



UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

CÁLCULO Y COMPARACIÓN DE P-VALORES MEDIANTE REMUESTREO
PARAMÉTRICO Y REMUESTREO CONDICIONAL EN BONDAD DE AJUSTE

DEFENSA DE TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
JESÚS IVÁN BELTRÁN BELTRÁN

TUTOR PRINCIPAL:
FEDERICO O'REILLY TOGNO
[INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS, UNAM](#)

MIEMBROS DEL COMITÉ TUTOR:
JOSÉ MARÍA GONZÁLEZ-BARRIOS MURGUÍA
[INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS, UNAM](#)
ELIANE REGINA RODRIGUES
[INSTITUTO DE MATEMÁTICAS, UNAM](#)

CIUDAD UNIVERSITARIA, CD. MX.

22 DE SEPTIEMBRE DE 2016



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

“Básicamente, no estoy interesado en hacer investigación. Estoy interesado en el entendimiento, que es algo muy diferente.”

David Harold Blackwell

Resumen

Generalmente en el caso de bondad de ajuste compuesto cuando la familia de distribuciones a probar es discreta, las estadísticas de prueba dependen fuertemente del valor del parámetro desconocido. Cuando la familia a probar admite una estadística suficiente, recientemente se ha propuesto utilizar la distribución condicional, de esta manera, la estadística de prueba ya no depende del parámetro desconocido. En este trabajo mediante la teoría de muestreo condicional se hace un estudio exhaustivo para la familia Poisson, binomial negativa y binomial, respectivamente, comparando la potencia de varias estadísticas de prueba. Se proponen dos nuevas estadísticas; una de las cuales resultó ser muy potente e igual de competitiva que las tradicionales estadísticas de prueba de la familia Cramér-von Mises.

En el capítulo dos se estudian las distribuciones exponencial truncada y la Gaussiana-inversa. Se hace un estudio comparativo de potencias de la estadística Anderson-Darling y una prueba muy reciente basada en la información de Kullback-Leibler. Se hace una comparación de los dos enfoques utilizados; uno que utiliza el remuestreo paramétrico y el otro que se basa en remuestreo condicional, mostrando una fuerte correlación entre ambos procedimientos.

En el capítulo tres se estudia la distribución exponencial para dos casos; cuando no existe censura y en presencia de datos censurados tipo II. Se hace un estudio de potencia para 4 estadísticas de prueba, dos estadísticas clásicas y dos muy recientes. Y por último, en el capítulo 4, se exploran dos artículos recientes en donde utilizan un enfoque de la estadística Bayesiana para probar bondad de ajuste. El primero usa el enfoque de la estadística Bayesiana paramétrica, es decir, utiliza una distribución inicial paramétrica, y se basa en la noción del p-valor predictivo posterior. En el segundo trabajo se basan en estadística Bayesiana no paramétrica, utilizando una distribución inicial no paramétrica, particularmente usando el proceso de Dirichlet.

Índice general

Resumen	v
Contenido	vi
Antecedentes	ix
1. Pruebas de bondad de ajuste caso discreto	1
1.1. Introducción	1
1.2. Nuevas estadísticas de prueba	10
1.3. Muestras condicionales	12
1.3.1. Distribución Rao-Blackwell	16
1.4. Estudio de potencia	18
1.4.1. Distribución Poisson	19
1.4.2. Distribución binomial negativa	20
1.4.3. Distribución binomial	21
1.5. Ejemplos	21
1.5.1. Muertes por patadas de caballo en el ejército prusiano	21
1.5.2. Ácaros rojos en las hojas de manzana	22
1.5.3. 10 muestras de cinco ratones	23
1.6. Comentarios finales	24
1.7. Tablas del estudio de simulación	25
2. Remuestreo paramétrico vs remuestreo condicional	41
2.1. Estadísticas de prueba	43
2.2. Distribución exponencial truncada	45
2.2.1. Estudio de potencia	46
2.3. Distribución Gaussiana-inversa	47
2.3.1. Estudio de potencia	51
3. Distribución exponencial con datos censurados tipo II	57
3.1. Introducción	57
3.2. La transformación J y la transformación N	58
3.3. Estadísticas de prueba	59

