



# COLEGIO DE POSTGRADUADOS

---

---

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN  
EN CIENCIAS AGRÍCOLAS

**CAMPUS MONTECILLO**

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E  
INFORMÁTICA ESTADÍSTICA

**PRUEBA PARA LA DISTRIBUCIÓN POISSON  
INFLADA CON CEROS**

Patricia Estrada Drouaillet

TESIS

PRESENTADA COMO REQUISITO PARCIAL PARA  
OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO  
2018

---

---



# COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

## CARTA DE CONSENTIMIENTO DE USO DE LOS DERECHOS DE AUTOR Y DE LAS REGALÍAS COMERCIALES DE PRODUCTOS DE INVESTIGACIÓN

En adición al beneficio ético, moral y académico que he obtenido durante mis estudios en el Colegio de Postgraduados, el que suscribe, **“Patricia Estrada Drouaillet”**, Alumna de esta Institución, estoy de acuerdo en ser partícipe de las regalías económicas y/o académicas, de procedencia nacional e internacional, que se deriven del trabajo de investigación que realicé en esta institución, bajo la dirección del Profesor **“Dr. Humberto Vaquera Huerta”**, por lo que otorgo los derechos de autor de mi tesis **“PRUEBA PARA LA DISTRIBUCIÓN POISSON INFLADA CON CEROS”**, y de los productos de dicha investigación al Colegio de Postgraduados. Las patentes y secretos industriales que se puedan derivar serán registrados a nombre del colegio de Postgraduados y las regalías económicas que se deriven serán distribuidas entre la Institución, El Consejero o Director de Tesis y el que suscribe, de acuerdo a las negociaciones entre las tres partes, por ello me comprometo a no realizar ninguna acción que dañe el proceso de explotación comercial de dichos productos a favor de esta Institución.

Montecillo, Mpio. de Texcoco, Edo. de México, a 10 de Julio de 2018.

---

Patricia Estrada Drouaillet

---

Vo. Bo. Dr. Humberto Vaquera Huerta

La presente tesis titulada: **PRUEBA PARA LA DISTRIBUCIÓN POISSON INFLADA CON CEROS**, realizada por la alumna: **Patricia Estrada Drouaillet**, bajo la dirección del Consejo Particular indicado ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**MAESTRA EN CIENCIAS**

**SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA**

**CONSEJO PARTICULAR**

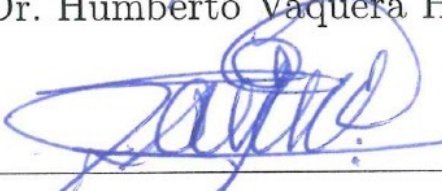
CONSEJERO



---

Dr. Humberto Vaquera Huerta

ASESOR



---

Dr. José Aurelio Villaseñor Alva

ASESOR



---

Dr. Barry C. Arnold

Montecillo, Texcoco, Estado de México, Junio de 2018

# PRUEBA PARA LA DISTRIBUCIÓN POISSON INFLADA CON CEROS

Patricia Estrada Drouaillet, Maestra en ciencias

Colegio de Postgraduados, 2018

## RESUMEN

En el presente trabajo se propone el uso de un intervalo de credibilidad para probar si una población pertenece a una distribución Poisson o una distribución Poisson inflada con ceros. Se desarrolla un estudio de la convergencia del algoritmo muestreador de Gibbs implementado para obtener la distribución posterior del parámetro  $\omega$  en el que esta basado la prueba. Se describe una aplicación que utiliza los datos recogidos por una encuesta realizada en Bangladesh en 2011 a mujeres y hombres en edad reproductiva, se les preguntó a las mujeres cuantos hijos menores de cinco años se les han muerto. Además, a través de un estudio de simulación se comparó el desempeño de la prueba del intervalo propuesto con otras pruebas conocidas.

**Palabras clave:** Distribución Poisson inflada con ceros, Muestreador de Gibbs, hdi.

# TEST FOR THE ZERO-INFLATED POISSON DISTRIBUTION

Patricia Estrada Drouaillet, Master of Science

Colegio de Postgraduados, 2018

## ABSTRACT

In this work we propose the use of a credibility interval to test if a population belongs to a Poisson distribution or to a Zero-Inflated Poisson. A study of the convergence of the implemented algorithm Gibbs sampler is developed to obtain the posterior distribution of the parameter  $\omega$  on which the test is based. An application that uses the data collected by a survey conducted in Bangladesh 2011 to women and men at reproductive age is described, they were asked how many children under five years have died. In addition, through a simulation study, the performance of the proposed interval test was compared with other tests.

**Key words:** Zero-Inflated Poisson, Gibbs sampler, hdi.

## AGRADECIMIENTOS

A **CONACYT**, por su apoyo económico otorgado para la realización de este trabajo de investigación.

A **COLPOS**, campus Montecillo, por brindarme la oportunidad de seguir esta formación continua.

A los integrantes de mi Consejo Particular:

Dr. Humberto Vaquera Huerta, por su paciencia, apoyo y valiosa dirección en la realización de éste trabajo.

Dr. José Aurelio Villaseñor Alva, por su tiempo dedicado, observaciones y comentarios para mejorar este trabajo de investigación.

Dr. Barry C. Arnold, por formar parte en la realización de esta tesis.

A mis profesores, por su esfuerzo dedicado a mi educación.

A mis compañeras, que supieron aceptarme para complementarnos con nuestra debilidades y fortalezas.

## DEDICATORIA

A **Dios**, que me hizo ser valiente en cada situación que se me presentó.

A mi **esposo**, por brindarme su inmenso amor al apoyarme y comprenderme a cada momento.

A mis **padres y hermanos**, por estar siempre presentes.

# CONTENIDO

<b>RESUMEN</b>	<b>iv</b>
<b>ABSTRACT</b>	<b>v</b>
<b>1. INTRODUCCIÓN</b>	<b>1</b>
<b>2. OBJETIVOS</b>	<b>3</b>
2.1. Objetivo general . . . . .	3
2.2. Objetivos particulares . . . . .	3
<b>3. MARCO TEÓRICO</b>	<b>4</b>
3.1. Prueba de hipótesis . . . . .	4
3.1.1. Lema de Neyman-Pearson . . . . .	5
3.1.2. Tipos de hipótesis . . . . .	6
3.1.3. Errores de decisión . . . . .	8
3.1.4. Función potencia . . . . .	10
3.1.5. Tamaño de la prueba . . . . .	11
3.2. Pruebas para probar distribución Poisson contra distribución Poisson inflada con ceros . . . . .	12
3.2.1. Estadístico de prueba Score . . . . .	13



## CONTENIDO

---

3.2.2. Estadístico de prueba Rao-Chakravarti . . . . .	13
3.2.3. Estadístico de prueba Cochran . . . . .	14
3.2.4. Estadístico de prueba de razón de verosimilitud generalizada . . . . .	15
3.3. Inferencia bayesiana . . . . .	15
3.3.1. Muestreador de Gibbs . . . . .	18
<b>4. DISTRIBUCIÓN POISSON INFLADA CON CEROS</b>	<b>22</b>
4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros . . . . .	23
4.2. Otras propiedades de la distribución Poisson inflada con ceros . . . . .	28
<b>5. PRUEBA PROPUESTA</b>	<b>30</b>
5.1. Distribución Poisson inflada con ceros . . . . .	30
5.2. Función de verosimilitud . . . . .	31
5.3. Inferencia Bayesiana . . . . .	31
<b>6. ESTUDIO DE PRUEBA PROPUESTA</b>	<b>34</b>
6.1. Diagnóstico de convergencia del algoritmo Gibbs . . . . .	34
6.2. Autocorrelación . . . . .	39
6.3. Función posterior . . . . .	41
6.4. Intervalo de credibilidad (Estadística Bayesiana) . . . . .	43
6.4.1. Potencia del intervalo de credibilidad . . . . .	43
6.4.2. Tabla de valores críticos . . . . .	45
6.4.3. Inferencia con el intervalo de credibilidad . . . . .	47
6.5. Intervalo de confianza (Estadística clásica) . . . . .	47

## CONTENIDO

---

6.5.1. Potencia del intervalo de confianza . . . . .	47
<b>7. EJEMPLO DE APLICACIÓN</b>	<b>49</b>
7.1. Encuesta sobre mortalidad de niños menores de 5 años . . . . .	49
7.2. Número de raíces en brotes micropropagados . . . . .	53
<b>8. CONCLUSIONES</b>	<b>55</b>
<b>REFERENCIAS</b>	<b>55</b>
<b>APÉNDICES</b>	<b>58</b>
Apéndice A: Programas de R usados en el trabajo . . . . .	58
Apéndice B: Programa de SAS usado en el trabajo . . . . .	64

# LISTA DE TABLAS

3.1. Tabla de potencia para el estadístico de prueba Score. . . . .	13
3.2. Tabla de potencia para el estadístico de prueba R. . . . .	14
3.3. Tabla de potencia para el estadístico de prueba Cochran. . . . .	14
3.4. Tabla de potencia para el estadístico de prueba de razón de verosimilitud generalizada. . . . .	15
6.1. Diagnóstico de Gelman para el parámetro $\omega$ . . . . .	37
6.2. Diagnóstico de Gelman para el parámetro $\lambda$ . . . . .	37
6.3. Potencia de la prueba. . . . .	44
6.4. Potencia de la prueba. . . . .	44
6.5. Potencia de la prueba. . . . .	44
6.6. Tabla de valores críticos. . . . .	45
6.7. Tabla de potencia para el estadístico basado en un intervalo de confianza. . . . .	48
7.1. Mortalidad de menores de 5 años en Bangladesh. . . . .	49
7.2. Estadísticas descriptivas para los datos de mortalidad de menores de 5 años en Bangladesh. . . . .	50
7.3. Estadísticas descriptivas para los datos de mortalidad de menores de 5 años en Bangladesh obtenidas en software SAS. . . . .	51

## LISTA DE TABLAS

---

7.4. Frecuencia de número de raíces en los brotes micropropagados. . . . .	53
--	----

# LISTA DE FIGURAS

3.1. Decisión estadística y errores al probar la hipótesis nula. . . . .	9
3.2. Representación gráfica de las probabilidades de error tipo I y tipo II, usando una prueba $t$ como ejemplo. . . . .	9
3.3. Función potencia para ejemplo. . . . .	11
4.1. Función de densidad con $\omega = 0.3$ . . . . .	24
4.2. Función acumulada con $\omega = 0.3$ . . . . .	25
4.3. Función de densidad con $\omega = 0.7$ . . . . .	26
4.4. Función acumulada con $\omega = 0.7$ . . . . .	27
6.1. Ejemplo de diagnóstico de Gelman y Rubin. . . . .	35
6.2. Ejemplo de diagnóstico de Gelman y Rubin. . . . .	36
6.3. Diagnóstico de Gelman y Rubin del parámetro $\omega$ . . . . .	37
6.4. Diagnóstico de Gelman y Rubin del parámetro $\lambda$ . . . . .	38
6.5. Cadena de Markov del parámetro $\omega$ . . . . .	39
6.6. Cadena de Markov del parámetro $\lambda$ . . . . .	39
6.7. Función de autocorrelación para el parámetro $\omega$ . . . . .	40
6.8. Función de autocorrelación para el parámetro $\lambda$ . . . . .	40
6.9. Función posterior del parámetro $\omega$ para $\lambda = 5$ . . . . .	41

## LISTA DE FIGURAS

---

6.10. Función posterior del parámetro $\omega$ para $\lambda = 10$ . . . . .	42
6.11. Función posterior para el parámetro $\lambda$ . . . . .	43
6.12. Ajuste de valores críticos para $\lambda = 1$ y $\alpha = 0.01$ . . . . .	46
6.13. Ajuste de valores críticos para $\lambda = 1$ y $\alpha = 0.05$ . . . . .	46
7.1. Función posterior para el parámetro $\omega$ . . . . .	50
7.2. Función posterior para el parámetro $\lambda$ . . . . .	51
7.3. Función posterior para el parámetro $\omega$ . . . . .	52
7.4. Función posterior para el parámetro $\lambda$ . . . . .	52
7.5. Función posterior para $\omega$ . . . . .	54

# Capítulo 1

## INTRODUCCIÓN

Las variables discretas surgen en disciplinas tales como la demografía ([Mamun, 2014](#)), la ingeniería ([Xie \*et al.\*, 2001](#)), la biología ([Thas y Rayner, 2005](#)), la agricultura ([Ridout y Hinde, 1998](#)), entre otras. Cuando la variable es de tipo conteo los datos observados se deben modelar estadísticamente con distribuciones discretas, por ejemplo, la Poisson o la binomial negativa, entre otras. Los datos de conteo a veces exhiben una mayor proporción de ceros y estos exceden a la frecuencia que se espera observar bajo la distribución que se ajusta. En este caso se dice que los datos presentan exceso de ceros o están inflados con ceros. Una alternativa para este tipo de datos es la distribución Poisson inflada con ceros (ZIP) (Véase el Capítulo 4).

Los ceros se clasifican en dos tipos: ceros estructurales y ceros muestrales. Los primeros ocurren con probabilidad  $\omega$  y produce sólo ceros, mientras que los otros ocurren con probabilidad  $(1 - \omega)$  y son de una distribución Poisson con media  $\lambda$ .

El interés de este estudio se centra en desarrollar y proponer una prueba para comparar o probar la distribución Poisson contra la distribución Poisson inflada con ceros, estudiar la potencia de prueba y posteriormente compararla con algunas pruebas ya establecidas, como la prueba de Score ([den Broek, 1995](#)) o la de Cochran ([Cochran, 1954](#)), entre otras.

La prueba propuesta está basada en un intervalo de credibilidad sobre el parámetro de forma de la distribución Poisson inflada con ceros, para lo que se recurre a métodos computacionales como el muestreador de Gibbs para poder obtener la distribución a posteriori del parámetro de interés, y de esta manera realizar inferencia sobre la muestra obtenida.

La estructura de este trabajo es la siguiente. En el capítulo 3 se tiene una pequeña revisión de literatura referente a prueba de hipótesis y debido a que éste trabajo se desarrolla desde el enfoque bayesiano de la estadística, se presenta una breve in-

## 1. INTRODUCCIÓN

---

roducción a los conceptos básicos de la estadística bayesiana. En el capítulo 4 se citan algunas características importantes de la distribución Poisson inflada con ceros. Posteriormente, en el capítulo 5 se describe la prueba propuesta para contrastar la distribución Poisson contra la distribución Poisson inflada con ceros. En el capítulo 6 se realiza un estudio de la prueba propuesta utilizando MCMC (o Markov Chain Monte Carlo), en el que se prueba el diagnóstico de Gelman y Rubin, y algunos otras pruebas gráficas a la cadena de Markov del muestreador de Gibbs. En el capítulo 7 se desarrolla un ejemplo de aplicación con datos de una encuesta que se realiza en Bangladesh, utilizando la frecuencia de muertes de niños menores de 5 años en el año 2011. En el último capítulo, se muestran los resultados obtenidos y las conclusiones.



# Capítulo 2

## OBJETIVOS

### 2.1. Objetivo general

- Proponer una prueba para contrastar la distribución Poisson contra la distribución Poisson inflada con ceros.

### 2.2. Objetivos particulares

- Escribir un programa computacional para obtener resultados de la prueba propuesta.
- Realizar un estudio de tamaño y potencia de la prueba propuesta.
- Comparar la prueba propuesta con pruebas ya establecidas.
- Implementar la prueba propuesta a un ejemplo de aplicación.

# Capítulo 3

## MARCO TEÓRICO

### 3.1. Prueba de hipótesis

La definición de una prueba de hipótesis es muy general, pero lo importante es que una hipótesis hace una declaración sobre la población. El objetivo de una prueba de hipótesis es decidir, basándose en una muestra de la población, cuál de las dos hipótesis complementarias es verdadera.

Las dos hipótesis complementarias en un problema de prueba de hipótesis se llaman hipótesis nula e hipótesis alternativa. Las hipótesis son denotadas por  $H_0$  y  $H_1$ , respectivamente.

Una importante subclase de problemas de decisión implica solo dos acciones posibles: rechazar o no rechazar una hipótesis. La hipótesis en sí es una conjetura sobre el estado de la naturaleza que prevalece, y normalmente puede formularse como una afirmación de que se presenta un estado específico. Por ejemplo, suponga que se tiene una moneda cargada, para la cual la probabilidad de cara es  $p_1 = 1/4$  o  $p_2 = 3/4$ . Lo que se contrasta es  $p_1$  y  $p_2$ . La afirmación: “se cumple  $p_1$ ”, es una hipótesis, como lo es la afirmación: “en una larga secuencia de lanzamientos de monedas, el número de colas dominará aproximadamente 3 : 1”. La segunda afirmación es equivalente a la primera, en el sentido de que las colas dominarán si y solo si  $p_1$  se cumple.

A  $H_0$  se le llama hipótesis nula. La terminología deriva de una aplicación frecuente en la que  $H_0$  afirma que un tratamiento particular no tiene ningún efecto. Suponga, por ejemplo, que alguien propone un tratamiento para la moneda sesgada, que se conoce que  $p = 1/4$ . El tratamiento pretende eliminar el sesgo. La hipótesis nula,  $H_0$ , afirma que el tratamiento en realidad no tiene ningún efecto. Es decir, la probabilidad permanece  $p = 1/4$ . Una prueba, que involucra una muestra de lanzamientos de la moneda tratada, aceptará o rechazará la hipótesis nula. Dependiendo de otras

### 3.1. Prueba de hipótesis

---

afirmaciones hechas por el proveedor del tratamiento, la alternativa puede ser simple ( $p = 1/2$ ) o compuesta ( $p \neq 1/4$ ). Es decir, si el tratamiento tiene algún efecto, ¿funciona como se anuncia para entregar una moneda imparcial, o simplemente cambia el sesgo existente?

Note que  $H_0$  es un subconjunto de un todo.  $H_1$  es el complemento, en el sentido de que  $H_0 \cup H_1$  abarca (todo el espacio paramétrico) todas las opciones posibles según lo determine la aplicación. También es tradicional suponer pérdida nula cuando se toma la decisión correcta. Es decir, no se tiene ninguna pérdida al aceptar  $H_0$  si  $\theta \in H_0$ , y no se tiene ninguna pérdida al no aceptar  $H_0$  si  $\theta \in H_1$ . Para cada conjunto que puede tomar  $\theta$ , se puede tener una función,  $c(\theta)$ , que es el costo de tomar la decisión equivocada cuando  $\theta$  prevalece como el estado actual. La función de pérdida  $L(\text{acepta/no acepta}, \theta)$  es

$$L(\text{acepta}, \theta) = \begin{cases} 0, & \theta \in H_0 \\ c(\theta), & \theta \in H_1 \end{cases}$$

$$L(\text{no acepta}, \theta) = \begin{cases} c(\theta), & \theta \in H_0 \\ 0, & \theta \in H_1 \end{cases}$$

#### 3.1.1. Lema de Neyman-Pearson

Considérese probar la hipótesis

$$H_0 : \theta = \theta_0 \quad \text{contra} \quad H_1 : \theta = \theta_1$$

donde la función de distribución de probabilidad o función masa de probabilidad correspondiente a  $\theta_i$  es  $f(x|\theta_i)$ ,  $i = 0, 1$ , usando una prueba con región de rechazo  $R$  que satisface

$$\begin{cases} x \in R & \text{si } f(x|\theta_1) > kf(x|\theta_0) \\ x \in R^c & \text{si } f(x|\theta_1) < kf(x|\theta_0) \end{cases} \quad (3.1)$$

para todo  $k \geq 0$ , y

$$\alpha = P_{\theta_0}(X \in R). \quad (3.2)$$

Entonces

- (Suficiente) Cualquier prueba que satisfaga (3.1) y (3.2) es una prueba mas potente de nivel  $\alpha$ .
- (Necesariamente) Si existe una prueba que satisface (3.1) y (3.2) con  $k > 0$ , entonces toda prueba mas potente de nivel  $\alpha$  es una prueba de tamaño  $\alpha$  y toda prueba uniformemente mas potente de nivel  $\alpha$  satisface (3.1), excepto en un conjunto  $A$  que satisface  $P_{\theta_0}(x \in A) = P_{\theta_1}(X \in A) = 0$ .

### 3.1. Prueba de hipótesis

---

**Definición 3.1 *P-value*.** El valor  $p$  o el nivel de significancia observado de una prueba estadística es el valor más pequeño de  $\alpha$  para el cual  $H_0$  puede rechazarse. Es el riesgo real de en lugar de cometer el error tipo I, si  $H_0$  se rechaza en función del valor observado de la estadística de las pruebas. El valor  $p$  mide la fuerza de la evidencia contra  $H_0$ .

