Xiaoli F. Channel change at Toudaoguai Station and

its responses to the operation of upstream reservoirs in the upper Yellow River. *J. Geogr. Sci.*, 20(2): 231-247. 2010.

Sarate C. Dass, Chae Young Lim, Tapabrata Maiti. Hierarchical Spatial Regression Models for change point Analysis. Section on Survey Research Methods – JSM. 2011.

Schoenberg, and Gotway. Statistical methods for spatial data analysis. CRC Press. 2004.

Shrutilipi, B., Titas, A., and Soumya, K. G. Spatio-temporal Change Detection of Urban Heat Islands Using Spatial Interpolation. LNCS 8251, pp. 222–228, 2007.

Skliris, N., Marsh, R., Josey, S. A., Good, S. A., Liu, C., and Allan, R. P. Salinity changes in the World Ocean since 1950 in relation to changing surface freshwater fluxes. *Clim Dyn*, 43, 709-736. 2014.

Snedecor, George W. and Cochran, William G. Statistical Methods, Eighth Edition, Iowa State University Press. 1989.

Sneyers. On the Statistical Analysis of series of Observations. Geneva. 1990.

Srivastava, M. S. and Worsley K. J. Likelihood ratio tests for a change in the multivariate normal mean. *Journal of the American Statistical Association*, Vol. 81, No. 393, pp.199-204. 1986.

Taylor Wayne A. *Change-Point Analysis: A Powerful New Tool For Detecting Changes*, Baxter Healthcare Corporation, Round Lake, IL 60073. 2000.

Tesfay G. G., Yasir A. M., Pieter V. Z. and Eyasu Y. H. Temporal and spatial changes of rainfall and streamflow in the Upper Tekeze-Atbara river basin, Ethiopia. *Hydrol. Earth Syst. Sci.*, 21, 2127–2142. 2017

Thorsten Pohlert. *Non-Parametric Trend Tests and Change-Point Detection*, Package R. 2020.

Tierney, T., and Cushman, J. Temporal changes in native exotic vegetation and soil characteristics following disturbances by feral pigs in a California grassland. *Biological Invasions*, 8, 1073-1089. 2005.

Tlili, A., Corcoll, N., Bonet, B., Moirin, S., Montuelle, B., Bérard, A., and Guasch, H. In situ spatio-temporal changes in pollution-induced community tolerance to zinc in autotrophic and heterotrophic biofilm communities. *Ecotoxicology*, 20, 1823-1839. 2011.

Tomozeiu, R., Busuioc, A., Marletto, V., Zinoni, F. and Cacciamani C. Detection of changes in the summer precipitation time series of the region Emilia-Romagna, Italy. *Theor. Appl. Climatol.* 67. 193-200. 2013.

Von Newman, J. Distribution of the Ratio of the Mean Square Successive Difference to the variance. *The Annals of Mathematical Statistics*, 12(4), 367-395. 1941.

Wang Suiji, Yan Yunxia and Yan Ming, Zhao. Quantitative estimation of the impact of precipitation and human activities on runoff change of the Huangfuchuan River Basin. *J. Geogr. Sci.*, 22(5): 906-918. 2012.

William M. Boldstad . *Introduction to Bayesian Statistic*. Second Edition, Wiley. 2007.

Won-Ho Nam, Eun-Mi Hong and Guillermo A. Baigorria. How climate change has affected the spatiotemporal patterns of precipitation and temperature at various time scales in North Korea. *International Journal of Climatology*. 2015.

Xiaomeng Song, Jianyun Zhang¹, Amir AghaKouchak, Shouraseni Sen Roy, Yunqing Xuan, Guoqing Wang¹, Ruimin He, Xiaojun Wang, and Cuishan Liu¹. Rapid urbanization and changes in spatiotemporal characteristics of precipitation in Beijing metropolitan area. *Journal of Geophysical Research: Atmospheres*, 119, 11, 250–11, 271. 2014.

Xin Zhao and Pao Shin Chu. Bayesian Multiple Changepoint Analysis of Hurricane Activity in the Eastern North Pacific: A Markov Chain Monte Carlo Approach. *Journal of Climate*. 2005.

Xiujing, M., Shifeng, Z., Yongyong, Z., and Cuicui, W. Temporal and spatial changes of temperature and precipitation in Hexi Corridor during 1955-2011. *J. Geogr. Sci.*, 23(4), 653-667. 2013.

Xun, Z., Shashi, S., and Reem, Y. Spatial temporal change footprint pattern discovery: an inter-disciplinary survey. *WIRES Data Mining Knowl Discov*, 4, 1-23. 2014.