3.4. Estudio de potencia	60
4. Enfoque Bayesiano	67
4.1. Uso del p-valor predictivo posterior	67
4.1.1. Distribución normal	69
4.1.2. Estudio de potencia	70
4.2. Proceso de Dirichlet y su medida base	72
4.2.1. Medida de probabilidad aleatoria y el proceso de Dirichlet	73
4.2.2. Distancia entre el proceso de Dirichlet y su medida base	76
4.2.3. Bondad de ajuste	77
4.2.4. Ejemplos y un estudio de potencia	80
5. Conclusiones y perspectivas	85
Bibliografía	87

Antecedentes

Para resolver el problema compuesto de bondad de ajuste, donde el objetivo es determinar la distribución de la estadística de prueba bajo \mathcal{H}_0 , en una población con parámetros desconocidos, generalmente se estudia las propiedades asintóticas de la distribución de la estadística usando los estimadores de los parámetros. Dicha estadística de prueba suele ser función del llamado proceso empírico con parámetros estimados. Este proceso requiere de un gran esfuerzo para identificar su comportamiento asintótico. Aunado a lo anterior, en la mayoría de los casos el comportamiento asintótico de la distribución depende de los parámetros desconocidos con lo que, al tomar una decisión (respecto a la hipótesis nula) en estas condiciones, se introduce una mayor incertidumbre. Considerando además que las muestras utilizadas son de tamaño finito y las propiedades en que se fundamenta la prueba son asintóticas (cuando n tiende a infinito), es importante tener procedimientos que sean eficientes.

Una manera para disminuir la incertidumbre la proporcionan los recursos actuales de cómputo, junto con técnicas de re-muestreo paramétrico (bootstrap). Otra posibilidad es trabajar con los valores de la estadística de prueba pero utilizando la distribución condicional de la estadística de prueba dada una estadística suficiente; así en lugar de una simulación de la distribución, se simula la distribución condicional. Esto se fundamenta en las propiedades que tiene la distribución condicional.

Con este procedimiento, se evalúa el p-valor (condicional) cuyo valor puede ser tan exacto como se desee, porque la exactitud depende del número de simulaciones de las muestras condicionales.

En este contexto se ubican los trabajos publicados por [Engen & Lillegard \(1997\)](#), [Lindqvist & Taraldsen \(2005\)](#), [O'Reilly & Gracia-Medrano \(2006\)](#), [González-Barrios](#)

[et al. \(2010\)](#) y otros citados en la bibliografía, en donde el paradigma, es simular el p-valor (condicional) con simulaciones condicionales.

En el presente trabajo, se ha dividido el problema de bondad de ajuste en el caso discreto y el caso continuo. Sobre el caso discreto se realizaron estudios comparativos que se describen en el capítulo 1. En capítulo 2 se estudia la distribución exponencial truncada y la Gaussiana-inversa, donde se usan dos enfoques, el remuestreo paramétrico y el remuestreo condicional. En el capítulo 3 se aborda el problema de bondad de ajuste para la distribución exponencial en el caso de no tener censura, y en el caso de datos censurados tipo II. Por último en el capítulo 4, se exploran dos artículos recientes. El primero usa el enfoque de la teoría Bayesiana paramétrica. Mientras que en el segundo se basa en la estadística Bayesiana no paramétrica, particularmente utilizan el proceso de Dirichlet.

Capítulo 1

Pruebas de bondad de ajuste compuestas para distribuciones discretas

1.1. Introducción

Las pruebas de bondad de ajuste se pueden dividir en dos áreas de investigación, la primera es cuando la distribución a probar es discreta, que es llamado caso discreto y la segunda cuando la distribución a probar es continua, que se refiere como caso continuo. Esto se debe a que las estadísticas de prueba en el caso discreto y continuo tienen propiedades diferentes, por ejemplo, en el caso continuo, las estadísticas basadas en un proceso empírico (por ejemplo, las estadísticas de prueba de la familia Crámer-von Mises), bajo condiciones muy generales, este proceso converge a un Browniano atado, y para el caso discreto no se cumple esta convergencia. Aunque para el caso discreto existen resultados análogos donde se define un símil al proceso empírico del caso continuo. También, por ejemplo en el caso continuo, la transformación de la muestra, mediante la función de distribución garantiza la uniformidad. Es por esto que en este capítulo se discute el caso discreto, y en los siguiente capítulos se discute el caso continuo.