#### 3.1.2. Tipos de hipótesis

Suponga que  $X_1, \dots, X_n$  forma una muestra aleatoria de una distribución para la cual la función masa o función de densidad es  $f(x|\theta)$ , donde el valor del parámetro  $\theta$  debe estar en el espacio paramétrico  $\Omega$ ;  $\Omega_0$  y  $\Omega_1$  son conjuntos disjuntos con  $\Omega_0 \cup \Omega_1 = \Omega$ ; y se desea probar el siguiente contraste de hipótesis:

$$H_0 : \theta \in \Omega_0 \quad \text{contra} \quad H_1 : \theta \in \Omega_1,$$

Para  $i = 0$  o  $i = 1$ , el conjunto  $\Omega_i$  puede contener solo un valor de  $\theta$  o podría ser un conjunto más grande.

**Definición 3.2 *Hipótesis simple*:** Una hipótesis  $H : \theta \in \omega$  se llama simple si  $\omega$  contiene un punto único. Así, si  $\omega_1$  consta del punto  $\theta_1$  y si  $\omega_2$  es el punto  $\theta_2$ , el problema se denomina probar una hipótesis simple contra una alternativa simple.

**Definición 3.3 *Hipótesis compuesta*:** Una hipótesis  $H : \theta \in \omega$  se llama compuesta si  $\omega$  contiene más de un valor de  $\theta$ . Por ejemplo,  $H_0 : \theta \in \Omega$  se llama hipótesis compuesta si  $\omega$  contiene más de un elemento, similarmente  $H_1 : \theta \in \Omega - \omega$  es una hipótesis compuesta si  $\Omega - \omega$  contiene más de un elemento.

Bajo una hipótesis simple, la distribución de las observaciones está completamente especificada. Bajo una hipótesis alternativa, se especifica solo que la distribución de las observaciones pertenecen a una cierta clase.

La hipótesis alternativa puede ser unilateral o bilateral.

- **Bilateral:** Una hipótesis alternativa bilateral (también conocida como hipótesis no direccional) se utiliza para determinar si el parámetro de población es mayor que o menor que el valor hipotético. Una prueba bilateral puede detectar cuándo el parámetro de población difiere en cualquier dirección, pero tiene menos potencia que una prueba unilateral.
- **Unilateral:** Una hipótesis alternativa unilateral (también conocida como hipótesis direccional) se utiliza para determinar si el parámetro de población difiere

### 3.1. Prueba de hipótesis

---

del valor hipotético en una dirección específica. Se puede especificar la dirección para que sea mayor que o menor que el valor hipotético. Una prueba unilateral tiene mayor potencia que una prueba bilateral, pero no puede detectar si el parámetro de población difiere en la dirección opuesta.

#### Simple contra simple

El estudio de los métodos para determinar la mejor prueba inicia para el caso en que ambas hipótesis, nula y alterna, son simples. Anteriormente se describió que es una hipótesis simple, así el caso de que  $\Omega = \theta_0, \theta_1$  y se tenga el contraste de hipótesis

$$H_0 : \theta = \theta_0 \quad \text{contra} \quad H_1 : \theta = \theta_1.$$

Se dice que se tiene un juego de hipótesis simple contra simple. Luego, si  $\phi$  es una función de prueba, entonces las probabilidades de los errores están dadas de la siguiente forma:

$$\begin{aligned} P(\text{Error tipo I, usando } \phi) &= P(\phi(X) = 1 | \theta = \theta_0) = \alpha \\ P(\text{Error tipo II, usando } \phi) &= P(\phi(X) = 0 | \theta = \theta_1) = \beta \end{aligned}$$

**Ejemplo 3.1 Simple contra simple.** *Suponga que  $X_1, \dots, X_n$  son i.i.d. (independientes e idénticamente distribuidos) con distribución común  $P$ . El problema es probar la hipótesis nula simple  $P = P_0$  contra la alternativa simple  $P = P_1$ . Sea  $p_i$  la densidad de  $P_i$  con respecto a una medida  $\mu$ . En el lema Neyman-Pearson, la prueba óptima rechaza para valores grandes de  $\sum_{i=1}^n \log[p_1(X_i)/p_0(X_i)]$ . La distribución nula exacta de esta estadística de prueba puede ser difícil de obtener dado que, en general, se requiere una integración de orden  $n$ . Por otro lado, dado que la estadística toma la forma simple de una suma de variables independientes e idénticamente distribuidas, grandes aproximaciones de muestra al valor crítico y potencia se obtienen fácilmente del teorema del límite central.*

#### Simple contra compuesta

Hasta ahora se ha tratado el caso más sencillo de contraste: dos hipótesis simples. En la práctica, las situaciones realmente interesantes comportan por lo menos una hipótesis compuesta. Uno de los contrastes de hipótesis más habituales consiste en:

$$H_0 : \theta = \theta_0 \quad \text{contra} \quad H_1 : \theta \neq \theta_0$$

es decir, la hipótesis alternativa es la negación de la hipótesis nula. Se conoce a dicho contraste como el de la alternativa bilateral.

## 3.1. Prueba de hipótesis

---

### 3.1.3. Errores de decisión

Cuando se usa el Lema de Neyman-Pearson para probar la hipótesis  $H_0$ , hay cuatro resultados posibles basados en si la hipótesis  $H_0$  es realmente verdadera (sin efecto) o no (efecto real) para la población (ver figura 3.1). Al rechazo de la hipótesis nula  $H_0$  generalmente se le denomina como resultado significativo e implica que la hipótesis alternativa  $H_1$  es verdadera. Claramente, una de las dos hipótesis resultan en la correcta decisión estadística; se rechaza correctamente la hipótesis nula  $H_0$  cuando esta es falsa o no se rechaza correctamente la hipótesis nula  $H_0$  cuando esta es verdadera. Pero cuando no se toma la decisión correcta se comete un error.

**Definición 3.4 Error tipo I:** *Es cuando se rechaza por error la hipótesis nula  $H_0$  siendo esta correcta (por ejemplo, cuando se concluye de una muestra y una prueba  $t$  que el parámetro de población no es igual a cero cuando, de hecho, el parámetro de población es igual a cero) y se denota con la letra griega  $\alpha$ . El error Tipo I solo puede ocurrir cuando  $H_0$  es verdadera o no es falsa.*

**Definición 3.5 Error tipo II:** *Es cuando se acepta erróneamente la hipótesis nula  $H_0$  cuando esta es falsa (por ejemplo, cuando se concluye de una muestra y una prueba  $t$  que el parámetro de población es igual a cero cuando, de hecho, el parámetro de población es diferente de cero). El error tipo II se indica con la letra griega  $\beta$  y solo puede ocurrir cuando  $H_0$  es falsa.*

Ambos errores son el resultado de un riesgo. Una muestra aleatoria puede proporcionar información engañosa sobre la población, especialmente si los tamaños de muestra son pequeños. Por ejemplo, dos poblaciones pueden tener el mismo valor medio pero la muestra de una población puede, por azar, contener todos los valores grandes y la muestra de la otra población puede, por azar, contener todos los valores pequeños, lo que resulta en una diferencia estadísticamente significativa entre las medias de las poblaciones.

Note que el error tipo I es posible incluso si  $H_0 : \mu_1 = \mu_2$  es verdadero, aunque es poco probable. Las probabilidades de error tipo I y tipo II no se aplican necesariamente a una prueba estadística específica, pero representan la probabilidad de errores a largo plazo si se muestrean repetidamente de la misma población y se repite la prueba muchas veces.

### 3.1. Prueba de hipótesis

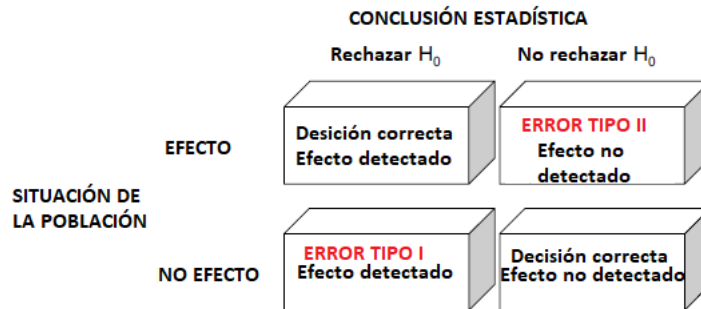


Figura 3.1: Decisión estadística y errores al probar la hipótesis nula.

La figura 3.2, muestra la probabilidad de la distribución muestral  $t$  (de Student) cuando  $H_0$  es verdadera (curva izquierda) y la probabilidad de la distribución muestral  $t$  cuando la hipótesis alternativa  $H_1$  es verdadera (curva derecha). Por supuesto, nunca se sabe cómo se ve esta última distribución en la práctica porque si la hipótesis nula  $H_0$  es falsa, no se conoce cuál es la verdadera hipótesis alternativa  $H_1$ .

Para un caso particular, habrá una distribución diferente para cada posible  $H_1$ , pero solo una distribución muestral para  $H_0$ . Suponga que el valor crítico de la distribución  $t$  para  $\alpha = 0.05$  es indicado. Entonces si la hipótesis nula  $H_0$  es verdadera, cualquier valor de  $t$  mayor al valor crítico conlleva a un rechazo de  $H_0$  y un error tipo I. Si  $H_0$  es falsa y  $H_1$  es verdadera, cualquier valor igual o menor que el valor crítico conducirá a no rechazar  $H_0$  y un error tipo II. Note que si por ejemplo  $H_0$  indica igualdad entre medias, entonces  $H_1$  indicará diferencia entre medias (Quinn y Keough, 2002).

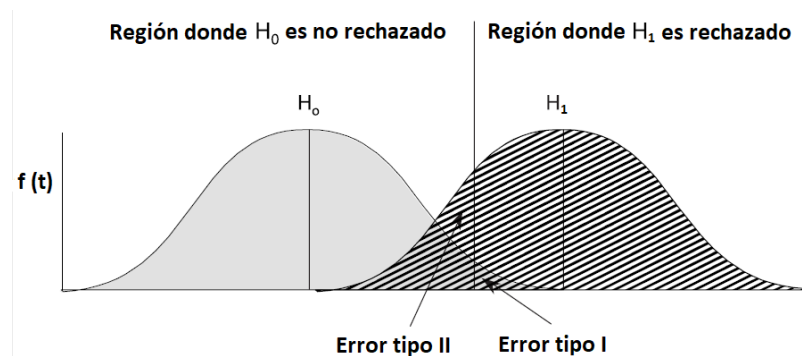


Figura 3.2: Representación gráfica de las probabilidades de error tipo I y tipo II, usando una prueba  $t$  como ejemplo.

## 3.1. Prueba de hipótesis

---

### 3.1.4. Función potencia

**Definición 3.6 *Función potencia:*** Es la probabilidad de que la hipótesis nula sea rechazada cuando la hipótesis alternativa es verdadera (es decir, la probabilidad de no cometer un error de tipo II). La función potencia de una prueba de hipótesis con región de rechazo  $R$  es la función de  $\theta$  definida por  $\beta(\theta) = P(X \in R)$ .

La función potencia ideal es 0 para todo  $\theta \in \Omega_0$  y 1 para todo  $\theta \in \Omega_1$ . Excepto en situaciones triviales, este ideal no puede ser alcanzado. Cualitativamente, una buena prueba tiene una función de potencia cercana a 1 para  $\theta \in \Omega_1$  y cerca de 0 para  $\theta \in \Omega_0$ .

**Ejemplo 3.2** Sea  $X \sim \text{Binomial}(5, \theta)$ . Considere probar la hipótesis

$$H_0 : \theta \leq \frac{1}{2} \quad \text{contra} \quad H_1 : \theta > \frac{1}{2}$$

considere primero la prueba que rechaza  $H_0$  si y solo si se observan todos los éxitos. La función de potencia para esta prueba es

$$\beta_1(\theta) = P(X \in R) = P(X = 5) = \theta^5.$$

La gráfica de  $\beta_1(\theta)$  se muestra en la figura 3.3. Al examinar esta función potencia, se puede decidir que, aunque la probabilidad del error tipo I es aceptablemente baja ( $\beta_1(\theta) \leq (\frac{1}{2})^5 = 0.0312$ ) para todo  $\theta \leq \frac{1}{2}$ , la probabilidad del error tipo II es alta ( $\beta_1(\theta)$  es pequeña) para la mayoría  $\theta > \frac{1}{2}$ . La probabilidad del error tipo II es menor que  $\frac{1}{2}$  solo si  $\theta > (\frac{1}{2})^5 = 0.87$ . Para lograr probabilidades de error tipo II más pequeñas, se podría considerar usar la prueba que rechaza la hipótesis nula  $H_0$  si  $X = 3, 4, \text{ o } 5$ . La función poder para esta prueba es

$$\beta_2(\theta) = P(X = 3, 4, \text{ o } 5) = \binom{5}{3}\theta^3(1-\theta)^2 + \binom{5}{4}\theta^4(1-\theta)^1 + \binom{5}{5}\theta^5(1-\theta)^0.$$

La gráfica de  $\beta_2(\theta)$  también está en la figura 3.3. Se puede ver en la figura 3.3 que la segunda prueba ha logrado una probabilidad de error de tipo II más pequeña,  $\beta_2(\theta)$  es mayor para  $\theta > \frac{1}{2}$ . Pero la probabilidad de error tipo I es mayor para la segunda prueba,  $\beta_2(\theta)$  es mayor para  $\theta \leq \frac{1}{2}$ . Si se debe hacer una elección entre estas dos pruebas, el investigador debe decidir que tipo de error, la descrita por  $\beta_1(\theta)$  o la descrita por  $\beta_2(\theta)$ , es más aceptable (Casella y Berger, 2002).



### 3.1. Prueba de hipótesis

---

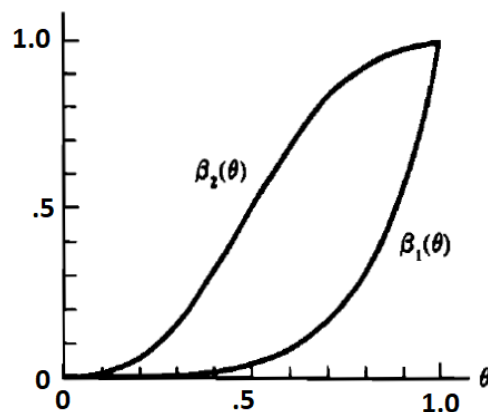


Figura 3.3: Función potencia para ejemplo.

#### 3.1.5. Tamaño de la prueba

Para que una prueba de hipótesis sea buena, se debe diseñar de una forma que minimice los errores de decisión. Esto no es tan sencillo como puede parecer, puesto que para un tamaño de muestra fijo, generalmente es imposible hacer que ambos tipos de probabilidades de error sean pequeños, como ya se ha comentado anteriormente el disminuir un tipo de error va generalmente acompañado por un incremento en el otro tipo de error.

En la práctica, un tipo de error puede tener más importancia que el otro. Una forma de reducir al mismo tiempo ambos tipos de errores es incrementar el tamaño de la muestra, lo cual no siempre puede ser posible. Así en la búsqueda de una buena prueba, es común restringir la consideración a las pruebas que controlan la probabilidad de error tipo I en un nivel específico. Dentro de esta clase de pruebas, se buscan pruebas que tengan una probabilidad de error de tipo II lo más pequeña posible. Los siguientes dos términos son útiles cuando se discuten pruebas que controlan probabilidades de error tipo I.

**Definición 3.7** Para  $0 \leq \alpha \leq 1$ , una prueba con la función potencia  $\beta(\theta)$  es una prueba de tamaño  $\alpha$  si  $\sup_{\theta \in \Omega_0} \beta(\theta) = \alpha$ .

**Definición 3.8** Para  $0 \leq \alpha \leq 1$ , una prueba con la función potencia  $\beta(\theta)$  es una prueba de nivel  $\alpha$  si  $\sup_{\theta \in \Omega_0} \beta(\theta) \leq \alpha$ .

Algunos investigadores no hacen distinción entre los términos tamaño y nivel de la prueba, y algunas veces estos términos se usan indistintamente. Pero de acuerdo con

### 3.2. Pruebas para probar distribución Poisson contra distribución Poisson inflada con ceros

---

definiciones mencionadas, el conjunto de pruebas de nivel  $\alpha$  contiene el conjunto de pruebas de tamaño  $\alpha$ . Además, la distinción se vuelve importante en modelos complicados y situaciones de prueba complicadas, donde a menudo es computacionalmente imposible construir una prueba de tamaño  $\alpha$ .

## 3.2. Pruebas para probar distribución Poisson contra distribución Poisson inflada con ceros

Como se ha hecho referencia la distribución Poisson se usa ampliamente en los datos relacionados con el conteo, como en estudios de calidad en la industria, en la genética y en otras áreas de la ciencia. En ocasiones ocurre que en muchos casos se comete el error de estimar el modelo bajo la distribución Poisson cuando en las frecuencias de los datos se encuentran demasiados ceros. Una situación que se está volviendo más y más común es la ocurrencia de un gran número de ceros en conteos.

Se ha observado que cuando hay una gran cantidad de datos de recuento cero, la distribución Poisson no cuenta bien los datos en muchos casos, y se prefieren algunos modelos alternativos (Campbell *et al.* (1991); Freund *et al.* (1999); Dankmar (1998); Miaou (1994)). De hecho, el modelo Poisson a menudo subestima la dispersión observada, y el resultado es que las estimaciones calculadas son indebidamente obtenidas. Por lo tanto, se debe investigar la extensión de la distribución Poisson a una que pueda usarse para modelar un mayor efecto de dispersión.

En la literatura existen varias pruebas para probar hipótesis del tipo simple contra compuesto, específicamente, para el caso en estudio la hipótesis que se va a desarrollar es la siguiente;

$$H_0 : \text{Distribución Poisson} \quad \text{contra} \quad H_1 : \text{Distribución ZIP},$$

es decir,

$$H_0 : \omega = 0 \quad \text{contra} \quad H_1 : \omega \neq 0, \quad \omega \in (0, 1).$$

Cuando no se rechaza la hipótesis nula, no es necesario utilizar el modelo Poisson inflado con ceros, ya que es más fácil trabajar con la distribución Poisson. En esta sección se presentan varias pruebas encontradas en la literatura para realizar una comparación con la prueba propuesta.

Para obtener la función potencia de cada prueba se siguió el siguiente procedimiento:

1. Generar una muestra de tamaño  $n = 10, 20, 50$ .

## 3.2. Pruebas para probar distribución Poisson contra distribución Poisson inflada con ceros

---

2. Calcular el estadístico de prueba y su p-value.
3. Realizar  $nSIM = 1000$  simulaciones de paso 1 y 2.
4. Se cuenta el número de rechazos, comparar el p-value con respecto al tamaño nominal de la prueba  $\alpha = 0.05$ .

### 3.2.1. Estadístico de prueba Score

Jan Van den Broek en 1995, propuso el estadístico de Score, el cual presenta una prueba de puntaje para probar si el número de ceros es demasiado grande para que una distribución de Poisson se ajuste bien a los datos.

$$S_1 = \frac{(n_0 - np_0)^2}{np_0(1 - p_0) - n\bar{y}p_0^2},$$

donde  $n_0$  es el número de ceros en la muestra,  $n$  es el tamaño de muestra,  $\bar{y}$  es la media de la muestra y  $p_0 = e^{-\hat{\mu}}$ , en el cual  $\hat{\mu}$  es la estimación del parámetro de la distribución Poisson bajo la hipótesis nula. El estadístico de Score asintóticamente se distribuye como una distribución Ji-cuadrada con una grado de libertad bajo la hipótesis nula (den Broek, 1995).

En la tabla 3.1 se muestran las potencias para  $\lambda = 10$ ,  $\alpha = 0.05$ , diferentes tamaños de muestra y diferentes valores de  $\omega$ , se realizaron 1000 simulaciones.

**Tabla 3.1:** Tabla de potencia para el estadístico de prueba Score.

N	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0$
10	0.998	0.993	0.966	0.881	0.001
20	1.000	0.999	0.999	0.992	0.000
50	1.000	1.000	1.000	1.000	0.003
100	1.000	1.000	1.000	1.000	.003
300	1.000	1.000	1.000	1.000	.015
500	1.000	1.000	1.000	1.000	.054
700	1.000	1.000	1.000	1.000	.053

### 3.2.2. Estadístico de prueba Rao-Chakravarti

Rao-Chakravarti en 1956, propuso el estadístico conocido como  $R$ , el cual se muestra en la ecuación 3.3, el estadístico se distribuye asintóticamente bajo la hipótesis nula como una distribución normal estándar.