Yang, T., Chen, X., Xu, C., and Zhang, Z.-C. Estimation as a Method Describing Spatio Temporal Changes in the Upper Austrian Food Retailing Market. *Stochastic Environmental Research and Risk Assessment*, 23, 1071-1087. 2009.

YanJun Xua, Qingzhao Yua, Richard Scribnerb, Katherine Theallc, Scott Scribnerb, and Neal Simonsenb. Multilevel Spatiotemporal Change-point Models for Evaluating the Effect of an Alcohol Outlet Control Policy on Changes in Neighborhood Assaultive Violence Rates. *Spat Spatiotemporal Epidemiol*, 3(2), 121–128. 2012.

Yao, Y. C. and Davis, R. A. The asymptotic behavior of the likelihood ratio statistics for testing shift in mean in a sequence of independent normal variates. *Sankhya*, A48, 339-353. 1986.

Yankova, Y., Villiger, J., Pernthaler, J., Schanz, F., and Posch, T. Prolongation, deepening and warming of the metalimnion change habitat condi-

tions of the harmful filamentous cyanobacterium *Planktothrix rubescens* in a prealpine lake. *Hydrobiologia*, 776, 125-138. 2017.

Yanyu Yin, Hui Liu, Xiangsheng Yi, and Weidong Liu. Spatiotemporal Variation and Abrupt Change Analysis of Temperature from 1960 to 2012 in the Huang-Huai-Hai Plain, China. *Advances in Meteorology*, Vol. 2015. 2015.

Yu Sun, Susanna T. Y. Tong, Mao Fang and Y. Jeffrey Yang. Exploring the effects of population growth on future land se change in the Vegas Wash watershed: an integrated approach of geospatial modeling and analytics. *Environ Dev Sustain* 15:1495–1515. 2013.

Yin, Y., Liu, H., Yi, X., and Liu, W. Spatiotemporal Variation and Abrupt Analysis of Temperature from 1960 to in the Huang-Huai-Hai Plain, China. *Advances in Meteorology*. 2015.

Zhao Huiying, Gong Lijuan, Qu Huihui, Zhu Haixia, Li Xiufen and Zhao Fang Heilongjiang . The climate change variations in the northern Greater Khingan Mountains during the past centuries Institute of Meteorological Science, Harbin 150030, China. *J. Geogr. Sci.* 26(5): 585-602. 2016.

Zun, Y. Y., Qinfeng G., Hai R. and Shao-Lin P.. Seasonal changes in spatial patterns of two annual plants in the Chihuahuan Desert, USA. *Plant Ecology*. 178: 189–199. 2005.

Apéndice

Programa 1

```
rm(list = ls())
library(poisson)
library(INLA)
vector <- -hpp.event.times(2, 15, num.sims = 1, t0 = 0)
plot(vector)
vector1 <- -hpp.event.times(1, 15, num.sims = 1, t0 = 0)
vector2 <- -hpp.event.times(4, 15, num.sims = 1, t0 = 0)
vector4 <- -hpp.event.times(6, 15, num.sims = 1, t0 = 0)
datos <- -c(vector, vector1, vector2, vector4)
plot(datos, type = "l")
datosA = datos[1 : 30]
datosB = datos[31 : 60]
datos
datosA
datosB
vec_res = rep(0, 20)
factor_b <- -function(datosA, datosB, datos){
  p1 <- -data.frame("num" = seq(1, length(datosA), 1), "datosA" = datosA)
  p2 <- -data.frame("num" = seq(1, length(datosB), 1), "datosB" = datosB)
  mp1 <- -inla(num f(datosA, model = "ar1"), data = p1, family = "poisson")
  mp2 <- -inla(num f(datosB, model = "ar1"), data = p2, family = "poisson")
  midf <- -data.frame("num" = seq(1, length(datos), 1), "datos" = datos)
  mp <- -inla(num f(datos, model = "ar1"), data = midf, family = "poisson")
  respuesta <- -as.numeric(mp1$mlik[1, 1]) + as.numeric(mp2$mlik[1, 1]) -
  as.numeric(mp$mlik[1, 1])
  return(respuesta)
}

vec_res[1] = factor_b(datosA, datosB, datos)
vec_res

datosC = datosA[1 : 15]
datosD = datosA[16 : 30]
vec_res[2] = factor_b(datosC, datosD, datosA)
vec_res
datosE = datosB[1 : 15]
```

```

datosF = datosB[16 : 30]
vec_res[3] = factor_b(datosE, datosF, datosB)
vec_res
datosG = datosC[1 : 8]
datosH = datosC[9 : 15]
vec_res[4] = factor_b(datosG, datosH, datosC)
vec_res
datosI = datosD[1 : 8]
datosJ = datosD[9 : 15]
vec_res[5] = factor_b(datosI, datosJ, datosD)
vec_res
datosK = datosE[1 : 8]
datosL = datosE[9 : 15]
vec_res[6] = factor_b(datosK, datosL, datosE)
vec_res
datosM = datosF[1 : 8]
datosN = datosF[9 : 15]
vec_res[7] = factor_b(datosM, datosN, datosF)
vec_res
datosO = datosG[1 : 4]
datosP = datosG[5 : 8]
vec_res[8] = factor_b(datosO, datosP, datosG)
vec_res
datosQ = datosH[1 : 4]
datosR = datosH[5 : 7]
vec_res[9] = factor_b(datosQ, datosR, datosH)
vec_res
datosS = datosI[1 : 4]
datosT = datosI[5 : 8]
vec_res[10] = factor_b(datosS, datosT, datosI)
vec_res
datosU = datosJ[1 : 4]
datosV = datosJ[5 : 7]
vec_res[11] = factor_b(datosU, datosV, datosJ)
vec_res
datosX = datosK[1 : 4]
datosY = datosK[5 : 8]
vec_res[12] = factor_b(datosX, datosY, datosK)
vec_res
datosW = datosL[1 : 4]
datosZ = datosL[5 : 7]
vec_res[13] = factor_b(datosW, datosZ, datosL)
vec_res
datosAA = datosM[1 : 4]
datosBB = datosM[5 : 8]
vec_res[14] = factor_b(datosAA, datosBB, datosM)
vec_res
datosCC = datosN[1 : 4]
datosDD = datosN[5 : 7]
vec_res[15] = factor_b(datosCC, datosDD, datosN)