Sea X_1, \dots, X_n una muestra aleatoria de una distribución discreta F . Supongamos sin pérdida de generalidad que los valores que toma la variable aleatoria son $0, 1, 2, \dots$; y se quiere probar $\mathcal{H}_0 : F \in \mathcal{F}_\theta$ vs. $\mathcal{H}_1 : F \notin \mathcal{F}_\theta$, con \mathcal{F}_θ alguna familia

de distribuciones discretas con parámetro θ (conocido o desconocido), cuando θ es conocido al problema se le llama caso simple y cuando θ es desconocido es llamado caso compuesto o con parámetros desconocidos. En este trabajo tratamos solo el último caso, ya que el caso simple ha sido estudiado y resuelto de muchas maneras. Particularmente, se estudia los casos cuando \mathcal{F}_θ corresponde a la distribuciones Poisson, binomial negativa y binomial, respectivamente. Para las distribuciones binomial negativa y binomial se supone que se conoce un parámetro.

A continuación se da la notación para las distribuciones Poisson, binomial y binomial negativa, para luego dar algunas estadísticas de prueba conocidas en la literatura para dichas distribuciones.

Se dice que X tiene una distribución Poisson con parámetro $\lambda > 0$, denotado por $X \sim Poisson(\lambda)$ si y sólo si

$$P(X = x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \quad \text{si } x = 0, 1, \dots$$

Se dice que X tiene una distribución binomial negativa con parámetros $r \geq 1$ y $0 < p < 1$, con r conocido, denotado por $X \sim NB(r, p)$ si y sólo si

$$P(X = x; p) = \binom{r+x-1}{x} p^r (1-p)^x \quad \text{si } x = 0, 1, \dots$$

Se dice que X tiene una distribución binomial con parámetros $m \geq 1$ y $0 < p < 1$, con m conocido, denotado por $X \sim Bin(m, p)$ si y sólo si

$$P(X = x; p) = \binom{m}{x} p^x (1-p)^{m-x} \quad \text{si } x = 0, 1, \dots, m.$$

Resulta que las tres distribuciones definidas anteriormente son casos particulares de la distribución de series de potencia, definida por:

$$P(X = x; \theta) = \frac{a_x \theta^x}{b(\theta)} \quad \text{para } x = 0, 1, \dots \quad (1.1)$$

en donde $\theta > 0$, $a_x \geq 0$ y $b(\theta) = \sum_{x=0}^{\infty} a_x \theta^x$. Para una muestra de esta distribución, la estadística suficiente minimal es $\sum_{i=1}^n X_i$.

A continuación, damos una breve revisión de estadísticas de prueba existentes para la distribución Poisson.

[Spinelli & Stephens \(1997\)](#) definen en general las estadísticas de prueba de la familia Cramér-von Mises. Estos autores consideran un vector de dimensión infinita $\mathbf{o}^T = (o_0, o_1, \dots, o_j, \dots)$ tal que,

$$o_j = \{ \text{número of } X'_i s = j \}, \quad j = 0, 1, \dots; \quad (1.2)$$

y $\mathbf{e}^T = (e_0, e_1, \dots, e_j, \dots)$ el correspondiente vector esperado (estimado) de \mathbf{o}^T , es decir,

$$e_j = \begin{cases} nF_0(0; \hat{\theta}) & \text{si } j = 0, \\ n[F_0(j; \hat{\theta}) - F_0(j-1; \hat{\theta})] & \text{si } j > 0. \end{cases} \quad (1.3)$$

donde $\hat{\theta}$ es el estimador de máxima verosimilitud de θ . [Spinelli & Stephens \(1997\)](#) definen

$$p_j = \frac{e_j}{n}, \quad Z_j = \sum_{i=0}^j (o_i - e_i), \quad \text{y} \quad H_j = \sum_{i=0}^j p_i \quad \text{where } i, j = 0, 1, \dots \quad (1.4)$$