### 3.2. Pruebas para probar distribución Poisson contra distribución Poisson inflada con ceros

---

$$R = \frac{n_0 - n((n-1)/n)^{n\bar{y}}}{\{n((n-1)/n)^{n\bar{y}} - n^2((n-1)/n)^{2n\bar{y}} + n(n-1)((n-2)/n)^{n\bar{y}}\}^{1/2}} \quad (3.3)$$

donde  $n_0$  es el número de ceros en la muestra,  $n$  es el tamaño de muestra y  $\bar{y}$  es la media de la muestra (Rao y Chakravarti, 1956).

En la tabla 3.2, se muestran las potencias para el estadístico  $R$ , se realizaron 1000 simulaciones, para diferentes tamaños de muestra,  $\lambda = 10$ ,  $\alpha = 0.05$  y diferentes valores de  $\omega$ .

**Tabla 3.2:** Tabla de potencia para el estadístico de prueba R.

N	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0$
10	0.998	0.989	0.972	0.897	0.001
20	1.000	1.000	0.998	0.989	0.002
50	1.000	1.000	1.000	1.000	0.001
100	1.000	1.000	1.000	1.000	.028
500	1.000	1.000	1.000	1.000	.04

#### 3.2.3. Estadístico de prueba Cochran

Cochran en 1954, propuso el estadístico que se muestra en la ecuación 3.4, bajo la hipótesis nula el estadístico se distribuye como una distribución normal estándar (Cochran, 1954).

$$C = \frac{n_0 - ne^{-\bar{y}}}{[ne^{-\bar{y}}(1 - e^{-\bar{y}} - \bar{y}e^{-\bar{y}})]^{1/2}} \quad (3.4)$$

En la tabla 3.3, se muestran las potencias para el estadístico Cochran para varios tamaños de muestra,  $\lambda = 10$ ,  $\alpha = 0.05$ .

**Tabla 3.3:** Tabla de potencia para el estadístico de prueba Cochran.

N	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0$
10	0.999	0.994	0.972	0.896	0.01
20	1.000	1.000	1.000	0.988	0.01
50	1.000	1.000	1.000	1.000	0.02
100	1.000	1.000	1.000	1.000	.062
500	1.000	1.000	1.000	1.000	.057
700	1.000	1.000	1.000	1.000	.052

### 3.3. Inferencia bayesiana

---

#### 3.2.4. Estadístico de prueba de razón de verosimilitud generalizada

Usando notación de el-Shaarawi, en 1985 se propuso el estadístico de razón de verosimilitud generalizada, el cual puede ser representado como en la ecuación 3.5 (el Shaarawi, 1985),

$$-2 \log \Delta = 2 \left( n_0 \ln \left( \frac{n_0}{n} \right) + (n - n_0) \left( \ln \left( \frac{\bar{y}}{\hat{\lambda}_2} \right) - \hat{\lambda}_2 \right) + n\bar{y}(\ln \hat{\lambda}_2 + 1 - \ln \bar{y}) \right), \quad (3.5)$$

donde  $\Delta = \frac{L(0, \hat{\lambda}_0)}{L(\hat{\omega}_1, \hat{\lambda}_1)}$ .

En la tabla 3.4 se muestran las potencias para el estadístico de razón de verosimilitud generalizada para diversos tamaños de muestra,  $\lambda = 10$ ,  $\alpha = 0.05$ .

**Tabla 3.4:** Tabla de potencia para el estadístico de prueba razón de verosimilitud generalizada.

N	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0$
10	0.998	0.994	0.976	0.896	0.01
20	1.000	1.000	0.999	0.992	0.01
50	1.000	1.000	1.000	1.000	0.02
100	1.000	1.000	1.000	1.000	.017
500	1.000	1.000	1.000	1.000	.022
700	1.000	1.000	1.000	1.000	.051

### 3.3. Inferencia bayesiana

La inferencia bayesiana es el proceso de ajustar un modelo a un conjunto de datos y resumir el resultado mediante una distribución de probabilidad sobre los parámetros del modelo y sobre cantidades no observadas tales como predicciones para nuevas observaciones.

Mientras que las dificultades relacionadas con los métodos de máxima verosimilitud son principalmente problemas de optimización (modos múltiples, solución de ecuaciones de verosimilitud, enlaces entre ecuaciones de probabilidad y modos globales, etc.), el enfoque bayesiano resulta más a menudo en problemas de integración.

En el paradigma bayesiano, la información presentada por los datos  $x$ , una realización de  $X \sim f(x|\theta)$ , se combina con información previa que se especifica en una

### 3.3. Inferencia bayesiana

distribución a priori con densidad  $\pi(\theta)$  y se resume en una distribución de probabilidad,  $\pi(\theta|x)$ , llamada distribución posterior. Se deriva de la distribución conjunta  $f(x|\theta)\pi(\theta)$ , según la fórmula de Bayes

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)}$$

donde  $m(x) = \int f(x|\theta)\pi(\theta)d\theta$  es la densidad marginal de  $X$ . Para la estimación de un parámetro en particular  $h(\theta)$  el enfoque teórico de decisión para la inferencia estadística requiere la especificación de la función de pérdida  $L(\delta, \theta)$  que representa la pérdida incurrida al estimar  $h(\theta)$  con  $\delta$ . La versión bayesiana de este enfoque conduce a la minimización del riesgo de Bayes,

$$\int \int L(\delta, \theta) f(x|\theta) \pi(\theta) dx d\theta.$$

*Ejemplo 3.1* Asuma  $X \sim N(\theta, \sigma^2)$ , donde  $\theta$  es desconocida y  $\sigma$  conocida. Sea  $\pi(\theta)$  una densidad normal con media  $\mu$  y varianza  $\tau^2$ , donde  $\mu$  y  $\tau$  son conocidas. Entonces

$$f(x|\theta)\pi(\theta) = (2\pi\sigma\tau)^{-1} \exp\left\{-\frac{1}{2}\left[\frac{(\theta - \mu)^2}{\tau^2} + \frac{(x - \theta)^2}{\sigma^2}\right]\right\}.$$

Para encontrar  $m(x)$ , se define

$$\rho = \tau^{-2} + \sigma^{-2} = \frac{\tau^2 + \sigma^2}{\tau^2\sigma^2}$$

al completar cuadrados se obtiene,

$$\begin{aligned} \frac{1}{2}\left[\frac{(\theta - \mu)^2}{\tau^2} + \frac{(x - \theta)^2}{\sigma^2}\right] &= \frac{1}{2}\left[\left(\frac{1}{\tau^2} + \frac{1}{\sigma^2}\right)\theta^2 - 2\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\theta + \left(\frac{\mu^2}{\tau^2} + \frac{x^2}{\sigma^2}\right)\right] \\ &= \frac{1}{2}\rho\left[\theta^2 - \frac{2}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\theta\right] + \frac{1}{2}\left(\frac{\mu^2}{\tau^2} + \frac{x^2}{\sigma^2}\right) \\ &= \frac{1}{2}\rho\left[-\frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2 - \frac{1}{2\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)^2 + \frac{1}{2}\left(\frac{\mu^2}{\tau^2} + \frac{x^2}{\sigma^2}\right) \\ &= \frac{1}{2}\rho\left[-\frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2 + \frac{(\mu - x)^2}{2(\sigma^2 + \tau^2)}. \end{aligned}$$

Por lo tanto

$$f(x|\theta)\pi(\theta) = (2\pi\sigma\tau)^{-1} \exp\left\{-\frac{1}{2}\rho\left[-\frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2\right\} \exp\left\{\frac{(\mu - x)^2}{2(\sigma^2 + \tau^2)}\right\}$$

y

$$m(x) = \int_{-\infty}^{\infty} f(x|\theta)\pi(\theta)d\theta = (2\pi\rho)^{-1/2}(\sigma\tau)^{-1} \exp\left\{\frac{(\mu - x)^2}{2(\sigma^2 + \tau^2)}\right\}.$$

### 3.3. Inferencia bayesiana

---

Así,

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)} = \left(\frac{\rho}{2\pi}\right)^{1/2} \exp\left\{-\frac{1}{2}\rho\left[-\frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2\right\}.$$

Note que a partir de las ecuaciones anteriores, la distribución marginal de  $X$  es  $N(\mu, (\sigma^2 + \tau^2))$  y la distribución posterior de  $\theta$  dado  $x$  es  $N(\mu(x), \rho^{-1})$ , donde

$$\mu(x) = \frac{1}{\rho} \left( \frac{\mu}{\tau^2} + \frac{x}{\sigma^2} \right) = \frac{\sigma^2}{\sigma^2 + \tau^2} \mu + \frac{\tau^2}{\sigma^2 + \tau^2} x$$

$$\therefore \mu(x) = x - \frac{\sigma^2}{\sigma^2 + \tau^2} (x - \mu).$$

Como ejemplo concreto, considere la situación en la que un niño realiza una prueba de inteligencia. Suponga que el resultado de la prueba  $X$  es  $N(\theta, 100)$ , donde  $\theta$  es el verdadero nivel del cociente intelectual del niño, medido por la prueba. (En otras palabras, si el niño tomara una gran cantidad de pruebas independientes similares, su puntuación promedio sería aproximadamente  $\theta$ ).

Suponga la población como un todo,  $\theta$  se distribuye como una distribución  $N(100, 225)$ . Usando las ecuaciones anteriores, se deduce que, marginalmente,  $X$  es  $N(100, 325)$ , mientras que la distribución posterior de  $\theta$  dado  $x$  es normal con media

$$\mu(x) = \frac{(100)(100) + x(225)}{(100 + 225)} = \frac{400 + 9x}{13}$$

y varianza

$$\rho^{-1} = \frac{(100)(225)}{(100 + 225)} = 69.23.$$

Por lo tanto, si un niño obtiene un puntaje de 115 en la prueba, su cociente intelectual verdadero  $\theta$  tiene una distribución posterior  $N(110.39, 69.23)$  ([Berger, 1985](#)).

**Ejemplo 3.2** Se debe realizar un análisis de sangre para ayudar a indicar si una persona tiene o no una enfermedad en particular. Si el resultado de la prueba es positivo (se denota  $x = 1$ ) o negativo (denotado  $x = 0$ ).  $\theta_1$  denota que “la enfermedad está presente” y  $\theta_2$  denota el estado “no hay enfermedad presente”, suponga que se sabe que  $f(1|\theta_1) = 0.8$ ,  $f(0|\theta_1) = 0.2$ ,  $f(1|\theta_2) = 0.3$  y  $f(0|\theta_2) = 0.7$ .

Se cuenta con la información a priori,  $\pi(\theta_1) = 0.05$  y  $\pi(\theta_2) = 0.95$ . Entonces

$$m(1) = f(1|\theta_1)\pi(\theta_1) + f(1|\theta_2)\pi(\theta_2) = 0.04 + 0.285 = 0.325,$$

$$m(0) = f(0|\theta_1)\pi(\theta_1) + f(0|\theta_2)\pi(\theta_2) = 0.01 + 0.665 = 0.675,$$

$$\pi(\theta|x=1) = \frac{f(1|\theta)\pi(\theta)}{m(1)} = \begin{cases} \frac{0.04}{0.325} = 0.123 & \text{si } \theta = \theta_1, \\ \frac{0.285}{0.325} = 0.877 & \text{si } \theta = \theta_2 \end{cases}$$

### 3.3. Inferencia bayesiana

---

y

$$\pi(\theta|x=0) = \frac{f(0|\theta)\pi(\theta)}{m(0)} = \begin{cases} \frac{0.01}{0.675} = 0.0148 & \text{si } \theta = \theta_1, \\ \frac{0.665}{0.675} = 0.9852 & \text{si } \theta = \theta_2 \end{cases}.$$

Es interesante observar que, incluso si la prueba de sangre es positiva, todavía hay una probabilidad del 12.3% de que la enfermedad esté presente. Note que  $m(1)$  y  $m(0)$  dan las proporciones generales de pruebas positivas y negativas que pueden anticiparse.

Estos podrían ser útiles para fines logísticos, es decir, la prueba positiva se siguiera con pruebas más elaboradas; el 32.5% de los inicialmente evaluados requeriría las pruebas más elaboradas.

En situaciones discretas, como en el ejemplo anterior, la fórmula  $\pi(\theta|x)$  es comúnmente conocida como el teorema de Bayes, y fue descubierta por Bayes en 1763. La expresión del teorema de Bayes es en términos de eventos disjuntos  $A_1, A_2, \dots, A_n$ , cuya unión tiene probabilidad uno (es decir, uno de los eventos  $A_i$  es seguro que ocurrirá).

Las probabilidades a priori  $P(A_i)$ , para los eventos, son asumidas conocidas. Se produce un evento  $B$ , para el cual  $P(B|A_i)$  (la probabilidad condicional de  $B$  dado  $A_i$ ) es conocida para cada  $A_i$ . El teorema de Bayes dice que ([Berger, 1985](#)),

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}.$$

#### 3.3.1. Muestreador de Gibbs

El muestreador de Gibbs inició con gran popularidad con el artículo de Geman y Geman en 1984, quienes estudiaron modelos de procesamiento de imágenes. El muestreador de Gibbs es una técnica para generar variables aleatorias desde una distribución (marginal) indirectamente sin tener que calcular la densidad.

Mediante el uso de técnicas como el muestreador de Gibbs, es posible evitar cálculos difíciles, sustituyéndolos por una secuencia de cálculos más fáciles, estas metodologías han tenido un amplio impacto en problemas prácticos. Aunque la mayoría de las aplicaciones del muestreador de Gibbs ha sido en modelos bayesianos, es extremadamente útil en cálculos clásicos como la verosimilitud. Además, estas metodologías de cálculo también han tenido un impacto en la teoría. Al liberar a los estadísticos de ocuparse de cálculos complicados, el aspecto estadístico de un problema puede convertirse en el foco principal.

Suponga que se tiene una distribución conjunta  $f(x, y_1, \dots, y_p)$  y se está interesado en



### 3.3. Inferencia bayesiana

---

obtener características de la distribución marginal

$$f(x) = \int \cdots \int f(x, y_1, \dots, y_p) dy_1 dy_2 \cdots dy_p$$

como su media y su varianza. Tal vez, el mas natural y sencillo enfoque seria calcular  $f(x)$  y utilizarlo para obtener la característica deseada. Sin embargo, hay muchos casos donde la solución de la integral es muy difícil obtener, ya sea numéricamente o analíticamente. En estos casos el muestreador de Gibbs proporciona un método alternativo para obtener  $f(x)$ .

Mas bien en lugar de computar o aproximar directamente  $f(x)$ , el muestreador de Gibbs permite generar una muestra  $X_1, \dots, X_m \sim f(x)$  sin requerir  $f(x)$ . Mediante la simulación de una muestra suficientemente grande, se puede calcular la media, la varianza o cualquier otra característica de  $f(x)$  hasta el grado de precisión deseado.

Es importante darse cuenta de que, en efecto, el resultado final de cualquier cálculo, aunque basado en simulaciones, son las cantidades de población. Por ejemplo, para calcular la media de  $f(x)$ , se puede utilizar  $\frac{1}{m} \sum_{i=1}^m X_i$  y el hecho de que

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m X_i = \int_{-\infty}^{\infty} x f(x) dx = E[X].$$

Por lo tanto, al tomar  $m$  suficientemente grande, cualquier característica de la población, incluso la propia densidad, se puede obtener con cualquier grado de exactitud.

Para entender el funcionamiento del muestreador de Gibbs, se explicará en el caso de dos variables. Comenzando con un par de variables aleatorias  $(X, Y)$ , el muestreador de Gibbs genera una muestra de  $f(x)$  por muestreo en lugar de las distribuciones condicionales  $f(x|y)$  y  $f(y|x)$ , distribuciones que se conocen con frecuencia en modelos estadísticos. Esto se hace generando una “secuencia de Gibbs” de variables aleatorias

$$Y'_0, X'_0, Y'_1, X'_1, Y'_2, X'_2, \dots, Y'_k, X'_k. \quad (3.6)$$

El valor inicial  $Y'_0$  es especificado y el resto de la secuencia se obtiene iterativamente generando alternativamente valores de

$$X'_j \sim f(x|Y'_j = y'_j)$$

$$Y'_{j+1} \sim f(y|X'_j = x'_j).$$

Se refiere a esta generación de 3.6 como muestreo de Gibbs. Resulta que bajo condiciones razonablemente generales, la distribución de  $X'_k$  converge a  $f(x)$  (el verdadero valor marginal de  $X$ ) cuando  $k \rightarrow \infty$ . Por tanto, para  $k$  lo suficientemente grande,

### 3.3. Inferencia bayesiana

---

la observación final en 3.6, es decir  $X'_k = x'_k$ , es efectivamente un valor puntual de la muestra de  $f(x)$ . La convergencia (en distribución) de la secuencia de Gibbs 3.6 puede explotarse de diversas maneras para obtener una muestra aproximada de  $f(x)$ . Por ejemplo, Gelfand y Smith (1990) sugieren generar  $m$  secuencias de Gibbs independientes de longitud  $k$ , y luego usar el valor final de  $X'_k$  de cada secuencia. Si  $k$  se elige lo suficientemente grande, esto produce una muestra independiente e idénticamente distribuida de  $f(x)$ . Existen otros métodos para extraer información de la secuencia de Gibbs.

Algunos de los temas más importantes en el muestreo de Gibbs rodean la implementación y comparación de los diversos enfoques para extraer información de la secuencia de Gibbs. Estos temas son actualmente un tema de mucho debate e investigación.

Como se ha mencionado, la muestra de Gibbs genera una cadena de variables aleatorias de Markov que convergen a la distribución de interés  $f(x)$ . Muchos de los enfoques populares para extraer información de la secuencia de Gibbs explotan esta propiedad seleccionando un valor grande para  $k$ , y luego tratar cualquier  $X_j$  para  $j > k$  como una muestra de  $f(x)$ . El problema entonces se convierte en elegir el valor apropiado de  $k$ .

Una estrategia general para elegir  $k$  es observar la convergencia de algún aspecto de la secuencia de Gibbs. Por ejemplo, Gelfand y Smith (1990) y Gelfand *et al.* (1990) sugieren estudiar las estimaciones de densidad de  $m$  secuencias independientes de Gibbs, y elegir  $k$  como el primer punto en el que estas densidades parecen ser las mismas bajo una “prueba de rotulado (felt-tip pen)”. Tanner (1991) sugiere analizar una secuencia de pesos que mide la discrepancia entre la distribución muestreada y la deseada. Geweke (en prensa) sugiere un análisis basado en consideraciones de series temporales. Desafortunadamente, tales enfoques no son infalibles, ilustrados por Gelman y Rubin (1991). Una alternativa puede ser elegir  $k$  sobre la base de consideraciones teóricas, como en Raftery y Banfield (1991). M. T. Wells sugirió una conexión (un enlace) entre la selección de  $k$  y el parámetro (cooling) de enfriamiento en la simulación (annealing).

Una alternativa natural para el muestreo del  $k$ -ésimo o valor final de muchas repeticiones independientes de la secuencia de Gibbs, como se mencionó en la sección del muestreador de Gibbs, es generar una larga secuencia de Gibbs y luego extraer cada  $t$ -ésima observación. Para  $k$  suficientemente grande, esto también arrojaría una muestra independiente e idénticamente distribuida aproximada de  $f(x)$ . Una ventaja de este enfoque es que disminuye la dependencia de los valores iniciales. Una posible desventaja es que la secuencia de Gibbs puede permanecer en un pequeño subconjunto del espacio de muestra durante un tiempo prolongado (Gelman y Rubin, 1991).

Para problemas grandes, computacionalmente costosos, un enfoque menos derrochador para explotar la secuencia de Gibbs es usar todas las realizaciones de  $X'_j$ , para  $j \leq k$ , como en George y McCulloch (1993). Aunque los datos resultantes serán de-

### 3.3. Inferencia bayesiana

---

pendientes, todavía será el caso de que la distribución empírica de  $X'_j$ , converja a  $f(x)$ . Tenga en cuenta que desde este punto de vista se puede ver que la “eficiencia de la muestra de Gibbs ” está determinada por la velocidad de esta convergencia. Intuitivamente, esta tasa de convergencia será más rápida cuando  $X'_j$  se mueva rápidamente a través del espacio de la muestra, una característica que puede considerarse mezcla.