```

vec_res

Para utilizar otra función a priori se cambia en el programa por la siguiente instrucción, donde se incluye la distribución a priori a utilizar y sus parámetros, ya el paquete INLA tiene las distribuciones a priori a utilizar.

```
prec.prior <- list(prec = list(prior = "logtgaussian", param = c(0, 0.001)))
mp1 <- inla(num f(datosA, model = "ar1", hyper = prec.prior), data =
p1, family = "poisson")
```

Programa 2

```
rm(list = ls())
library(poisson)
library(INLA)
esp_t = matrix(0, 60, 16)
matriz_resp = matrix(0, 15, 16)
mlambdas = matrix(0, 16, 4)

factor_b <- function(datosA, datosB, datos){
p1 <- data.frame("num" = seq(1, length(datosA), 1), "datosA" = datosA)
p2 <- data.frame("num" = seq(1, length(datosB), 1), "datosB" = datosB)
mp1 <- inla(num f(datosA, model = "ar1"), data = p1, family = "poisson")
mp2 <- inla(num f(datosB, model = "ar1"), data = p2, family = "poisson")
midf <- data.frame("num" = seq(1, length(datos), 1), "datos" = datos)
mp <- inla(num f(datos, model = "ar1"), data = midf, family = "poisson")

respuesta <- as.numeric(mp1$mlik[1, 1]) + as.numeric(mp2$mlik[1, 1]) -
as.numeric(mp$mlik[1, 1])
return(respuesta)
}

for(i in 1:16){
vlambda = sample(1:15, 4)
vector <- hpp.event.times(vlambda[4], 15, num.sims = 1, t0 = 0)
plot(vector)
mlambdas[i,] = vlambda
vector1 <- hpp.event.times(vlambda[1], 15, num.sims = 1, t0 = 0)
vector2 <- hpp.event.times(vlambda[2], 15, num.sims = 1, t0 = 0)
vector4 <- hpp.event.times(vlambda[3], 15, num.sims = 1, t0 = 0)
datos <- c(vector, vector1, vector2, vector4)
esp_t[, i] = datos
}
for(indice in 1:16){
datosA = esp_t[1:30, indice]
datosB = esp_t[31:60, indice]
datos = esp_t[, indice]
datosA
datosB
matriz_resp[1, indice] = factor_b(datosA, datosB, datos)
```

```

datosC = datosA[1 : 15]
datosD = datosA[16 : 30]
matriz_resp[2, indice] = factor_b(datosC, datosD, datosA)
datosE = datosB[1 : 15]
datosF = datosB[16 : 30]
matriz_resp[3, indice] = factor_b(datosE, datosF, datosB)
datosG = datosC[1 : 8]
datosH = datosC[9 : 15]
matriz_resp[4, indice] = factor_b(datosG, datosH, datosC)
datosI = datosD[1 : 8]
datosJ = datosD[9 : 15]
matriz_resp[5, indice] = factor_b(datosI, datosJ, datosD)
datosK = datosE[1 : 8]
datosL = datosE[9 : 15]
matriz_resp[6, indice] = factor_b(datosK, datosL, datosE)
datosM = datosF[1 : 8]
datosN = datosF[9 : 15]
matriz_resp[7, indice] = factor_b(datosM, datosN, datosF)
datosO = datosG[1 : 4]
datosP = datosG[5 : 8]
matriz_resp[8, indice] = factor_b(datosO, datosP, datosG)
datosQ = datosH[1 : 4]
datosR = datosH[5 : 7]
matriz_resp[9, indice] = factor_b(datosQ, datosR, datosH)
datosS = datosI[1 : 4]
datosT = datosI[5 : 8]
matriz_resp[10, indice] = factor_b(datosS, datosT, datosI)
datosU = datosJ[1 : 4]
datosV = datosJ[5 : 7]
matriz_resp[11, indice] = factor_b(datosU, datosV, datosJ)
datosX = datosK[1 : 4]
datosY = datosK[5 : 8]
matriz_resp[12, indice] = factor_b(datosX, datosY, datosK)
datosW = datosL[1 : 4]
datosZ = datosL[5 : 7]
matriz_resp[13, indice] = factor_b(datosW, datosZ, datosL)
datosAA = datosM[1 : 4]
datosBB = datosM[5 : 8]
matriz_resp[14, indice] = factor_b(datosAA, datosBB, datosM)
datosCC = datosN[1 : 4]
datosDD = datosN[5 : 7]
matriz_resp[15, indice] = factor_b(datosCC, datosDD, datosN)
indice = indice + 1
}
matriz_resp
mlambdas