Las estadísticas de prueba de la familia Cramér-von Mises W^2, A^2 y W_m^2 , respectivamente, son definidas como:

$$W^2 = n^{-1} \sum_{j=0}^{\infty} Z_j^2 p_j, \quad (1.5)$$

$$A^2 = n^{-1} \sum_{j=0}^{\infty} \frac{Z_j^2 p_j}{H_j(1 - H_j)}, \quad (1.6)$$

$$W_m^2 = n^{-1} \sum_{j=0}^{\infty} Z_j^2, \quad (1.7)$$

Observe que en A^2 , p_j da más peso a las colas, lo cual, proporciona mejor potencia contra distribuciones de cola pesada. Nótese que la estadística W_m^2 no depende de m , solo se sigue la notación utilizada en [Spinelli & Stephens \(1997\)](#).

En la práctica, la suma en (1.5), (1.6) y (1.7) debe ser finita, [Spinelli & Stephens \(1997\)](#) proponen que la suma sea truncada por abajo y arriba, para acotar por abajo se elige un número M_l , tal que $o_j = 0$ y $p_j < \frac{10^{-3}}{n}$ para todo $j < M_l$. Similarmente, para acotar por arriba se elige un número M_u , tal que $o_j = 0$ y $p_j < \frac{10^{-3}}{n}$ para todo $j > M_u$. Naturalmente, si se tienen frecuencias muy pequeñas es probable que $p_0 > \frac{10^{-3}}{n}$, por lo que la suma iniciará desde el índice $j = 0$.

Desde otro punto de vista, [Rueda et al. \(1991\)](#) definen una estadística de prueba d , la cual es función de un proceso empírico basado en la función generadora de probabilidades. Posteriormente, en [Rueda & O'Reilly \(1999\)](#), se demuestra que dicha estadística de prueba puede ser expresada como una forma cuadrática:

$$d = \frac{1}{n}(\mathbf{o} - \mathbf{e})^T \mathbf{H}(\mathbf{o} - \mathbf{e}), \quad (1.8)$$

donde \mathbf{H} es la matriz de Hilbert, es decir, $[\mathbf{H}]_{ij} = \frac{1}{i+j-1}$ para $i, j = 1, 2, \dots$; y \mathbf{o} y \mathbf{e} son definidos como en (1.2) y (1.3), respectivamente.

Se tiene que la estadística de prueba d es una suma infinita, en la práctica para calcular esta estadística en [Rueda & O'Reilly \(1999\)](#) proponen que el vector $(\mathbf{o} - \mathbf{e})$ infinito y la matriz de Hilbert infinita sean truncadas a un vector de dimensión K y una matriz cuadrada de tamaño K , respectivamente. Para elegir un tamaño apropiado para K , simplemente se verifica que $o_K = 0$ y $F_0(K, \hat{\theta})$ sea cercano a 1. Por ejemplo un K tal que $o_K = 0$ y $F_0(K, \hat{\theta}) > 1 - \frac{1}{10^6 n}$.

Otra estadística de prueba clásica en bondad de ajuste para la distribución Poisson con λ desconocida, es dada en [Rayner & Best \(1989\)](#), definida como:

$$Z = \frac{(D_n - n)^2}{2n}, \quad (1.9)$$

donde $D_n = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\bar{X}}$ es el índice de dispersión de Fisher.

Usando una caracterización de los momentos para la distribución Poisson, binomial negativa y binomial, otras estadísticas son definidas en [Kyriakoussis et al. \(1998\)](#). Está caracterización es usada posteriormente para probar bondad de ajuste para estas distribuciones. Sea

$$\hat{c} = \frac{\frac{1}{n} \sum_{i=1}^n X_i (X_i - 1)}{\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2},$$

el estimador de momentos de,

$$c = \frac{E[X(X-1)]}{(EX)^2}.$$

Para probar si una muestra proviene de una distribución Poisson (λ desconocida), proponen la siguiente estadística (ver [Kyriakoussis et al. 1998](#)),

$$T_p = \frac{\sqrt{n}(\hat{c} - 1)}{\frac{\sqrt{2}}{\bar{X}}}, \quad (1.10)$$

donde $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Bajo \mathcal{H}_0 , esta estadística converge a una distribución normal con media 0 y varianza 1. Estos resultados están estrechamente relacionados con los que se basan en el índice de dispersión de Fisher. Es fácil ver que $T_p = (D_n - n)/\sqrt{2n}$. Por lo tanto, $Z = T_p^2$, es decir, estas estadísticas son equivalentes. Nosotros usaremos la estadística de prueba Z para el estudio de potencia que se realizará en la sección 1.4.