Las variaciones en estos y otros enfoques para explotar la secuencia de Gibbs han sido sugeridas por [Gelman y Rubin \(1991\)](#), [Müller \(1991\)](#), [Ritter y Tanner \(1990\)](#) y [Tierney \(1991\)](#).

# Capítulo 4

## DISTRIBUCIÓN POISSON INFLADA CON CEROS

La distribución Poisson (la distribución lleva el nombre de Siméon Poisson) es una distribución de probabilidad discreta para los recuentos de eventos que ocurren aleatoriamente en un intervalo de tiempo (o espacio) dado.

La distribución Poisson es extensamente usada en datos relacionados a conteos, pero como se ha mencionado que existen muestras donde se presenta exceso de ceros, por lo que se ha propuesto en la literatura usar el modelo Poisson inflado con ceros.

La distribución Poisson inflada con ceros (ZIP) se dio a conocer en 1992 por Lambert. Los ceros que se presentan en el modelo ZIP se clasifican en dos tipos: ceros estructurales, los que ocurren con probabilidad  $\omega$  y ceros muestrales, los que ocurren con probabilidad  $(1 - \omega)$  y son de una distribución Poisson con media  $\lambda$ .

Este proceso de dos estados proporciona una distribución simple de la mezcla de dos componentes con función de densidad de probabilidad, la cual es dada por:

$$P[Y = y] = \begin{cases} \omega + (1 - \omega)e^{-\lambda}, & y = 0, \\ (1 - \omega)\frac{e^{-\lambda}\lambda^y}{y!}, & y = 1, 2, \dots, \quad 0 \leq \omega \leq 1, \end{cases}$$

con media

$$E[Y] = (1 - \omega)\lambda = \mu$$

y varianza

$$Var[Y] = \mu + \left(\frac{\omega}{1 - \omega}\right)\mu^2.$$

#### 4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros

---

Dicha distribución se denota por  $Y \sim ZIP(\lambda, \omega)$ , la distribución marginal de  $Y$  exhibe sobredispersión, si  $\omega > 0$ . Además se observa que si  $\omega = 0$  se tiene la distribución Poisson. El modelo ZIP corresponde cuando  $\omega$  se encuentra entre 0 y 1, y podría llamarse Poisson deflada con ceros si  $(1 - \exp^{-\lambda})^{-1} \leq \omega < 1$ . La distribución ZIP fue descrita por Singh en 1963, él consideró un modelo biológico en la que la categoría “no progenie” (clase de ceros) combinó las categorías “estéril” y “fertil con no progenie”. Este ejemplo ilustra que la distribución ZIP puede ser considerada como una mezcla de una distribución degenerada por ceros y una distribución Poisson.

Para observaciones  $y_1, \dots, y_n$  la función log-verosimilitud es,

$$l(\lambda, \omega, y) = \sum_{i=1}^n \{I_{(y=0)} \log[\omega + (1 - \omega)e^{-\lambda}] + I_{(y=1,2,\dots)} [\log(1 - \omega) - \lambda + y\{\log\lambda - \log(y!)\}]\}.$$

Muchas pruebas estadísticas se han desarrollado para probar la hipótesis nula de que los datos pertenecen a una distribución Poisson. Sin embargo, la distribución ZIP fue mencionada explícitamente por primera vez en la literatura en 1963, resulta que ya existían pruebas para probar Poisson contra ZIP. Una prueba fue propuesta por Cochran en 1954 la cual fue mas tarde probada ser equivalente a la prueba de Score para probar Poisson contra ZIP. Otra prueba fue dada por [Rao y Chakravarti \(1956\)](#). Recientemente [Ridout et al. \(2001\)](#) propusieron una prueba para probar ZIP contra la distribución binomial negativa inflada con ceros, [Thas y Rayner \(2005\)](#) presentaron mas generalmente pruebas para probar la distribución ZIP contra una gran clase de alternativas.

#### 4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros

En las figuras [4.1](#), [4.2](#), [4.3](#) y [4.4](#) se muestran la función de densidad y acumulada de la distribución Poisson inflada con ceros con algunos valores de los parámetros  $\omega$  y  $\lambda$ . Como se ha mencionado, la distribución ZIP puede ser representada como una mezcla de una distribución degenerada por ceros y una distribución Poisson.

En la figura [4.1](#) se presenta el caso para  $\omega = 0.3$ , note que en  $y = 0$  se muestra la probabilidad de tener la frecuencia de ceros de la parte degenerada, mientras que el parámetro  $\lambda = 6$ ,  $\lambda = 15$  y  $\lambda = 30$  concentra la otra parte de la probabilidad, con media en el valor del parámetro  $\lambda$ .

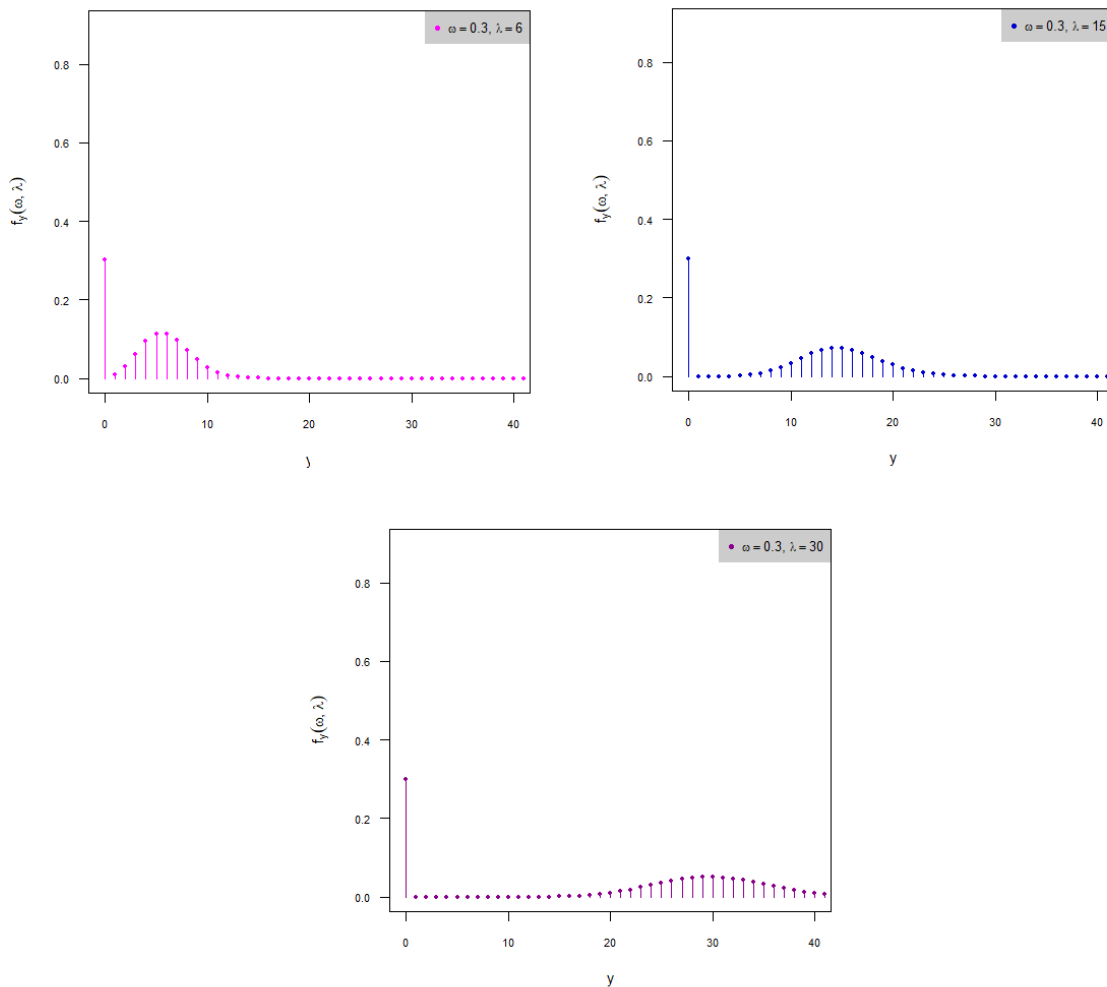
En la figura [4.2](#) se nota que la función de distribución acumulada converge mas rápido

#### 4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros

a 1 cuando el valor de  $\lambda$  es más pequeño.

La situación descrita anteriormente se repite sin importar la combinación de parámetros de  $\omega$  y  $\lambda$ , por ejemplo, en la figura 4.3 se presenta el caso para  $\omega = 0.7$  y valores de  $\lambda = 6$ ,  $\lambda = 15$  y  $\lambda = 30$ ; en  $y = 0$  se tiene la probabilidad de tener la frecuencia de ceros de la parte degenerada mientras que en los valores de  $\lambda$  se concentra la otra parte de la probabilidad con media en el valor del parámetro.

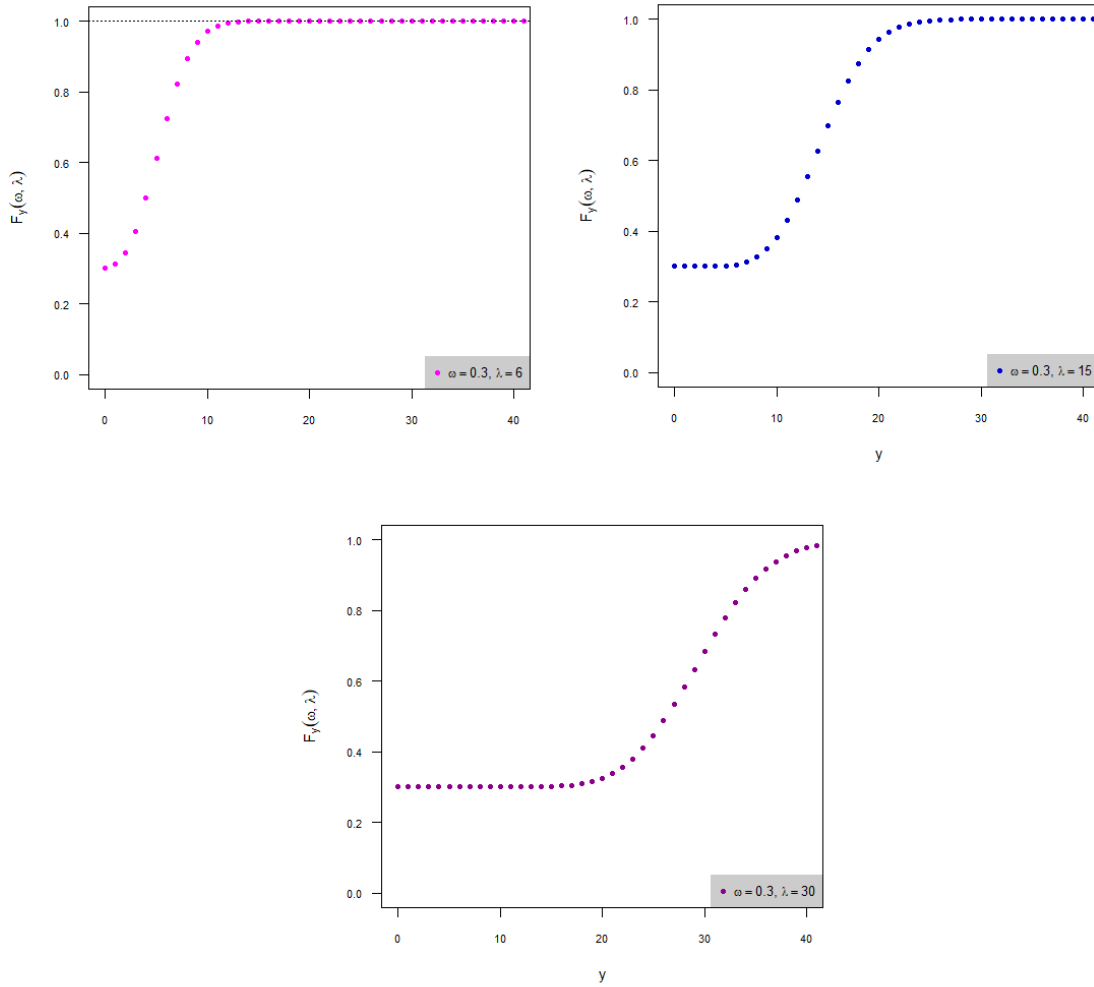
De la misma forma, la función de distribución acumulada converge más rápido a 1 cuando el valor de  $\lambda$  es más pequeño, como se muestra en la figura 4.4.



**Figura 4.1:** Función de densidad con  $\omega = 0.3$ .

#### 4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros

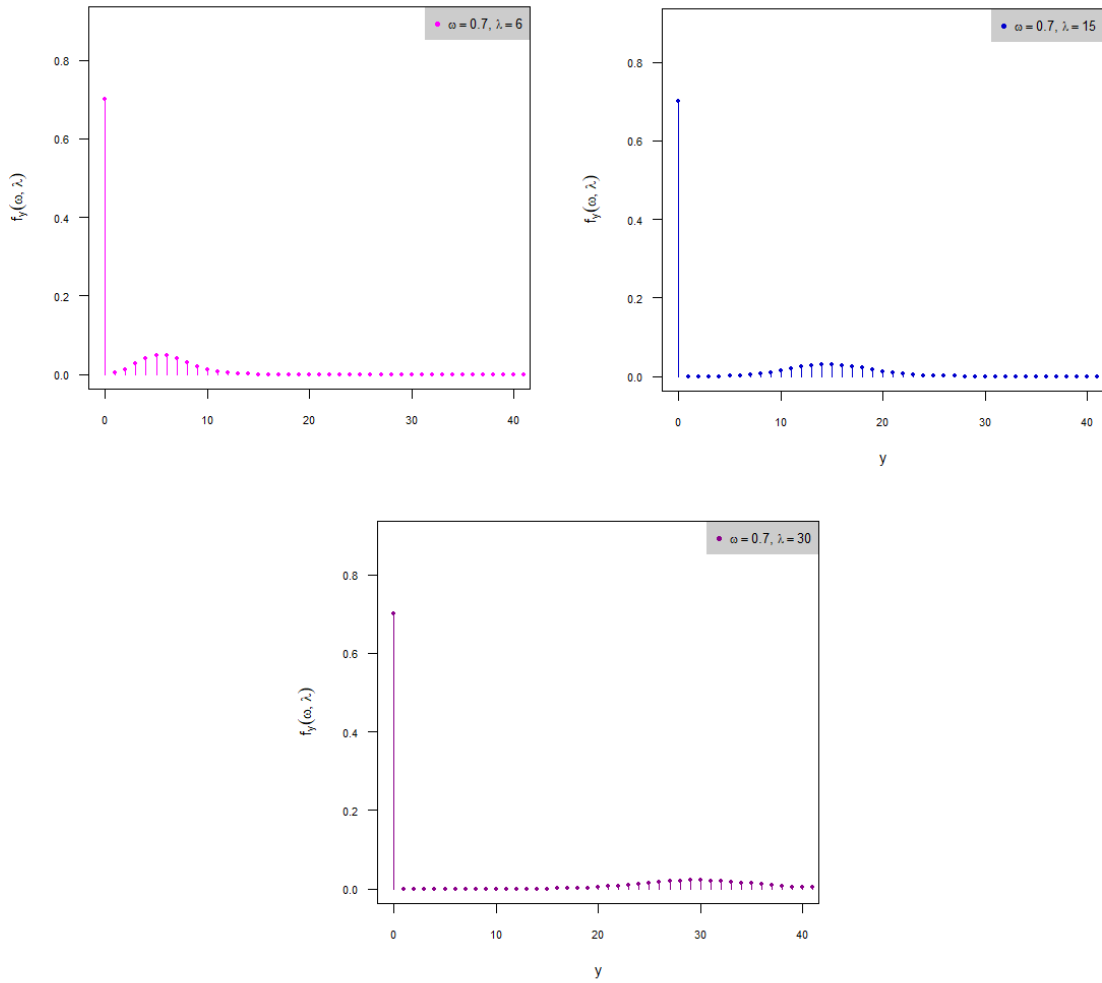
---



**Figura 4.2:** Función acumulada con  $\omega = 0.3$ .

#### 4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros

---

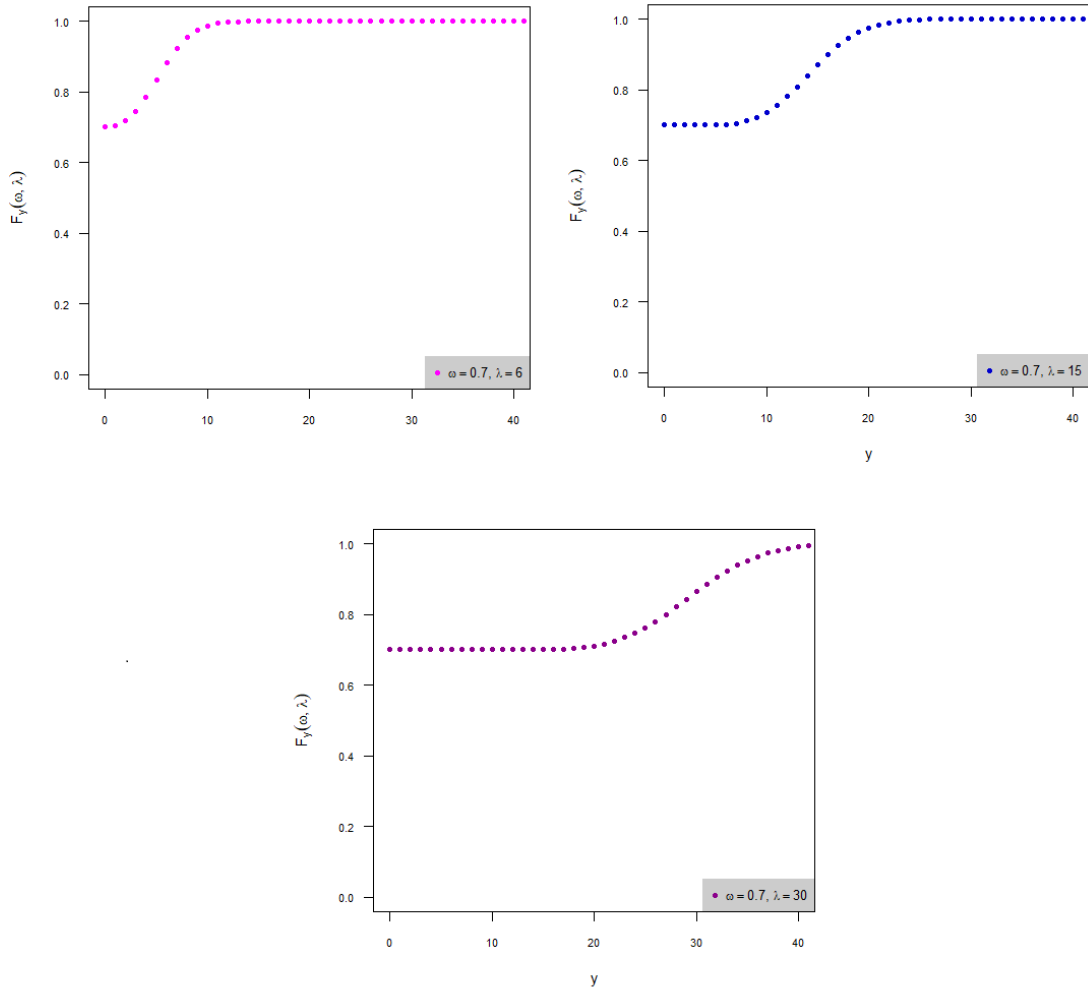


**Figura 4.3:** Función de densidad con  $\omega = 0.7$ .



#### 4.1. Función de densidad y función acumulada de la distribución Poisson inflada con ceros

---



**Figura 4.4:** Función acumulada con  $\omega = 0.7$ .

## 4.2. Otras propiedades de la distribución Poisson inflada con ceros

En esta sección se muestra que la función masa de probabilidad del modelo ZIP esta bien definida, lo cual se prueba a continuación;

$$\begin{aligned}
 \sum_{y=0}^{\infty} f_Y(y) &= 1 \\
 \sum_{y=0}^{\infty} f_Y(y) &= \sum_{y=0}^0 \omega + (1 - \omega)e^{-\lambda} + \sum_{y=1}^{\infty} (1 - \omega)e^{-\lambda} \frac{\lambda^y}{y!} \\
 &= \omega + (1 - \omega)e^{-\lambda} + (1 - \omega)e^{-\lambda} \sum_{y=1}^{\infty} \frac{\lambda^y}{y!} \\
 &= \omega + (1 - \omega)e^{-\lambda} + (1 - \omega)e^{-\lambda}(e^{\lambda} - 1) \\
 &= \omega + e^{-\lambda} - e^{-\lambda}\omega + (e^{-\lambda} - e^{-\lambda}\omega)(e^{\lambda} - 1) \\
 &= \omega + e^{-\lambda} - e^{-\lambda}\omega + e^{-\lambda}e^{\lambda} - e^{-\lambda} - e^{-\lambda}e^{\lambda}\omega + e^{-\lambda}\omega \\
 &= 1.
 \end{aligned}$$

Ahora se presentan los estimadores de momentos.