```

Programa 3

```
library(INLA)
```

cement

```
m1 <- inla(y x1 + x2 + x3 + x4, data = cement)
```

```
dato <- c(7, 1, 11, 11, 7, 11, 3, 1, 2, 21, 1, 11, 10)
```

```
m1$marginals.fixed$x1
```

```
1 - inla.pmargin(0, m1$marginals.fixed$x1)
```

```
probas = seq(1, 13)
```

```
for(h in 1 : 13){
```

```
  probas[h] = inla.pmargin(dato[h] + 0.2, m1$marginals.fixed$x1)
```

```
  -inla.pmargin(dato[h] - 0.2, m1$marginals.fixed$x1)
```

```
}
```

```
menores = seq(1, 13)
```

```
umbral = 3.041631e - 07
```

```
for(h in 1 : 13){
```

```
  menores[h] = probas[h] < umbral
```

```
}
```

```
menores
```

Programa 4

```
datos <- c(74, 60, 26, 24, 94, 22, 78, 270, 223, 104, 188, 56, 36
```

```
, 351, 49, 171, 68, 42, 229, 36, 206,
```

```
146, 69, 113, 278, 208, 41, 136, 80, 140, 37, 83)
```

```
promd <- -mean(datos)
```

```
promd
```

```
parciales <- -seq(1, 33)
```

```
parciales[1] = 0
```

```
for(iin2 : 32){
```

```
  parciales[i] = parciales[i - 1] + datos[i - 1] - promd
```

```
}
```

```
parciales
```

```
Smax = max(parciales)
```

```
Smin = min(parciales)
```

```
Sdif = Smax - Smin
```

```
Sdif
```

```
Smax
```

```
Smin
```

```
plot(parciales, type = "l")
```

```
remuestreo <- -function(rdatos){
```

```
  rpromd <- -mean(rdatos)
```

```
  rparciales <- -seq(1, 33)
```

```
  rparciales[1] = 0
```

```
  for(k in 2 : 33){
```

```
    rparciales[k] = rparciales[k - 1] + rdatos[k - 1] - rpromd
```

```
  }
```

```
  rSmax = max(rparciales)
```

```
  rSmin = min(rparciales)
```

```
  rSdif = rSmax - rSmin
```

```
return(rSdif)
}

diferencias = seq(1, 100)
for(m in 1:100){
diferencias[m] = remuestreo(sample(datos, 32))
}
diferencias

respuestas <- seq(1, 100)
for(m in 1:100){
respuestas[m] = diferencias[m] < Sdif
}
respuestas
diferencias
Sdif

NConfianza = 100 * sum(respuestas)/length(respuestas)
NConfianza
```