Para probar si una muestra proviene de una distribución binomial negativa con parámetros m (conocido) y p (desconocido), definen la estadística de prueba (ver [Kyriakoussis et al. 1998](#)),

$$T_{nb} = \frac{\sqrt{n}(\hat{c} - \frac{m+1}{m})}{\hat{\sigma}_{nb}}, \quad (1.11)$$

donde

$$\hat{\sigma}_{nb}^2 = \frac{4\sigma^6 + \mu^2\sigma^2 - 4\mu\sigma^4 - \mu^2\sigma^4 + \mu^2\mu_4 + 2\mu\mu_3(\mu - 2\sigma^2)}{\mu^6}, \quad (1.12)$$

con

$$\begin{aligned} \mu &= \frac{m(1 - \hat{p})}{\hat{p}}, \\ \sigma^2 &= \frac{m(1 - \hat{p})}{\hat{p}^2}, \\ \mu_3 &= \frac{m(2 - \hat{p})(1 - \hat{p})}{\hat{p}^3}, \\ \mu_4 &= \frac{m(1 - \hat{p})}{\hat{p}^4} [1 + 4(1 - \hat{p}) + (1 - \hat{p})^2 + 3m(1 - \hat{p})], \text{ donde } \hat{p} = \frac{m}{m + \bar{X}}. \end{aligned}$$

Bajo \mathcal{H}_0 , esta estadística converge a una distribución normal con media 0 y varianza 1.

Finalmente, para probar si una muestra proviene de una distribución binomial con parámetros m (conocido) y p (desconocido), proponen la siguiente estadística de prueba (ver [Kyriakoussis et al. 1998](#)),

$$T_b = \frac{\sqrt{n}(\hat{c} - \frac{m-1}{m})}{\hat{\sigma}_b}, \quad (1.13)$$

donde $\hat{\sigma}_b$ es determinado por la ecuación (1.12), con

$$\begin{aligned}\mu &= m\hat{p}, \\ \sigma^2 &= m\hat{p}(1 - \hat{p}), \\ \mu_3 &= m\hat{p}(1 - \hat{p})(1 - 2\hat{p}), \\ \mu_4 &= m\hat{p}(1 - \hat{p})[3\hat{p}^2(2 - m) + 3\hat{p}(m - 2) + 1], \quad \text{donde } \hat{p} = \frac{\bar{X}}{m}.\end{aligned}$$

De igual forma, bajo \mathcal{H}_0 , esta estadística converge a una distribución normal con media 0 y varianza 1.

Por otra parte, en teoría de la información, la divergencia Kullback-Leibler entre dos densidades discretas P y Q (relativa a P) es definida como:

$$KL = \sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right),$$

esta expresión es bien definida si el índice i está en el soporte de Q y el soporte de P es contenido en el soporte de Q ; que es el caso cuando se toma a P como la densidad empírica y Q como la densidad teórica (es decir bajo \mathcal{H}_0).

Con un simple cálculo se puede ver que cuando P es la densidad empírica y Q es la densidad teórica se tiene que $KL = -\frac{1}{n} \log(\Lambda)$, donde Λ es la estadística de cociente (o razón) de verosimilitudes generalizadas, es decir:

$$\Lambda = \prod_{j \in J} \left\{ \frac{Q(j)}{\frac{o_j}{n}} \right\}^{o_j},$$

siendo $Q(j) = P(X = j; \hat{\theta})$ (la probabilidad bajo \mathcal{H}_0), $\hat{\theta}$ el estimador de máxima verosimilitud y J el conjunto de los j 's observados en la muestra. Por lo tanto, la estadística Kullback-Leibler y la estadística de razón de verosimilitudes generalizadas son equivalentes.