Para obtener el primer momento, se siguieron los pasos siguientes:

$$\frac{\sum_{i=1}^n y_i}{n} = E[Y] = (1 - \omega)\lambda \tag{4.1}$$

de ecuación 4.1 se despeja  $\lambda$

$$\lambda = \frac{\sum y_i}{n(1 - \omega)}.$$

Para obtener el segundo momento, se siguieron los pasos siguientes:

$$\frac{\sum_{i=1}^n y_i^2}{n} = E[Y^2] = (1 - \omega)\lambda + \omega(1 - \omega)^2\lambda^2 + (1 - \omega)^2\lambda^2 \tag{4.2}$$

## 4.2. Otras propiedades de la distribución Poisson inflada con ceros

---

ahora se sustituye el despeje de  $\lambda$  en ecuación 4.2 y se despeja  $\omega$

$$\omega = \frac{\frac{\sum_{i=1}^n y_i^2}{n} - \bar{y} - \bar{y}^2}{\bar{y}^2} \Bigg/ 1 + \frac{\sum_{i=1}^n y_i^2}{n} - \bar{y} - \bar{y}^2$$

# Capítulo 5

## PRUEBA PROPUESTA

En este trabajo se estudia el comportamiento de un intervalo de credibilidad e intervalo de confianza para el modelo Poisson inflado con ceros para probar hipótesis simple contra compuesta, donde se utiliza la representación de la distribución Poisson inflada con ceros como lo sugiere [Ghosh \*et al.\* \(2006\)](#).

### 5.1. Distribución Poisson inflada con ceros

Como se ha mencionado a distribución Poisson inflada con ceros (ZIP) se dio a conocer en 1992 por Lambert, con función de densidad de probabilidad dada por:

$$P[Y = y] = \begin{cases} \omega + (1 - \omega)e^{-\lambda}, & y = 0, \\ (1 - \omega) \frac{e^{-\lambda} \lambda^y}{y!}, & y = 1, 2, \dots, \quad 0 \leq \omega \leq 1, \end{cases}$$

La cual pertenece a una familia de distribuciones llamada modelos inflados con ceros, los que pueden ser representados como  $Y = V(1 - B)$ , donde  $B$  es una distribución Bernoulli con parámetro  $\omega$  y la variable aleatoria  $V$  independiente a  $B$  puede ser una distribución Poisson con parámetro  $\lambda$ , una binomial negativa con parámetros  $(\lambda, r)$  o mas generalmente una serie de potencia con parámetro  $\lambda$ ,

$$E(Y) = (1 - \omega)E(V)$$
$$Var(Y) = \frac{\omega}{1 - \omega} [E(Y)]^2 + \delta E[Y]$$

donde  $\delta = \frac{Var[V]}{E[V]}$ .

## 5.2. Función de verosimilitud

---

Para el caso de interés, se representa la distribución Poisson inflada con ceros  $Y$  como una distribución Poisson por una distribución Bernoulli, como se muestra a continuación:

$$Y = V(1 - B)$$

donde  $V \sim \text{Poisson}(\lambda)$  y  $B \sim \text{Bernoulli}(\omega)$ , donde la variable  $V$  es independiente de  $B$ .

## 5.2. Función de verosimilitud

Para observaciones  $y_1, y_2, \dots, y_n$ , la función de verosimilitud puede ser representada de la siguiente forma, basada en el método aumentación de datos

$$L(y|\omega, \lambda) \propto \omega^S (1 - \omega)^{n-S} \lambda^{\sum_{y_i \neq 0} y_i} e^{-(n-S)\lambda}$$

donde  $S \sim \text{Binomial}(n, p(\lambda, \omega))$ ,  $p(\lambda, \omega) = \frac{\omega}{\omega + (1-\omega)e^{-\lambda}}$ .

## 5.3. Inferencia Bayesiana

Los modelos Bayesianos generalmente se refieren a inferencias en un conjunto de parámetros  $\theta = (\theta_1, \dots, \theta_d)$ , de dimensión  $d$ , que incluye cantidades inciertas, ya sean efectos fijos y aleatorios, parámetros jerárquicos, variables indicadoras no observadas y datos faltantes [Gelman y Rubin \(1992\)](#). El conocimiento previo sobre los parámetros se resume mediante la densidad a priori  $p(\theta)$ , la verosimilitud  $p(y|\theta)$ , y el conocimiento actualizado está contenido en la densidad posterior  $p(\theta|y)$ . El teorema de Bayes se indica en la siguiente ecuación

$$p(\theta|y) = p(y|\theta)p(\theta)/p(y),$$

donde el denominador en el lado derecho es la probabilidad marginal  $p(y)$ . Esta última es una integral sobre todos los valores de  $\theta$  del producto  $p(y|\theta)p(\theta)$  y puede considerarse como una constante de normalización para asegurar que  $p(\theta|y)$  sea una densidad adecuada. Esto significa que uno puede expresar el teorema de Bayes como

$$p(\theta|y) \propto p(y|\theta)p(\theta).$$

La distribución posterior  $f(\omega, \lambda|y)$  se obtiene de la siguiente manera:

$$f(\omega, \lambda|y) = \frac{L(y|\omega, \lambda)\pi(\omega)\pi(\lambda)}{\int_0^\infty \int_0^1 L(y|\omega, \lambda)\pi(\omega)\pi(\lambda)\partial\omega\partial\lambda}$$

### 5.3. Inferencia Bayesiana

---

$$\propto L(y|\omega, \lambda)\pi(\omega)\pi(\lambda),$$

donde se asume que las distribuciones a priori de  $\omega$  y  $\lambda$  son independientes y se distribuyen como se muestra a continuación,

$$\omega \sim \text{Beta}(\alpha, \beta)$$

$$\lambda \sim \text{Gama}(k, \theta).$$

Para el trabajo en cuestión la distribución posterior  $f(\omega, \lambda|y)$  se obtiene de la siguiente manera:

$$f(\omega, \lambda|y) \propto L(y|\omega, \lambda)\pi(\omega)\pi(\lambda),$$

$$f(\omega, \lambda|y) \propto \omega^S(1-\omega)^{n-S} \lambda^{\sum_{y_i \neq A} y_i} e^{-(n-S)\lambda} * \frac{\omega^{\alpha-1}(1-\omega)^{\beta-1}}{B(\alpha, \beta)} * \theta e^{-\theta\lambda} \frac{(\theta\lambda)^{k-1}}{\Gamma(k)}.$$

Dado que la distribución posterior no tiene forma cerrada, se encuentran las distribuciones condicionales para implementar el muestreador de Gibbs. La distribución condicional para el parámetro  $\lambda$  se obtiene de la forma siguiente:

$$f(\lambda|\omega, y) \propto \omega^S(1-\omega)^{n-S} \lambda^{\sum_{y_i \neq A} y_i} e^{-(n-S)\lambda} * \frac{\omega^{\alpha-1}(1-\omega)^{\beta-1}}{B(\alpha, \beta)} * \theta e^{-\theta\lambda} \frac{(\theta\lambda)^{k-1}}{\Gamma(k)}$$

$$\propto \lambda^{\sum_{y_i \neq A} y_i} e^{-(n-S)\lambda} e^{-\theta\lambda} \lambda^{k-1}$$

$$\propto \lambda^{\sum_{y_i \neq A} y_i + k - 1} e^{-\lambda(\theta + n - S)}.$$

De la última igualdad se puede deducir que la condicional de  $\lambda$ , tiene distribución Gama. Es decir,

$$\lambda|\omega, y \sim \text{Gama}\left(\sum_{y_i \neq A} y_i + k - 1, n + \theta - S\right).$$

Ahora, la distribución condicional para el parámetro  $\omega$  se obtiene como se muestra:

$$f(\omega|\lambda, y) \propto \omega^S(1-\omega)^{n-S} \lambda^{\sum_{y_i \neq A} y_i} e^{-(n-S)\lambda} * \frac{\omega^{\alpha-1}(1-\omega)^{\beta-1}}{B(\alpha, \beta)} * \theta e^{-\theta\lambda} \frac{(\theta\lambda)^{k-1}}{\Gamma(k)}$$

$$f(\omega|\lambda, y) \propto \omega^S(1-\omega)^{n-S} \omega^{\alpha-1}(1-\omega)^{\beta-1}$$

$$f(\omega|\lambda, y) \propto \omega^{\alpha+S-1}(1-\omega)^{n+\beta-S-1}$$

de donde se puede asumir que la condicional de  $\omega$ , tiene distribución *Beta*. Es decir,

$$\omega|\lambda, y \sim \text{Beta}(\alpha + S, n + \beta - S).$$

### 5.3. Inferencia Bayesiana

---

Con anterioridad se estableció la prueba de hipótesis de interés:

$$H_0 : \omega = 0 \quad \text{contra} \quad H_1 : \omega \in (0, 1)$$

en la que como se ha mencionado, se espera rechazar  $H_0$ .

La prueba está basada en el intervalo de credibilidad del parámetro  $\omega$  y se rechaza la hipótesis nula si el límite superior es mayor al cuantil 95 %, es decir, se rechaza  $H_0$  cuando límite superior  $\geq \hat{Q}(0.95)$ .

Para estudiar el comportamiento de la prueba del intervalo de credibilidad se recurrió a técnicas de simulación Monte Carlo utilizando el muestreador de Gibbs para diferentes tamaños de muestra y diferentes valores del parámetro  $\omega$ . Una vez obtenida la distribución posterior de  $\omega$ , se generaron 1000 muestras a posteriori del parámetro de interés, a cada muestra se obtuvo el intervalo de credibilidad usando la función **hdi** del paquete **HDInterval** del software estadístico **R**. Finalmente, con la regla de decisión “Rechazar  $H_0$  si el límite superior  $\geq \hat{Q}(0.95)$ ”, se cuenta el número de veces que se rechaza la hipótesis nula bajo la hipótesis nula y se cuenta las veces que se rechaza  $H_0$  bajo la hipótesis alternativa para obtener el tamaño de la prueba y la potencia de la prueba, respectivamente.

Se utilizó el algoritmo Gibbs para obtener la distribución posterior de  $\omega$ .

#### Algoritmo Gibbs para encontrar distribución del parámetro $\omega$

1. Iniciar con valores  $\omega^t$  y  $\lambda^t$ .
2. Generar  $\omega_1^{t+1} \sim f(\omega_1|x, \lambda^t, \omega^t)$ .
3. Generar  $\omega_2^{t+1} \sim f(\omega_2|x, \lambda^t, \omega^{t+1})$ .
4.  $\vdots$
5. Generar  $\omega_k^{t+1} \sim f(\omega_k|x, \lambda^t, \omega^{t+1}, \dots, \omega^{k-1})$ .
6. Guardo  $\omega_k^{t+1}$ .
7. Repetir  $m$  veces paso 2, 3, 4, 5 y 6. Así se obtiene muestra posterior para parámetro omega de tamaño  $m$ .

Una vez obtenido la distribución posterior de  $\omega$ , es posible realizar inferencia en ella, así, con la ayuda del paquete **coda** del software **R** se crearon 1000 muestras de la distribución posterior y la función **hdi** proporcionó los intervalos de credibilidad para cada muestra.

# Capítulo 6

## ESTUDIO DE PRUEBA PROPUESTA

### 6.1. Diagnóstico de convergencia del algoritmo Gibbs

Las cadenas múltiples casi siempre son necesarias para evaluar la convergencia de un algoritmo, ahora se considera la herramienta de convergencia central de [Gelman y Rubin \(1992\)](#), implementada en el software [R](#), haciendo uso de la librería `coda` con el código `gelman.diag` y `gelman.plot`.

[Gelman y Rubin \(1992\)](#) proponen un enfoque general para monitorear la convergencia del rendimiento de métodos Monte Carlo usando cadenas de Markov en el que  $m > 1$  cadenas paralelas se ejecutan con valores iniciales que están sobredispersados en relación con la distribución posterior. La convergencia se diagnostica cuando las cadenas han “olvidado” sus valores iniciales, y la salida de todas las cadenas es indistinguible. El diagnóstico de `gelman.diag` se aplica a una sola variable de la cadena. Se basa en una comparación de varianzas dentro de la cadena y entre cadenas, y es similar a un análisis de varianza clásico.

Usando  $M$  cadenas paralelas  $\{x_m^{(t)}\}_t (1 \leq m \leq M, 1 \leq t \leq T)$ , posiblemente transformado en  $\xi_m^{(t)} = h(x_m^{(t)})$ , la regla de detención de [Gelman y Rubin \(1992\)](#) se basa en la diferencia entre un estimador ponderado de la varianza y la varianza de los estimadores de las diferentes cadenas.

La varianza entre cadenas se define como la varianza de los promedios

$$B_T = \frac{1}{M-1} \sum_{m=1}^M (\bar{\xi}_m - \bar{\xi})^2,$$



## 6.1. Diagnóstico de convergencia del algoritmo Gibbs

mientras que la varianza dentro de la cadena es el promedio de las varianzas

$$W_T = \frac{1}{M-1} \sum_{m=1}^M s_m^2 = \frac{1}{M-1} \sum_{m=1}^M \frac{1}{T-1} \sum_{t=1}^T (\xi_m^{(t)} - \bar{\xi}_m)^2$$

con

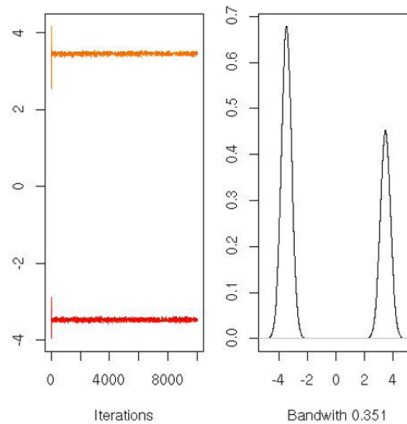
$$\bar{\xi}_m = \frac{1}{T} \sum_{t=1}^T \xi_m^{(t)}, \quad \bar{\xi} = \frac{1}{M} \sum_{m=1}^M \bar{\xi}_m.$$

Esas cantidades se usan comúnmente en análisis de varianza. Un primer estimador de la varianza posterior de  $\xi$  es

$$\hat{\sigma}_T^2 = \frac{T-1}{T} W_T + B_T.$$

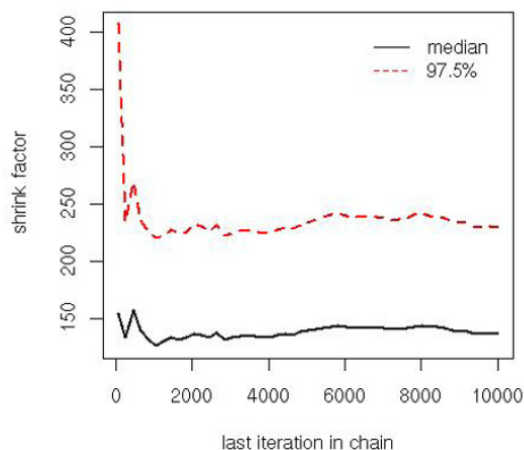
Las estimaciones de la varianza  $\hat{\sigma}_T^2$  y  $W_T$  son asintóticamente equivalentes, [Gelman y Rubin \(1992\)](#) construyeron su criterio basado en la comparación de estas dos cantidades. Mientras las diferentes secuencias  $(\xi_m^{(t)})$  permanezcan concentradas alrededor de sus valores iniciales (o en diferentes porciones del espacio de parámetros),  $\hat{\sigma}_T^2$  sobrestima la varianza de  $(\xi_m^{(t)})$  debido a la gran dispersión de la distribución inicial, mientras que  $W_T$  subestima esta varianza.

Por ejemplo, en la figura [6.1](#) y [6.2](#), se compara las estimaciones de cinco cadenas de Markov, estas gráficas son las proporcionadas por `gelman.plot`. El criterio (“factor de reducción”) está muy lejos de 1, para nada sorprendente que las cadenas aún no han convergido. Tenga en cuenta que el cuantil en el gráfico se basa en una distribución  $t$ , que equivale a suponer que la distribución objetivo está cerca de lo normal. Si claramente este no es el caso, la cadena debe transformarse primero.



**Figura 6.1:** Ejemplo de diagnóstico de Gelman y Rubin.

## 6.1. Diagnóstico de convergencia del algoritmo Gibbs



**Figura 6.2:** Ejemplo de diagnóstico de Gelman y Rubin.

El factor de reducción (shrink factor) de [Gelman y Rubin \(1992\)](#) es  $R_T$ , con

$$R_T^2 = \frac{\hat{\sigma}_T^2 + \frac{B_T}{M} \nu_T + 1}{W_T \nu_T + 3}$$

donde  $\nu_T = 2(\hat{\sigma}_T^2 + \frac{B_T}{M})^2/W_T$  es la estimación de grados de libertad. Mientras la distribución aproximada de  $R_T^2$  puede derivarse de una aproximación  $F$ ,  $TB_T/W_T \sim F(M-1, 2W_T^2/\omega_T)$  con

$$\omega_T = \frac{1}{M^2} \left[ \sum_{m=1}^M s_m^4 - \frac{1}{M} \left( \sum_{m=1}^M s_m^2 \right)^2 \right],$$

`gelman.diag` y `gelman.plot` solo reportan el valor del factor de reducción (shrink factor), junto con el límite superior de credibilidad del 97.5%.

Para el desarrollo del estudio, se trabajó con métodos de simulación para obtener cadenas de Markov, 2 precisamente, y a dichas cadenas se les realizó el diagnóstico de Gelman y Rubin sobre los parámetros  $\omega$  y  $\lambda$  de la distribución Poisson inflada con ceros, los cuales se muestran en las tablas [6.1](#) y [6.2](#), respectivamente. También se muestran los gráficos de Gelman y Rubin, donde se puede observar en que iteración aproximada se alcanza la convergencia de la cadena de Markov.

En la figura [6.3](#) es posible notar que el parámetro  $\omega$  converge a la distribución estacionaria a partir de la iteración 4000 aproximadamente, de tal forma que si se toman 10000 iteraciones es seguro que el último valor converja a la distribución posterior de  $\omega$ . La muestra posterior de omega se obtiene usando el método de ([Gelfand y Smith, 1990](#)). En la tabla [6.1](#) se presenta la estimación puntual del diagnóstico de Gelman y

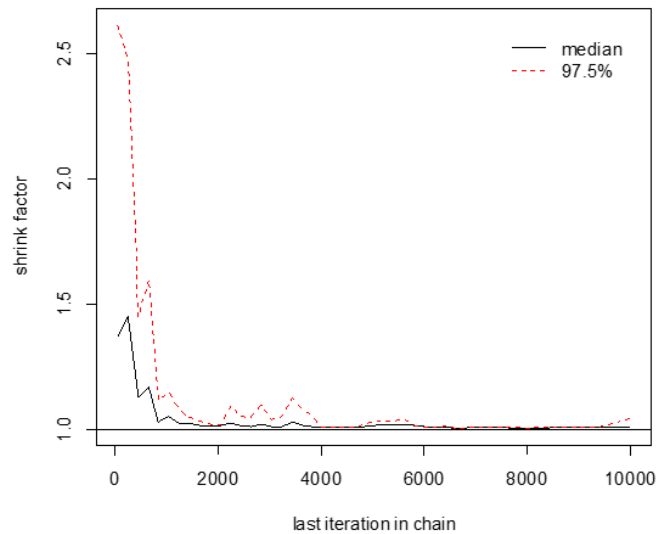
## 6.1. Diagnóstico de convergencia del algoritmo Gibbs

---

Rubin, en este caso es 1 e indica la convergencia de la cadena.