Recientemente, la distribución condicional (de la familia de distribuciones a probar) dada una estadística suficiente ha sido utilizada en las pruebas de bondad de ajuste, este enfoque ha tenido importancia debido a que estas pruebas condicionales permiten trabajar con la distribución condicional, la cual es independiente del parámetro (desconocido). El primer autor, del que se tiene conocimiento, que usó la probabilidad condicional dada una estadística suficiente para bondad de ajuste fue

Fisher (1950). Dicho autor utilizó la probabilidad condicional para probar bondad de ajuste para el caso de la distribución Poisson. Recientemente, González-Barrios et al. (2006) proponen un procedimiento general para probar bondad de ajuste usando la probabilidad condicional dada una estadística suficiente para las distribuciones de series de potencia (incluye la Poisson, binomial negativa y binomial). Para estas distribuciones, la estadística suficiente es $T_n = \sum_{i=1}^n X_i$, y la prueba de Fisher es reconstruida para la distribución Poisson. La probabilidad condicional usada es $P(X_1 = x_1, \dots, X_n = x_n | T_n = t_n)$ y será denotada por $\tilde{f}_n(x_1, x_2, \dots, x_n | t_n)$. El procedimiento usado en González-Barrios et al. (2006) para obtener los p -valores condicionales es equivalente a una enumeración de la probabilidad condicional de las posibles “muestras” dada la suma; un problema interesante de combinatoria.

La probabilidad condicional para la distribución Poisson con parámetro $\lambda > 0$ es (González-Barrios et al. 2006),

$$\begin{aligned} \tilde{f}_n(x_1, x_2, \dots, x_n | t_n) &= \binom{t_n}{x_n} \binom{t_n - x_n}{x_{n-1}} \binom{t_n - x_n - x_{n-1}}{x_{n-2}} \dots \binom{t_n - x_n - \dots - x_2}{x_1} \left(\frac{1}{n}\right)^{t_n} \\ &= \frac{t_n!}{x_1! \dots x_n!} \left(\frac{1}{n}\right)^{t_n}, \end{aligned}$$

la cual corresponde a la distribución multinomial con parámetros $p_1 = p_2 = \dots = p_n = \frac{1}{n}$.

La probabilidad condicional para la distribución binomial negativa con parámetros $r \geq 1$ y $0 < p < 1$, con r conocido, es

$$\begin{aligned} \tilde{f}_n(x_1, x_2, \dots, x_n | t_n) &= \frac{\binom{r + x_n - 1}{r - 1} \binom{r + x_{n-1} - 1}{r - 1} \dots \binom{r + x_1 - 1}{r - 1}}{\binom{nr + t_n - 1}{nr - 1}}. \end{aligned}$$

Por último, la correspondiente probabilidad condicional para la distribución binomial con parámetros $m \geq 1$ y $0 < p < 1$, con m conocido, es

$$\tilde{f}_n(x_1, x_2, \dots, x_n | t_n) = \frac{\binom{m}{x_n} \binom{m}{x_{n-1}} \dots \binom{m}{x_1}}{\binom{nm}{t_n}},$$

la cual corresponde a la versión multivariada de la distribución hipergeométrica, con n clases de tamaño m cada una.

A continuación para ilustrar el procedimiento usado en [González-Barríos et al. \(2006\)](#), se ejemplifica en el caso que la distribución hipotética sea la distribución Poisson. Para la distribución binomial y binomial negativa, el procedimiento será análogo. Supongamos que tenemos una muestra aleatoria (X_1, \dots, X_n) y se quiere probar:

$$\begin{aligned} \mathcal{H}_0 : X &\sim \text{Poisson}(\lambda) \text{ para algún } \lambda > 0 \text{ vs.} \\ \mathcal{H}_1 : X &\not\sim \text{Poisson}(\lambda) \text{ para ningún } \lambda > 0, \end{aligned} \quad (1.14)$$

con un nivel α . Como se mencionó, bajo \mathcal{H}_0 , la distribución condicional de (X_1, \dots, X_n) dada $T_n = t_n$ es una distribución multinomial con parámetro t_n y $p_1 = \dots = p_n = \frac{1}{n}$. Tomando los elementos del citado artículo de [González-Barríos et al. \(2006\)](#) se reproducen a continuación las ideas y ejemplos básicos.

Se dice que un vector \mathbf{Y} es un arreglo de (t_n, n) , si \mathbf{Y} es un vector de tamaño n , $\mathbf{Y} = (y_n, y_{n-1}, \dots, y_2, y_1)$ y cumple que $y_n \geq y_{n-1} \geq \dots \geq y_2 \geq y_1$ con $\sum_{i=1}^n y_i = t_n$. La probabilidad condicional de este arreglo \mathbf{Y} de (t_n, n) bajo \mathcal{H}_0 es,

$$\begin{aligned} P(X_n = y_n, X_{n-1} = y_{n-1}, \dots, X_2 = y_2, X_1 = y_1 | T_n = t_n) \\ = \binom{t_n}{y_n} \binom{t_n - y_n}{y_{n-1}} \binom{t_n - y_n - y_{n-1}}{y_{n-2}} \dots \binom{t_n - y_n - \dots - y_2}{y_1} \left(\frac{1}{n}\right)^{t_n}. \end{aligned}$$

La prueba consiste en encontrar todos los arreglos de (t_n, n) y calcular su respectiva probabilidad condicional, luego la probabilidad condicional de cualquier permutación de un arreglo será la misma. Por ejemplo suponga que $n = 10$ y $t_n = 5$, luego

la probabilidad condicional del arreglo $(2, 2, 1, 0, 0, 0, 0, 0, 0, 0)$ está dada por

$$\begin{aligned} P((2, 2, 1, 0, 0, 0, 0, 0, 0, 0) | T_{10} = 5) \\ &= \binom{10}{2} \binom{8}{1} \binom{7}{7} \binom{5}{2} \binom{3}{2} \binom{1}{1} \left(\frac{1}{10}\right)^5 \\ &= .108, \end{aligned}$$

en donde los primeros tres coeficientes binomiales corresponden a permutaciones del arreglo.

Una vez que se encuentra todos los arreglos de (t_n, n) , se procede a ordenarlos en forma decreciente, según su probabilidad condicional; de la más alta a la más baja. Al evaluar la probabilidad acumulada de los arreglos, se encuentra aquel arreglo con probabilidad acumulada más cercana a $1 - \alpha$. La región crítica (región de rechazo), la forman los arreglos subsecuentes cuya probabilidad es aproximadamente α .

Para ilustrar lo anterior, suponga \mathcal{H}_0 cierta, es decir, $X \sim Poisson(\lambda)$ para algún λ y sea $n = 10$ y con $t_n = 8$. En la tabla 1.1 se reportan todos los arreglos de (t_n, n) en orden decreciente, según sus respectivas probabilidades condicionales, al igual que sus probabilidades acumuladas. Si se quisiera probar (1.14), con un $\alpha \approx 0.1$. De la tabla (1.1) se ve que la región crítica es el complemento del conjunto de arreglos

$$\begin{aligned} R = \{ &(2, 2, 1, 1, 1, 1, 0, 0, 0, 0), (3, 2, 1, 1, 1, 0, 0, 0, 0, 0) \\ &(2, 1, 1, 1, 1, 1, 1, 0, 0, 0), (2, 2, 2, 1, 1, 0, 0, 0, 0, 0), \\ &(3, 1, 1, 1, 1, 1, 0, 0, 0, 0), (3, 2, 2, 1, 0, 0, 0, 0, 0, 0)\}. \end{aligned}$$

El nivel condicional actual de la prueba es $\alpha_{T_{10}=8} = 0.089776$. Si por ejemplo, se observa el arreglo $(3, 3, 1, 1, 0, 0, 0, 0, 0, 0)$, se rechaza \mathcal{H}_0 con el referido $\alpha_{T_{10}=8} = 0.089776$. Para los casos de bondad de ajuste de la distribución binomial negativa o binomial, el procedimiento a seguir es análogo a lo expuesto anteriormente. En lo que sigue, al procedimiento de para probar bondad ajuste visto en [González-Barrios et al. \(2006\)](#), donde se utiliza la probabilidad condicional \tilde{f}_n , nos referiremos indistintamente como estadística de prueba \tilde{f}_n ó uso de \tilde{f}_n .