**Tabla 6.1:** Diagnóstico de Gelman para el parámetro  $\omega$ .

Estimación puntual	1.01
Límite de confianza superior	1.04



**Figura 6.3:** Diagnóstico de Gelman y Rubin del parámetro  $\omega$ .

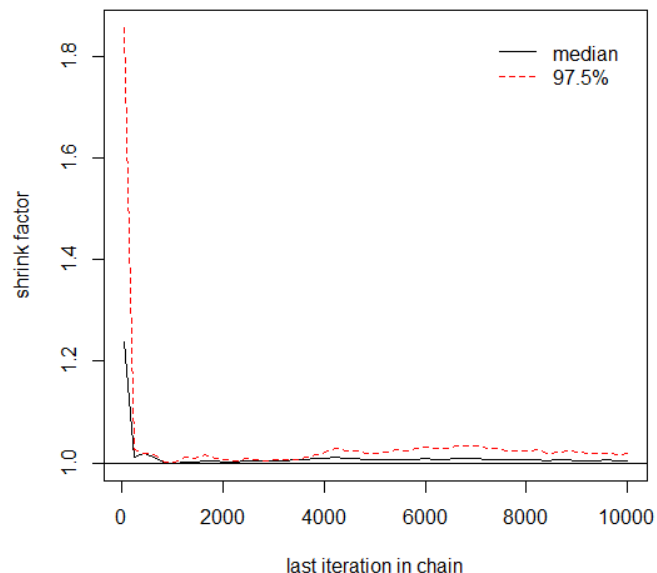
La gráfica 6.4 indica que el parámetro  $\lambda$  aproximadamente a partir de la iteración 2000 converge a la distribución estacionaria, es decir, el factor de reducción se aproxima al valor de referencia uno, así al realizar 10000 iteraciones es seguro que el último valor converge a la distribución posterior de  $\lambda$ . También se presenta en la tabla 6.2, donde se muestra la estimación puntual del diagnóstico de Gelman y Rubin, el cual es 1, lo que implica que la cadena converge.

**Tabla 6.2:** Diagnóstico de Gelman para el parámetro  $\lambda$ .

Estimación puntual	1.01
Límite de confianza superior	1.02

## 6.1. Diagnóstico de convergencia del algoritmo Gibbs

---



**Figura 6.4:** Diagnóstico de Gelman y Rubin del parámetro  $\lambda$ .

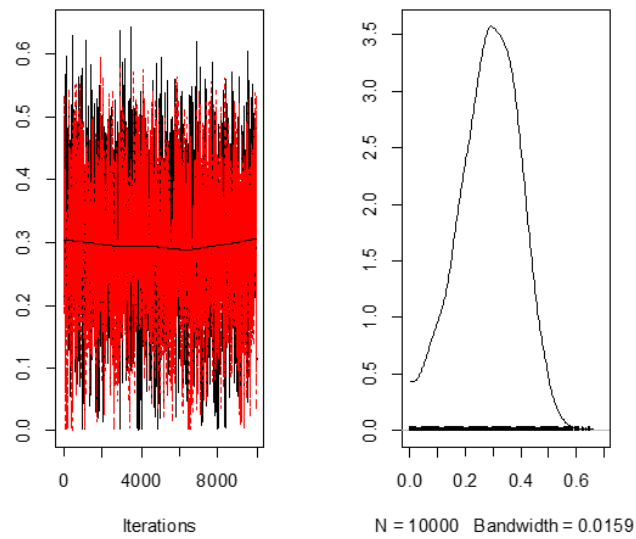
Existen varias posibilidades de realizar el diagnóstico de convergencia de las cadenas de Markov, en la literatura se sugiere examinar si la cadena está explorando bien el espacio, donde se grafican los valores simulados de  $\omega$  y  $\lambda$  frente a el tiempo y se analiza si se producen desviaciones de la estacionariedad.

La convergencia se obtiene cuando los valores están dentro del espacio paramétrico sin grandes periodicidades y especialmente sin tendencias o paralelismos entre las cadenas. Es decir, cuando, por ejemplo el estadístico Gelman y Rubin indica convergencia, la distribución posterior es única para las cadenas generadas con los valores dispersos aleatoriamente por todo el espacio muestral y no se pueden distinguir las diferentes cadenas.

El gráfico 6.5 muestra la traza del parámetro  $\omega$ , donde se observa que la cadena presenta una buena exploración del espacio paramétrico, dado que se encuentra entre  $(0, 1)$  y mas específicamente cerca del valor simulado  $\omega = 0.4$ .

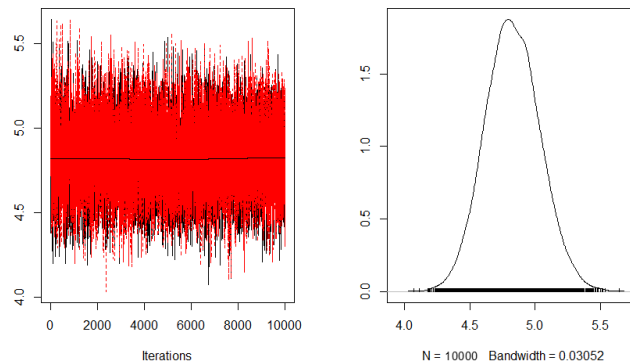
## 6.2. Autocorrelación

---



**Figura 6.5:** Cadena de Markov del parámetro  $\omega$ .

De manera similar para las secuencias del parámetro  $\lambda$ , mostradas en el gráfico 6.6, ya que las cadenas son similares y se encuentra cerca del valor simulado  $\lambda = 5$ , no presentan grandes periodicidades ni tendencias o paralelismos entre ellas.



**Figura 6.6:** Cadena de Markov del parámetro  $\lambda$ .

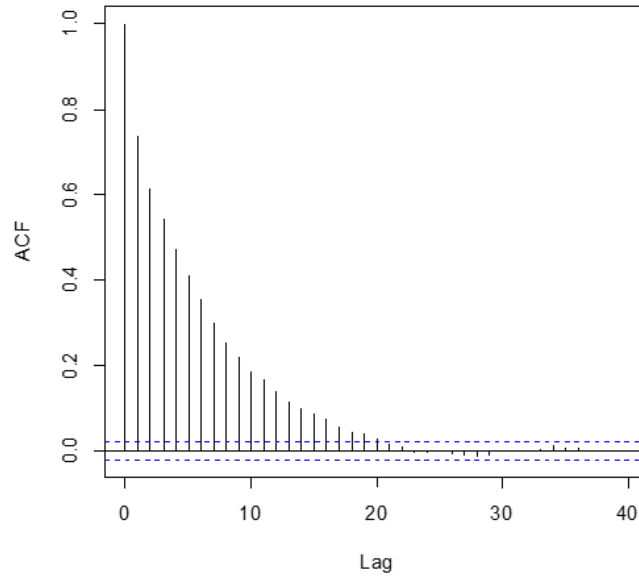
## 6.2. Autocorrelación

Es conveniente hacer un gráfico de la función de autocorrelación estimada. Bajos valores de autocorrelación indican rápida convergencia y valores elevados indican

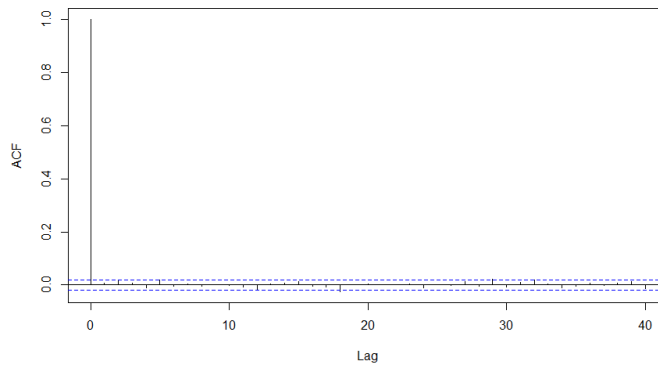
## 6.2. Autocorrelación

---

convergencia débil. En las figuras 6.7 y 6.8 se muestran las correlaciones para los parámetros  $\omega$  y  $\lambda$ , respectivamente, donde se observa que la convergencia es rápida, así se prueba que las observaciones de la muestra de la distribución posterior para los parámetros en cuestión son independientes.



**Figura 6.7:** Función de autocorrelación para el parámetro  $\omega$ .



**Figura 6.8:** Función de autocorrelación para el parámetro  $\lambda$ .

## 6.3. Función posterior

En esta sección se muestra la densidad de la función posterior de  $\omega$  y  $\lambda$  para diferentes tamaños de muestra y diferentes valores de los parámetros.

La gráfica 6.9 muestra la densidad posterior del parámetro  $\omega$  y para  $\lambda = 5$ , note que a medida que el tamaño de muestra crece, se tiene menos variabilidad y la masa de probabilidad se concentra alrededor del parámetro  $\omega$ . Además, conforme  $\omega$  se aleja de la hipótesis nula, la distribución posterior no contiene a  $\omega = 0$ .

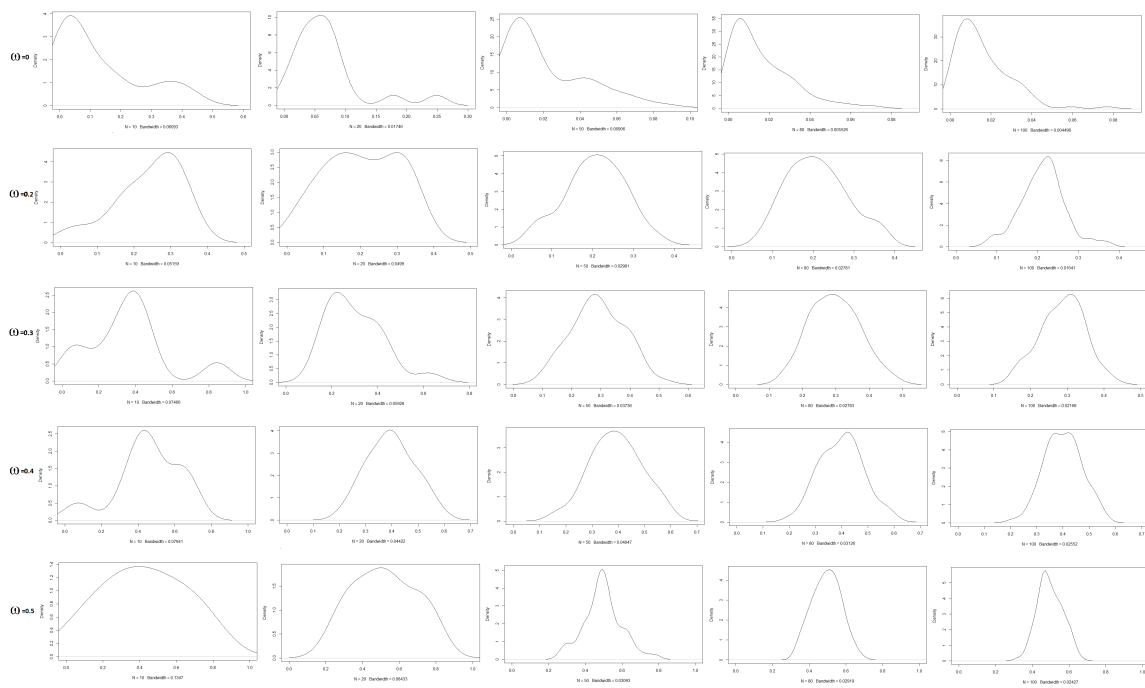


Figura 6.9: Función posterior del parámetro  $\omega$  para  $\lambda = 5$ .

### 6.3. Función posterior

La gráfica 6.10 muestra la densidad posterior del parámetro  $\omega$  y para  $\lambda = 10$ , note que a medida que el tamaño de muestra crece, se tiene menos variabilidad y la masa de probabilidad se concentra alrededor del parámetro  $\omega$ . Además, conforme  $\omega$  se aleja de la hipótesis nula, la distribución posterior no contiene a  $\omega = 0$ .

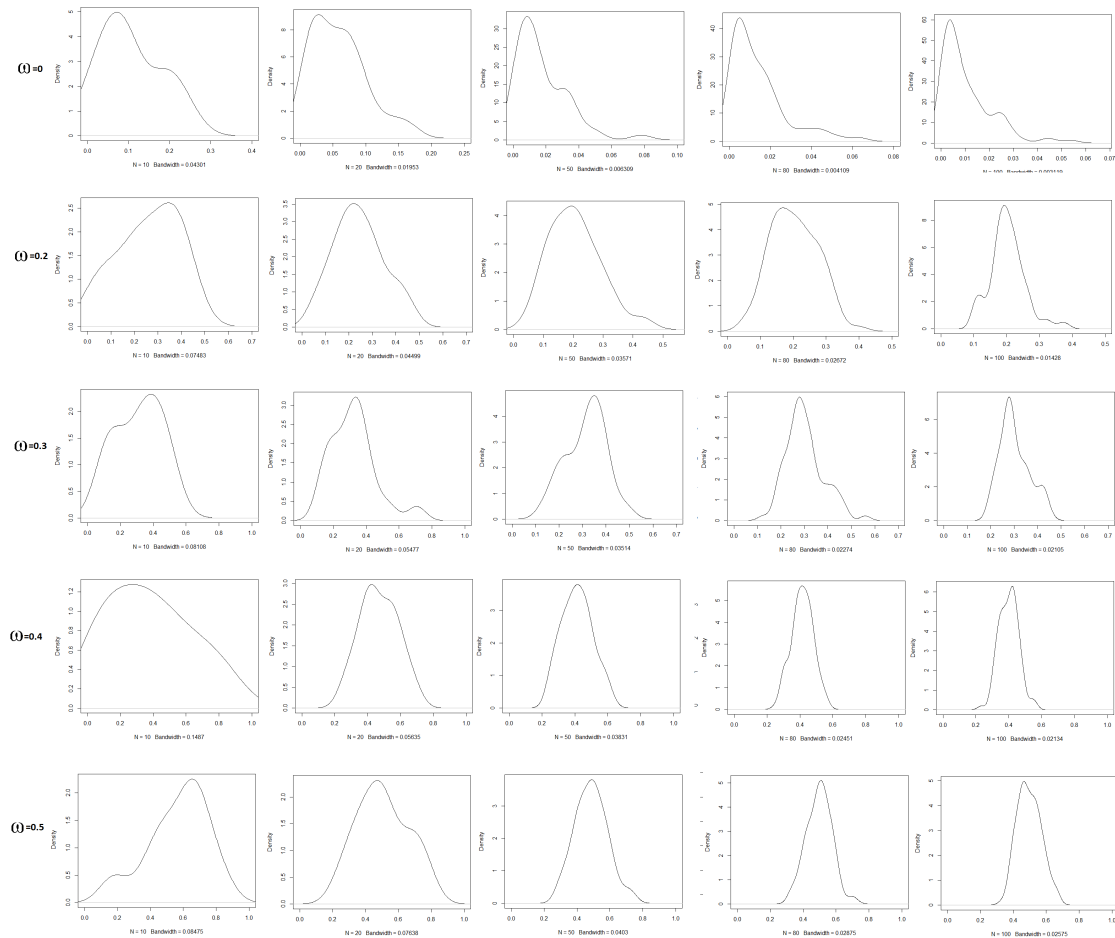


Figura 6.10: Función posterior del parámetro  $\omega$  para  $\lambda = 10$ .

La gráfica 6.11, muestra la densidad posterior para el parámetro  $\lambda$ , se observa que conforme crece el tamaño de muestra se tiene menos variabilidad y se concentra alrededor del parámetro  $\lambda$  fijado.



## 6.4. Intervalo de credibilidad (Estadística Bayesiana)

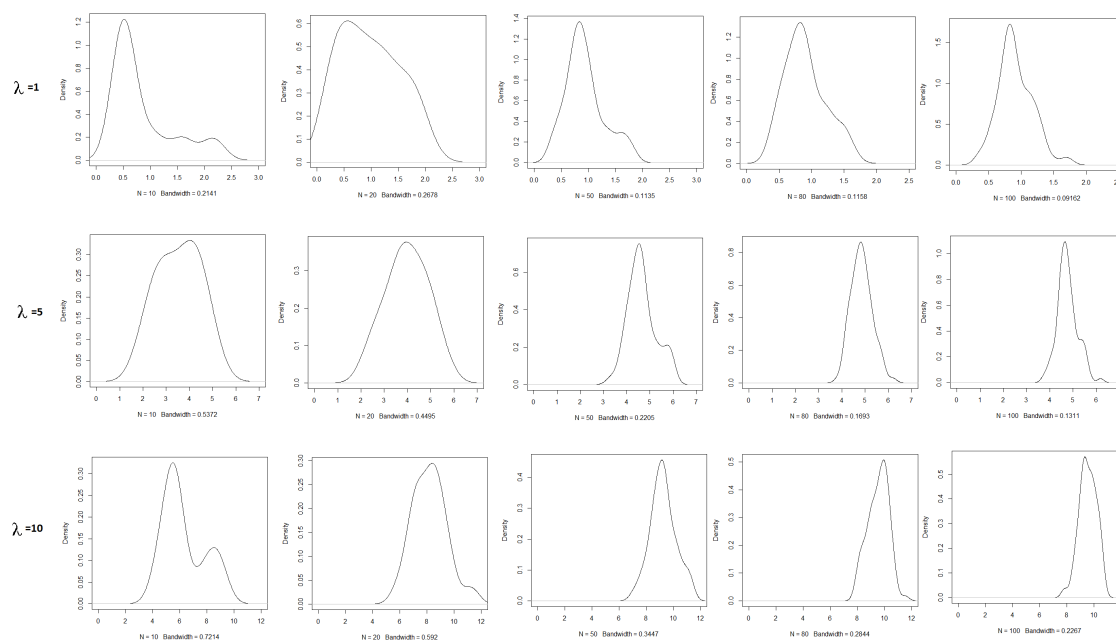


Figura 6.11: Función posterior para el parámetro  $\lambda$ .

## 6.4. Intervalo de credibilidad (Estadística Bayesiana)

Utilizando la función `hdi` del software `R` project, el cual calcula el intervalo de mayor densidad para una distribución de probabilidad o para una masa de probabilidad dada. A menudo se aplica a una distribución posterior bayesiana y se denomina “intervalo de densidad posterior más alto” o “intervalo de credibilidad”, pero se puede aplicar a cualquier distribución, incluidas las distribuciones a prioris. De esta manera se construyeron intervalos de credibilidad, para probar el juego de hipótesis:

$$H_0 : \omega = 0 \quad \text{contra} \quad H_1 : \omega \in (0, 1).$$

### 6.4.1. Potencia del intervalo de credibilidad

En la tabla 6.3 se presentan algunas potencias de la prueba propuesta, para diferentes niveles de tamaño de muestra, así como para valores diversos valores de  $\omega$  y  $\lambda = 1$ . La tabla ilustra el crecimiento de la potencia en función del tamaño de muestra, al igual que cuando los valores de  $\omega$  crecen.

## 6.4. Intervalo de credibilidad (Estadística Bayesiana)

---

Se esta probando el juego de hipótesis,

$$H_0 : \omega = 0 \quad \text{contra} \quad H_1 : \omega \neq 0, \quad \omega \in (0, 1).$$

Se rechaza la hipótesis nula si límite superior  $\geq \hat{Q}(.95)$ .

**Tabla 6.3:** Potencia de la prueba para  $\lambda = 1$ .

$n \backslash \omega$	0.5	0.4	0.3	0.2	0.1	0
10	1	1	0.948	.627	.594	0.051
20	1	1	1	.997	.995	0.052
50	1	1	1	1	1	0.051
80	1	1	1	1	1	0.052

En la tabla 6.5 se presentan algunas potencias de la prueba propuesta, para diferentes niveles de tamaño de muestra, así como para valores diversos valores de  $\omega$  y  $\lambda = 5$ . La tabla ilustra el crecimiento de la potencia en función del tamaño de muestra, al igual que cuando los valores de  $\omega$  crecen.

**Tabla 6.4:** Potencia de la prueba para  $\lambda = 5$ .

$n \backslash \omega$	0.5	0.4	0.3	0.2	0.1	0
10	1	1	0.999	0.963	0.624	0.051
20	1	1	1	1	0.994	0.052
50	1	1	1	1	1	0.051
80	1	1	1	1	1	0.051

En la tabla 6.5 se presentan algunas potencias de la prueba propuesta, para diferentes niveles de tamaño de muestra, así como para valores diversos valores de  $\omega$  y  $\lambda = 10$ . La tabla ilustra el crecimiento de la potencia en función del tamaño de muestra, al igual que cuando los valores de  $\omega$  crecen.

**Tabla 6.5:** Potencia de la prueba para  $\lambda = 10$ .

$n \backslash \omega$	0.5	0.4	0.3	0.2	0.1	0
10	1	1	0.992	0.936	0.573	0.052
20	1	1	1	1	0.963	0.051
50	1	1	1	1	1	0.051
80	1	1	1	1	1	0.051

## 6.4. Intervalo de credibilidad (Estadística Bayesiana)

### 6.4.2. Tabla de valores críticos

Los valores críticos de la prueba se obtienen utilizando simulación Monte-Carlo para varios tamaños de muestra y niveles de significancia bajo la hipótesis nula. En el cuadro 6.6 se presenta la tabla de valores críticos obtenida del proceso de simulación para la prueba propuesta basada en el intervalo de credibilidad, cuya regla de decisión es rechazar la hipótesis nula si el límite superior  $\geq \hat{Q}(0.95)$ .

**Tabla 6.6:** Tabla de valores críticos.

$\lambda$	1		5		10	
$\omega$	0.01	0.05	0.01	0.05	0.01	0.05
$n$						
5	0.523267728	0.529404397	0.5270759	0.4884988	0.5311996	0.4917339
10	0.37651563	0.369806033	0.3639475	0.3173742	0.387031	0.3237935
15	0.299797381	0.301094687	0.2981834	0.2408754	0.2816201	0.2240396
20	0.240691473	0.248195515	0.2172386	0.1739865	0.2532586	0.1968407
25	0.219341385	0.145506253	0.2179589	0.1606798	0.1935039	0.1446523
30	0.184534505	0.132347161	0.1906865	0.137928	0.164777	0.1196188
35	0.172628116	0.118011512	0.171747	0.1224041	0.1459168	0.1032385
40	0.156573114	0.108721164	0.1535949	0.1075983	0.1307914	0.08992483
45	0.134980627	0.084936106	0.1422953	0.09778467	0.1141791	0.07910606
50	0.127217537	0.075872062	0.1059824	0.07301086	0.1271386	0.08933486
60	0.105913533	0.065781969	0.1128021	0.07580311	0.09120013	0.05990278
70	0.096779042	0.052995662	0.09717287	0.065025	0.07775471	0.0517557
80	0.088030766	0.048659416	0.06705441	0.0448969	0.06935152	0.04447866
90	0.082286801	0.041169958	0.0600038	0.03924532	0.06019555	0.03946895
100	0.07717973	0.038148531	0.05461781	0.03560716	0.05506227	0.03601535

En la tabla 6.6 se tiene una tendencia logarítmica, aunque hay incrementos fuertes en el tamaño de muestra, los valores críticos no crecen excesivamente. Dada la situación, se realiza un estudio de regresión logarítmica para encontrar valores críticos que dependan del tamaño de muestra, así será posible encontrar el valor crítico para cualquier tamaño de muestra.

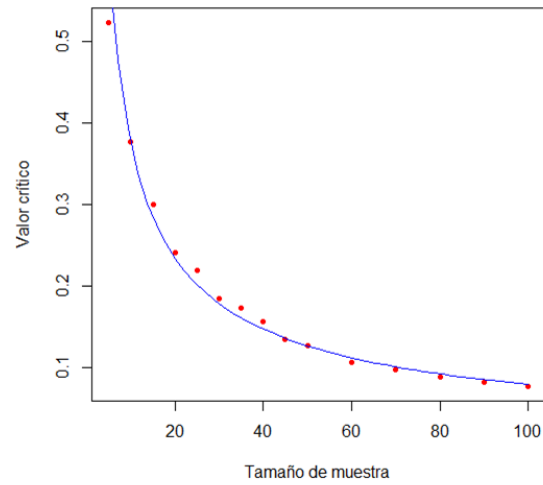
En la figura 6.12 se presenta la gráfica del análisis de regresión logarítmica para  $\lambda = 1$  y  $\alpha = 0.01$ , se obtiene la ecuación de regresión siguiente;

$$y = 1.759661x^{-0.672886}$$

donde  $y$  es el valor crítico y  $x$  el tamaño de muestra.

## 6.4. Intervalo de credibilidad (Estadística Bayesiana)

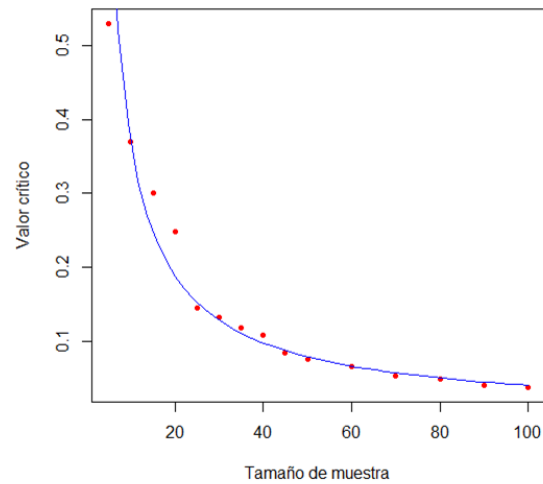
---



**Figura 6.12:** Ajuste de valores críticos para  $\lambda = 1$  y  $\alpha = 0.01$ .

De la misma forma, en la figura 6.13 se muestra el análisis de regresión logarítmico para  $\lambda = 1$  y  $\alpha = 0.05$ , la ecuación de regresión es la siguiente;

$$y = 3.280995x^{-0.9529831}.$$



**Figura 6.13:** Ajuste de valores críticos para  $\lambda = 1$  y  $\alpha = 0.05$ .

## 6.5. Intervalo de confianza (Estadística clásica)

---

### 6.4.3. Inferencia con el intervalo de credibilidad

Se simuló una muestra de tamaño  $n = 100$ ,  $\omega = 0.3$ ,  $\lambda = 2$ , utilizando una secuencia Gibbs de 10000 simulaciones y se obtuvo estadísticas descriptivas.

Parámetro	Media	sd	ECM	1er cuartil	mediana	3er cuartil
$\omega$	0.3605501	0.1628112	0.03014731	0.2369129	0.3477671	0.4705671
$\lambda$	2.031739	0.5792053	0.3361506	1.614798	1.999918	2.401530

## 6.5. Intervalo de confianza (Estadística clásica)

Es posible derivar una prueba basada en la normalidad asintótica de la estimación de los parámetros de la distribución Poisson inflada con ceros. A partir de las propiedades estadísticas del modelo ZIP, se puede obtener

$$E(\bar{Y}) = E(Y) = p\mu$$

y

$$Var(\bar{Y}) = \frac{1}{n}Var(Y) = \frac{1}{n}(p\mu + p\mu(\mu - p\mu)).$$

Aplicando el teorema límite central, asintóticamente (cuando  $n \rightarrow \infty$ ) la distribución de

$$Z = (\bar{Y} - p\mu)\sqrt{(p\mu + p\mu(\mu - p\mu))/n}$$

se distribuye como una distribución normal estándar. Así el intervalo de confianza se muestra en la siguiente ecuación;

$$\frac{\bar{y} - Z_{\alpha/2}\sqrt{(\bar{y} + \bar{y}(\hat{\lambda}_2 - \bar{y}))/n}}{\hat{\lambda}_2} \leq 1 - \omega \leq \frac{\bar{y} + Z_{\alpha/2}\sqrt{(\bar{y} + \bar{y}(\hat{\lambda}_2 - \bar{y}))/n}}{\hat{\lambda}_2}.$$

Y el estadístico basado en el intervalo de confianza es;

$$S = \frac{\bar{y} + Z_{\alpha}\sqrt{(\bar{y} + \bar{y}(\hat{\lambda}_2 - \bar{y}))/n}}{\hat{\lambda}_2}$$

donde se rechaza la hipótesis nula si  $S < 1$ .

### 6.5.1. Potencia del intervalo de confianza

En la tabla 6.7 se muestran las potencias para  $\lambda = 10$ ,  $\alpha = 0.05$ , diferentes tamaños de muestra y diferentes valores de  $\omega$ , se realizaron 1000 simulaciones.

## 6.5. Intervalo de confianza (Estadística clásica)

---

**Tabla 6.7:** Tabla de potencia para el estadístico basado en un intervalo de confianza.

N	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0$
10	0.944	0.851	0.599	0.317	0.000
20	0.999	0.990	0.895	0.580	0.000
50	1.000	1.000	1.000	0.972	0.000

# Capítulo 7

## EJEMPLO DE APLICACIÓN

### 7.1. Encuesta sobre mortalidad de niños menores de 5 años

Se ha considerado como ejemplo de aplicación una encuesta realizada en Bangladesh, como breve reseña sobre la encuesta de 2011 de BDHS (Encuesta Demográfica de salud de Bangladesh), BDHS(2011) es la sexta encuesta de su tipo realizada en Bangladesh, es una encuesta nacional de hombres y mujeres en edad reproductiva que proporciona información sobre los niveles de mortalidad infantil, las preferencias de fertilidad, el uso de métodos de planeación familiar y la salud materna, infantil y neonatal. La muestra para la BDHS(2011) es representativa a nivel nacional junto con representantes de cada división (hay 7 divisiones representativas en Bangladesh). 17842 mujeres en edad reproductiva (12-49 años de edad) y 3997 hombres aptos fueron entrevistados en la encuesta BDHS(2011). En la encuesta se les pregunta a las mujeres cuantos hijos menores de 5 años se han muerto, en la tabla 7.1 se muestran las frecuencias desde 0 muertes hasta 7 muertes.

**Tabla 7.1:** Mortalidad de menores de 5 años en Bangladesh.

Número de muertes	Frecuencia
0	12769
1	2442
2	563
3	155
4	44
5	23
6	4
7	4

## 7.1. Encuesta sobre mortalidad de niños menores de 5 años

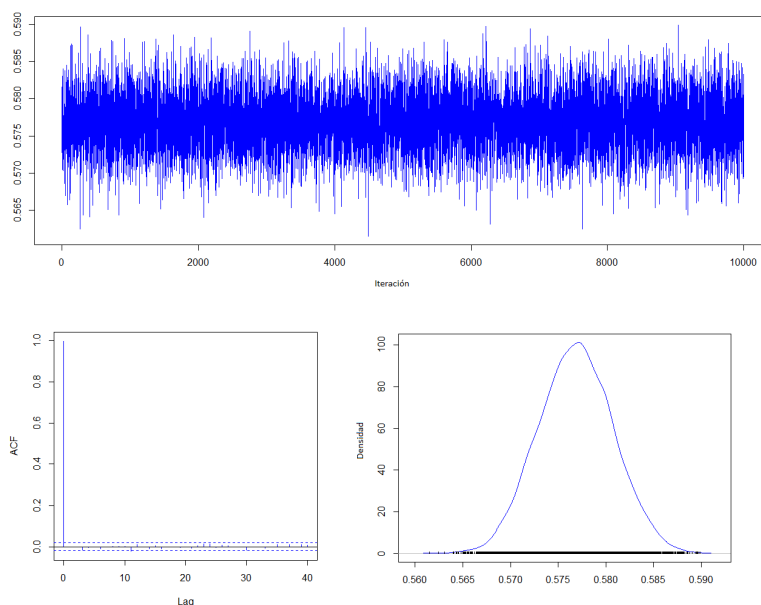
En la tabla 7.2 se muestran estadísticas descriptivas para la muestra de niños menores de 5 años que han muerto.

**Tabla 7.2:** Estadísticas descriptivas para los datos de mortalidad de menores de 5 años en Bangladesh.

Parámetro	Estimación	Media	sd	ECM	1er cuartil	mediana	3er cuartil
$\omega$	0.57143374	0.5714313	0.005507099	3.032629e-05	0.5676897	0.5714353	0.5751840
$\lambda$	0.63806205	0.6380938	0.008841491	7.816809e-05	0.6320765	0.6381293	0.6440869

El intervalo de credibilidad obtenido para el parámetro  $\omega$  es el siguiente; **(0.5609749,0.5826508)**. Se observa que el límite superior  $0.5826508 > 0.038148531$ , es mayor que el cuantil 95 para  $n = 100$ , por lo tanto, se rechaza la hipótesis nula y se concluye que los datos no tienen distribución Poisson. Por lo que se puede pensar en que se esta hablando de una distribución Poisson inflada con ceros.

En la figura 7.1 se muestran las iteraciones que se han generado con los datos de Bangladesh, ya que no se presentan grandes periodicidades, ni tendencia o paralelismo, se puede concluir la convergencia de la cadena. En la función de autocorrelación se tiene evidencia de independencia de observaciones. Respecto a la función de densidad posterior, ésta se centra aproximadamente en la estimación del parámetro  $\omega = 0.57143374$ .

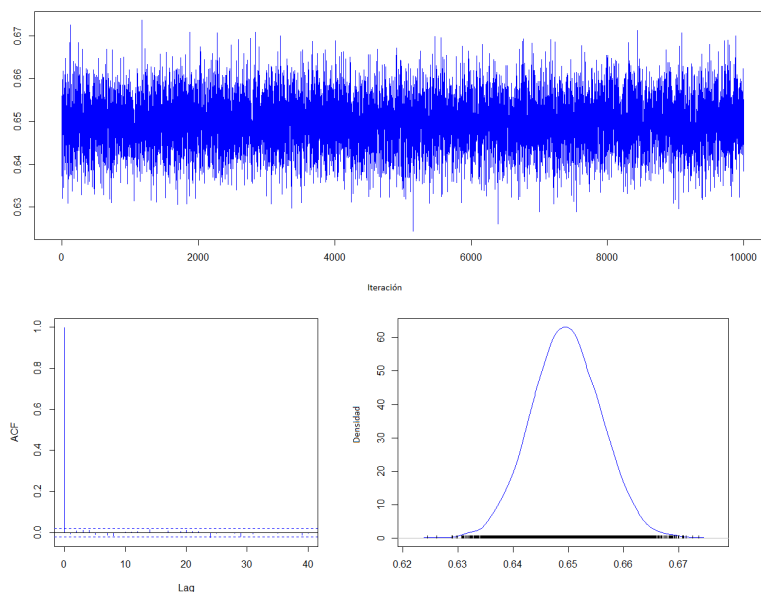


**Figura 7.1:** Función posterior para el parámetro  $\omega$ .

En la figura 7.2 se muestran las iteraciones que se han generado con los datos de Bangladesh, y se observa la convergencia de la cadena porque no se presentan grandes periodicidades, ni tendencias, también se obtuvo la función de autocorrelación y esta presenta observaciones independientes. Respecto a la función de densidad posterior, esta se centra aproximadamente en la estimación del parámetro  $\lambda = 0.63806205$ .



## 7.1. Encuesta sobre mortalidad de niños menores de 5 años



**Figura 7.2:** Función posterior para el parámetro  $\lambda$ .

El procedimiento para obtener la distribución posterior de  $\omega$  y  $\lambda$ , se programó también en el software SAS, con el que se obtuvieron resultados semejantes, los cuales se muestran en la tabla 7.3.

**Tabla 7.3:** Estadísticas descriptivas para los datos de mortalidad de menores de 5 años en Bangladesh obtenidas en software SAS.

Parámetro	Estimación	Media	sd	ECM	1er cuartil	mediana	3er cuartil
$\omega$	0.571433	0.5708	0.0110	0.000301	0.5634	0.5711	0.5783
$\lambda$	0.6380108	0.6378	0.0180	0.000485	0.6250	0.6380	0.6503

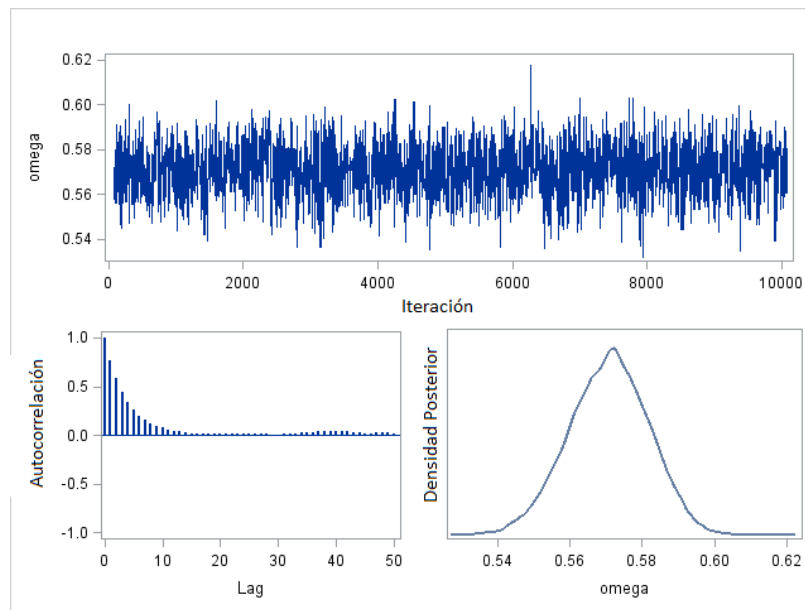
Se obtienen el intervalo de credibilidad para el parámetro  $\omega$  en SAS, **(0.5497, 0.5922)**, donde se observa que el límite superior  $0.5922 > 0.038148531$ , es mayor que el cuantil 95 para  $n = 100$ , por lo tanto, se rechaza la hipótesis nula y se concluye de la misma manera que en el software R.

En las figuras 7.3 y 7.4 se muestran los gráficos obtenidos en el programa SAS, para los parámetros  $\omega$  y  $\lambda$ , respectivamente. Como observación, el método que utiliza el procedimiento MCMC en SAS para generar las secuencias, es quemar o eliminar las primeras  $k - m$  observaciones, y se genera una muestra de tamaño  $m$  para la distribución posterior. Sobre esta muestra de tamaño  $m$  se realiza inferencias. Siendo éste un método distinto al seguido en el software R. Es intuitivo pensar que la función de autocorrelación y la cadena de Markov deben ser un poco diferentes, pero deben obtener una distribución posterior semejante, así como resultados que se pueda concluir de la misma forma .

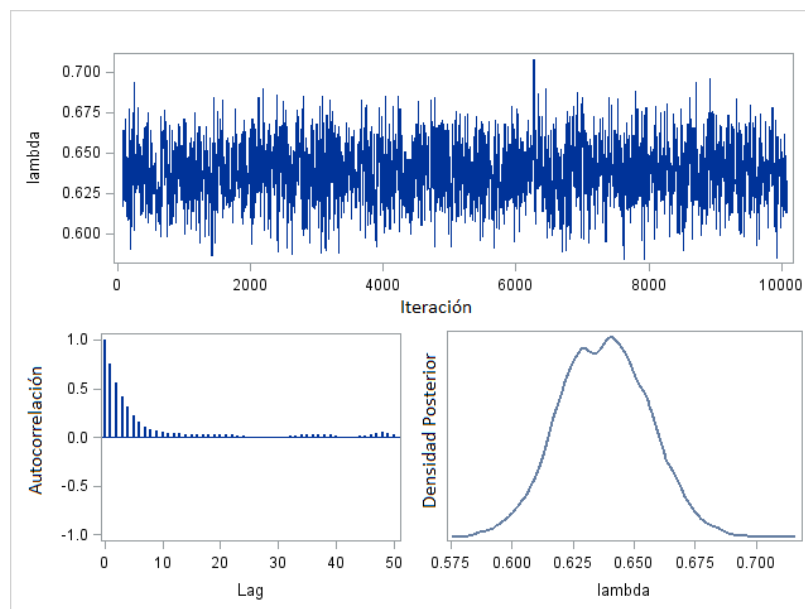
En las figuras 7.3 y 7.4 se muestran los gráficos obtenidos en el programa SAS, para los

## 7.1. Encuesta sobre mortalidad de niños menores de 5 años

parámetros  $\lambda$  y  $\omega$ , respectivamente. El método que utiliza el procedimiento MCMC en SAS para generar las secuencias es el siguiente; elimina las primeras  $k - m$  observaciones para generar una muestra de tamaño  $m$  de la distribución posterior. Sobre ésta muestra de tamaño  $m$  se realizan inferencias. Ya que éste método es distinto al seguido en el software R, es posible esperar que la función de autocorrelación y la cadena de Markov sean un poco diferentes, pero la distribución posterior obtenida y los resultados deben ser semejantes, se debe llegar a las mismas conclusiones .



**Figura 7.3:** Función posterior para el parámetro  $\omega$ .



**Figura 7.4:** Función posterior para el parámetro  $\lambda$ .

## 7.2. Número de raíces en brotes micropropagados

Como un ejemplo ilustrativo del modelo Poisson inflados con ceros, se considera los datos proporcionados por (Marin et al., 1993), los datos son el número de raíces producidas por 130 brotes micropropagados de la manzana columnar cultivar Trajan.

En la tabla 7.4 muestra la cantidad de brotes que produjeron 0, 1, ..., 12 raíces, los recuentos que excedieron 12 raíces se muestran individualmente.

**Tabla 7.4:** Frecuencia de número de raíces en los brotes micropropagados.

Número de raíces	Frecuencia
0	62
1	7
2	7
3	8
4	8
5	6
6	10
7	4
8	2
9	7
10	4
11	2
12	3

Se obtienen el intervalo de credibilidad para el parámetro  $\omega$ , (**0.3417398 0.5905079**), se observa que el límite superior  $0.5905079 > 0.03560716$ , es mayor que el cuantil 95 para  $n = 130$ , por lo tanto, se rechaza la hipótesis nula y se concluye que los datos no tienen distribución Poisson. Por lo que se puede pensar en que se esta hablando de una distribución Poisson inflada con ceros.

En la figura 7.5 se muestran las iteraciones que se han generado con los datos de número de raíces en brotes micropagados, y se observa la convergencia de la cadena porque no se presentan grandes periodicidades, ni tendencia o paralelismos, también se obtuvo la función de autocorrelación y esta presenta observaciones independientes. Respecto a la función de densidad posterior, esta se centra aproximadamente en la estimación del parámetro  $\omega = 0.4746593$ .

## 7.2. Número de raíces en brotes micropropagados

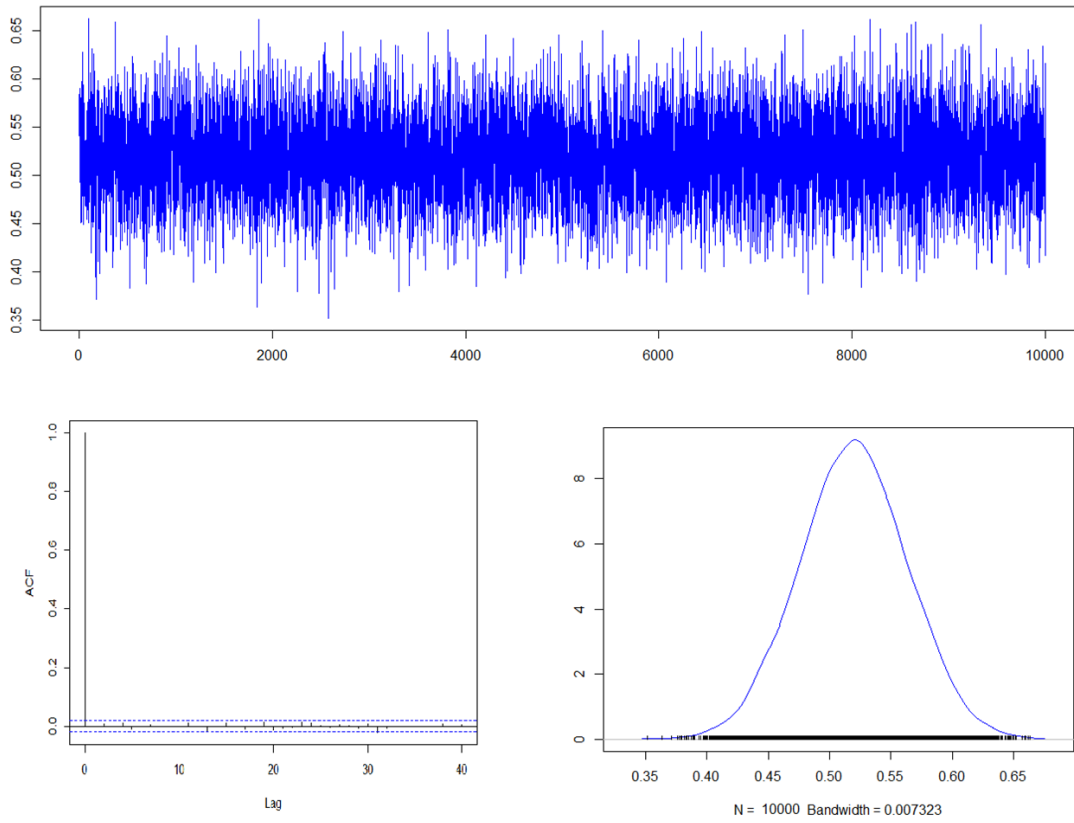


Figura 7.5: Función posterior para  $\omega$ .

# Capítulo 8

## CONCLUSIONES

Las conclusiones a las que se puede llegar en base al análisis realizado y tomando en cuenta los estudios de simulación MCMC y en el ejemplo de aplicación son las siguientes:

Fue posible construir una prueba basada en el intervalo de credibilidad del parámetro  $\omega$  de la distribución Poisson inflada con ceros para comparar el modelo Poisson clásico con el inflado con ceros.

Los diagnósticos de las cadenas de Markov fueron satisfactorios, con lo que se asegura la obtención de buenos resultados de la distribución posterior.

Como resultado del análisis de la prueba propuesta, se puede concluir que es una buena prueba con respecto a las que se encuentran en la literatura debido a que presenta buenas potencias.

El desempeño de la prueba propuesta fue aplicado a datos de una encuesta realizada en Bangladesh, y a datos de número de raíces en brotes micropropagados y los resultados fueron los esperados, dado que la hipótesis nula fue rechazada y se logró una buena estimación de los parámetros.

Fue posible desarrollar un modelo bayesiano y construir una prueba de la cual se obtuvieron resultados satisfactorios.

# REFERENCIAS

- Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag New York. ISBN 978-1-4757-4286-2.
- Campbell, I. C., Luczynski, K. J. y Hood, S. K. (1991). Putting knowledge-based concepts to work for mechanical design.
- Casella, G. y Berger, R. (2002). *Statistical Inference*. Duxbury advanced series in statistics and decision sciences. Thomson Learning, segunda edición. ISBN 9780534243128.
- Cochran, W. G. (1954). Some Methods for Strengthening the Common  $\chi^2$  Tests. *Biometrics*, 10, 4, 417–451. ISSN 0006341X, 15410420.
- Dankmar, B. (1998). Zero-Inflated Poisson Models and C.A.MAN: A Tutorial Collection of Evidence. *Biometrical Journal*, 40, 7, 833–843.
- den Broek, J. V. (1995). A Score Test for Zero Inflation in a Poisson Distribution. *Biometrics*, 51, 2, 738–743. ISSN 0006341X, 15410420.
- el Shaarawi, A. H. (1985). Some goodness-of-fit methods for the Poisson plus added zeros distribution. *Applied and Environmental Microbiology*, 49, 5, 1304–1306.
- Freund, A., Li, K. y Baltes, P. (1999). The role of selection, optimization, and compensation in successful aging. *Action and development: Origins and functions of intentional self-development*, 401–434.
- Gelfand, A. E., Hills, S. E., Racine-Poon, A. y Smith, A. F. M. (1990). Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *Journal of the American Statistical Association*, 85, 412, 972–985. ISSN 01621459.
- Gelfand, A. E. y Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 410, 398–409. ISSN 01621459.
- Gelman, A. y Rubin, D. (1991). An overview and approach to inference from iterative simulation. Inf. téc., University of California-Berkeley, Department of Statistic.
- Gelman, A. y Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 4, 457–472. ISSN 08834237.
- George, E. I. y McCulloch, R. E. (1993). Variable Selection via Gibbs Sampling. *Journal of the American Statistical Association*, 88, 423, 881–889.

## REFERENCIAS

---

- Ghosh, S. K., Mukhopadhyay, P. y Lu, J.-C. (2006). Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference*, 136, 4, 1360 – 1375. ISSN 0378-3758.
- Mamun, A. A. (2014). *Zero-inflated regression models for count data: an application to under-5 deaths*. Proyecto Fin de Carrera, Ball State University.
- Miaou, S.-P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention*, 26, 4, 471–482.
- Müller, P. (1991). A generic approach to posterior integration and gibbs sampling. Inf. téc.
- Quinn, G. y Keough, M. (2002). *Experimental Design and Data Analysis For Biologists*. Cambridge University Press. ISBN 9780511806384.
- R (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Raftery, A. E. y Banfield, J. D. (1991). Stopping the Gibbs Sampler, the Use of Morphology, and Other Issues in Spatial Statistics. *Annals of the Institute of Statistical Mathematics / edited by the Institute of Statistical Mathematics*, 43, 1, 32–43.
- Rao, C. R. y Chakravarti, I. M. (1956). Some Small Sample Tests of Significance for a Poisson Distribution. *Biometrics*, 12, 3, 264–282. ISSN 0006341X, 15410420.
- Ridout, D. C., M.S. y Hinde, J. (1998). Models for counts data with many zeros. *Proceedings of the XIXth International Biometric Conference*, Cape Town, Invited Papers, 179–192.
- Ridout, M., Hinde, J. y Demétrio, C. G. B. (2001). A Score Test for Testing a Zero-Inflated Poisson Regression Model Against Zero-Inflated Negative Binomial Alternatives. *Biometrics*, 57, 1, 219–223. ISSN 1541-0420.
- Ritter, C. y Tanner, M. A. (1990). The gridy gibbs sampler. Inf. téc., University of Rocheste.
- Tanner, M. A. (1991). *Tools for Statistical Inference*. 67. Springer-Verlag New York. ISBN 978-1-4684-0510-1.
- Thas, O. y Rayner, J. C. W. (2005). Smooth Tests for the Zero-Inflated Poisson Distribution. *Biometrics*, 61, 3, 808–815. ISSN 1541-0420.
- Tierney, L. (1991). Markov chains for exploring posterior distributions. Inf. téc., University of Minnesota, School of Statistic.
- Xie, M., He, B. y Goh, T. (2001). Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis*, 38, 2, 191 – 201. ISSN 0167-9473.

# APÉNDICES

## Apéndice A: Programas de R usados en el trabajo

Para realizar el ajuste de la distribución Poisson inflada con ceros es necesario instalar el paquete VGAM en el programa R. Las rutina funciona en el paquete R-3.5.0 ó superior.

```
#Estimación de los parámetros con la función optim
library(VGAM)
is.wholenumber=function(y,tol=.Machine$double.eps){abs(y-round(y))<tol}
ziip<-function(omega,lambda,x){
  n=length(x)
  v=matrix(NA,nrow=n,ncol=1)
  for (i in 1:n){
    a=x[i]
    if(omega>0 & omega<=1 & lambda>0 & a>0 & is.wholenumber(x==a)){
f_y=((1-omega)*(exp(-lambda)*lambda^(a))/factorial(a))
}
    if (omega>0 & omega<=1 & lambda>0 & a==0){
f_y=omega+((1-omega)*(exp(-lambda)))
}
    if(f_y<0 | f_y>1){f_y=0}
    v[i]=f_y
  }
  return(v)
}
vero=function(par){
  temp1=matrix(NA,n,1)
  temp=apply(datos,1,ziip,omega=par[1],lambda=par[2])
  temp1=log(temp)
  vero=sum(temp1)
  -vero
}
par.ajus=optim(c(.01,0.1),vero,method="L-BFGS-B",lower=c(0.0001,0.0001),
```



## APÉNDICES

---

```
upper=c(0.999,100))
par.ajus

#VGAM
ajus=coef(fit.0)
x=exp(ajus[1])
theta=x/(1+x)
lambda=exp(ajus[2])
theta
lambda

#Ajuste realizado a datos de Bangladesh utilizando los estimadores
#de momentos como valores iniciales
datosBan<-as.matrix(read.csv(file="datosBan.csv",head=TRUE))
#Ajuste con la función optim
is.wholenumber=function(y,tol=.Machine$double.eps){abs(y-round(y))<tol}
ziip<-function(omega,lambda,x){
  n=length(x)
  v=matrix(NA,nrow=n,ncol=1)
  for (i in 1:n){
    a=x[i]
    if(omega>0 & omega<=1 & lambda>0 & a>0 & is.wholenumber(x==a)){
f_y=((1-omega)*(exp(-lambda)*lambda^(a))/factorial(a))
}
    if (omega>0 & omega<=1 & lambda>0 & a==0){
f_y=omega+((1-omega)*(exp(-lambda)))
}
    if(f_y<0 | f_y>1){f_y=0}
    v[i]=f_y
  }
return(v)
}
vero=function(par){
  temp1=matrix(NA,n,1)
  temp=apply(datosBan,1,ziip,omega=par[1],lambda=par[2])
  temp1=log(temp)
  vero=sum(temp1)
  -vero
}
omegamomentos<-(((sum(datosBan^2)/n)-mean(datosBan)-(mean(datosBan))^2)/
((mean(datosBan))^2))/(1+(((sum(datosBan^2)/n)-mean(datosBan)-
(mean(datosBan))^2)/((mean(datosBan))^2)))
lambdamomentos<-(sum(datosBan))/(n*(1-omegamomentos))
momentos<-c(omegamomentos,lambdamomentos)
par.ajus=optim(momentos,vero,method="L-BFGS-B",lower=c(0.0001,0.0001),
upper=c(0.999,100))
par.ajus
```

## APÉNDICES

---

```
#Ajuste de datos Bangladesh con VGAM
library(VGAM)
fit.0<-vglm(datosBan[,1]~1,zipoisson(lpstr0="logit",llambda="loge"),
trace=TRUE)
summary(fit.0)
```

Para la implementación del algoritmo Gibbs es necesario instalar los paquetes **coda** y **HDInterval** del software [R](#).

```
library(coda)
library(HDInterval)

gibbs1<-function(n,omega,lambda,alpha,beta,k,theta,s){
v<-rpois(n,lambda)
ber<-rbinom(n,1,omega)
datos<-v*(1-ber)
muestreador<-matrix(0,ncol=4,nrow=s+1)
colnames(muestreador)=c("p","binomial",
"omega","lambda")
rownames(muestreador)=seq(0,s,1)
a<-which(datos==0)
yi<-which(datos>0)
yii<-datos[yi]
p<-(omega)/(omega+(1-omega)*exp(-lambda))
m<-length(a==0)
muestreador[1,1]<-p
muestreador[1,2]<-rbinom(1,m,p)
muestreador[1,3]<-rbeta(1,shape1=muestreador[1,2]
+alpha,n-muestreador[1,2]+beta)
muestreador[1,4]<-rgamma(1,shape=sum(yii)+k,scale=
(n-muestreador[1,2]+theta)^-1)
for(i in 1:s){
muestreador[i+1,1]<-(muestreador[i,3])/(muestreador[i,3]
+(1-muestreador[i,3])*exp(-muestreador[i,4]))
muestreador[i+1,2]<-rbinom(1,m,muestreador[i+1,1])
muestreador[i+1,3]<-rbeta(1,shape1=muestreador[i+1,2]
+alpha,n-muestreador[i+1,2]+beta)
muestreador[i+1,4]<-rgamma(1,shape=sum(yii)+k,scale=
(n-muestreador[i,2]+theta)^-1)
}
muestra<-muestreador[s+1,3]
return(muestra)
}

MGibbs<-function(n,omega,lambda,alpha,beta,k,theta,s,m){
```

## APÉNDICES

---

```
mg<-matrix(NA,ncol=1,nrow=m)
for(i in 1:m){
mg[i,1]=gibbs1(n,omega,lambda,alpha,beta,k,theta,s)
}
return(mg)
}

muestrasinter<-function(n,omega,lambda,alpha,beta,k,theta,s,m,columnas){
mue<-matrix(NA,ncol=columnas,nrow=m)
for(i in 1:columnas){
mue[,i]=MGibbs(n,omega,lambda,alpha,beta,k,theta,s,m)
}
return(mue)
}

START=Sys.time()
a<-muestrasinter(10,0.1,1,1,1,1,2,10000,10,1000)
int<-t(apply(a, 2,hdi,credMass=.95))
omega=0.1
m=1000
l=1
contador=0
for (i in 1:m){
if (int[i,1]<=omega & int[i,2]>=omega) {contador=contador+1}
}
potencia<-(contador/1000)
potencia
#END=Sys.time()
#print(START)
write.csv(potencia,"contador_0.1_10.csv")
write.csv(a,"muestra_0.1_10.csv")
write.csv(int,"intervalos_0.1_10.csv")
print(END)
```

En la siguiente sección, se muestran los comandos utilizados para obtener las potencias de las pruebas encontradas en la literatura:

```
rm(list=ls(all=TRUE))
install.packages("VGAM")
library(VGAM)
#Ho:Poisson(lambda)
#Ha:ZIP(omega,lambda)
#Prueba de Cochran
cochran=(n0-n*p0)/sqrt((n*p0*(1-p0-mu1*p0)))
#p-value
pval=pnorm(cochran, mean = 0, sd = 1, lower.tail = F, log.p = F)
```

## APÉNDICES

---

```
return(pval)
}
nSIM=50000
alpha=0.05
pSIM=matrix(rep(NA,nSIM),ncol=1)
START=Sys.time()
for(i in 1:nSIM){
#print(i)
pSIM[i,]=fp(n=20,lambda0=5,omega0=0,alpha=0.05)
}
rH0=which(pSIM[,1]<=alpha)
tamaño=length(rH0)/nSIM
END=Sys.time()
tamaño
START
END
#Simulaciones (POTENCIA DE LA PRUEBA)
nSIM=1000
alpha=0.05
pSIM=matrix(rep(NA,nSIM),ncol=1)
START=Sys.time()
for(i in 1:nSIM){
#print(i)
pSIM[i,]=fp(n=50,lambda0=10,omega0=0,alpha=0.05)
}
H0=which(pSIM[,1]>alpha)
beta=length(H0)/nSIM
potencia=1-beta
END=Sys.time()
potencia
START
END
#Prueba Score
s1=(n0-n*p0)^2/(n*p0*(1-p0)-n*mu1*p0^2)
#p-value
pval=pchisq(s1, 1, ncp = 0, lower.tail = FALSE, log.p = FALSE)
return(pval)
}
#Simulaciones (TAMAÑO DE LA PRUEBA)
nSIM=50000
alpha=0.05
pSIM=matrix(rep(NA,nSIM),ncol=1)
START=Sys.time()
for(i in 1:nSIM){
#print(i)
pSIM[i,]=fp(n=20,lambda0=5,omega0=0,alpha=0.05)
}
```

## APÉNDICES

---

```
rH0=which(pSIM[,1]<=alpha)
tamaño=length(rH0)/nSIM
END=Sys.time()
tamaño
START
END
#Simulaciones (POTENCIA DE LA PRUEBA)
nSIM=30000
alpha=0.05
pSIM=matrix(rep(NA,nSIM),ncol=1)
START=Sys.time()
for(i in 1:nSIM){
#print(i)
pSIM[i,]=fp(n=20,lambda0=5,omega0=0,alpha=0.05)
}
H0=which(pSIM[,1]>alpha)
beta=length(H0)/nSIM
potencia=1-beta
END=Sys.time()
potencia
START
END
#Prueba de intervalo de confianza
fp=function(n,lambda0,omega0,alpha){
datos=matrix(rzipois(n,lambda=lambda0,pstr0=omega0),n,1)
datos=as.matrix(datos)
mu1=mean(datos)
n=length(datos)
ceros=which(datos==0)
n0=length(ceros)
zval<-qnorm(alpha, mean = 0, sd = 1, lower.tail = FALSE, log.p = FALSE)
p0=exp(-mu1)
mu2=mu1/((1-(n0/n))/(1-p0))
#interval
interval=(mu1+zval*sqrt((mu1+mu1*(mu2-mu1))/n))/mu2
#p-value
return(interval)
}
#Simulaciones (POTENCIA DE LA PRUEBA)
nSIM=1000
alpha=0.05
pSIM=matrix(rep(NA,nSIM),ncol=1)
START=Sys.time()
for(i in 1:nSIM){
#print(i)
pSIM[i,]=fp(n=10,lambda0=10,omega0=0.3,alpha=0.05)
}
```

## APÉNDICES

---

```
H0=which(pSIM[,1]>=1)
beta=length(H0)/nSIM
potencia=1-beta
END=Sys.time()
potencia
START
END
```

### Apéndice B: Programas de SAS usado en el trabajo

```
*Ajuste de datos de Bangladesh;
FILENAME REFFILE '/folders/myfolders/datosBan.csv';
PROC IMPORT DATAFILE=REFFILE
DBMS=CSV
OUT=DATOS;
GETNAMES=YES;
PROC CONTENTS DATA=DATOS;
PROC PRINT DATA=DATOS;
proc genmod data=DATOS;
model v1=/ dist=zip;
zeromodel / link=logit;
run;
```

```
*Implementación del algoritmo Gibbs;
ODS graphics ON;
proc mcmc data=DATOS nbi=80 nmc=10000 outpost=regOut
  STATS = all DIAG = all plots=all; *seed=1;
parms omega .4 lambda 1;
prior omega~beta(1,1);
prior lambda~gamma(1,scale=2);
llike=LOG(omega*(V1 eq 0) + (1-omega)*PDF("poisson", V1, lambda));
model GENERAL(llike);
run;
```