TABLA 1.1: Arreglos y sus probabilidades para $T_{10} = 8$ en el caso Poisson.

Arreglo	Probabilidad bajo \mathcal{H}_0	Probabilidad acumulada
(2,2,1,1,1,1,0,0,0,0)	0.31752	0.31752
(3,2,1,1,1,0,0,0,0,0)	0.169344	0.486864
(2,1,1,1,1,1,1,0,0,0)	0.169344	0.656208
(2,2,2,1,1,0,0,0,0,0)	0.127008	0.783216
(3,1,1,1,1,1,0,0,0,0)	0.084672	0.867888
(3,2,2,1,0,0,0,0,0,0)	0.042336	0.910224
(4,1,1,1,1,0,0,0,0,0)	0.021168	0.931392
(4,2,1,1,0,0,0,0,0,0)	0.021168	0.952560
(1,1,1,1,1,1,1,1,0,0)	0.018144	0.970704
(3,3,1,1,0,0,0,0,0,0)	0.014112	0.984816
(2,2,2,2,0,0,0,0,0,0)	0.005292	0.990108
(5,1,1,1,0,0,0,0,0,0)	0.0028224	0.9929304
(3,3,2,0,0,0,0,0,0,0)	0.002016	0.9949464
(4,3,1,0,0,0,0,0,0,0)	0.002016	0.9969624
(4,2,2,0,0,0,0,0,0,0)	0.001512	0.9984744
(5,2,1,0,0,0,0,0,0,0)	0.0012096	0.9996840
(6,1,1,0,0,0,0,0,0,0)	0.0002016	0.9998856
(5,3,0,0,0,0,0,0,0,0)	0.0000504	0.999936
(4,4,0,0,0,0,0,0,0,0)	0.0000315	0.9999675
(6,2,0,0,0,0,0,0,0,0)	0.0000252	0.9999927
(7,1,0,0,0,0,0,0,0,0)	0.0000072	0.9999999
(8,0,0,0,0,0,0,0,0,0)	0.0000001	1.0000000

En la sección 1.2 proponemos dos nuevas estadísticas de prueba. En la sección 1.3 se da el procedimiento para generar muestras condicionalmente independientes. Mediante el uso de las muestras condicionales, en la sección 1.4 se realiza un estudio de simulación para comparar la potencia de las nuevas estadísticas de prueba con las estadísticas de prueba mencionadas anteriormente para la distribución Poisson, binomial negativa y binomial, respectivamente.

1.2. Nuevas estadísticas de prueba

En esta sección, se proponen dos estadísticas de prueba de bondad de ajuste para la distribución Poisson. Una es una variación de la prueba basada en la razón de verosimilitud originalmente usada en el caso continuo, en Zhang (2002). La otra es una modificación de la prueba basada en la probabilidad condicional para las distribuciones de series de potencia visto en González-Barrios et al. (2006),

completando lo que puede ser una prueba basada en la razón de probabilidades condicionales, que podría ser algo similar a la idea basada en la prueba de razón de verosimilitudes generalizadas.

En [Zhang \(2002\)](#), se propone una estadística de prueba Z , definida por,

$$Z = \int_{-\infty}^{\infty} G_t^2 dw(t), \quad (1.15)$$

con

$$G_t^2 = 2n \left[F_n(t) \log \frac{F_n(t)}{F_0(t)} + \{1 - F_n(t)\} \log \frac{1 - F_n(t)}{1 - F_0(t)} \right],$$

y $w(t)$ alguna función de peso. Donde $F_n(t)$ es la función de distribución empírica de la muestra original X_1, \dots, X_n y F_0 es la función de distribución bajo \mathcal{H}_0 . Este autor comenta que una familia grande que incluye a G_t^2 fue dada en [Cressie & Read \(1984\)](#). En [Zhang \(2002\)](#) en el caso de probar normalidad, se concluye que la estadística de prueba Z es más potente que las estadísticas de prueba tradicionales (Kolmogorov-Smirnov, Shapiro-Wilk, Cramér-von Mises y Anderson-Darling).

Más recientemente, en [Potas et al. \(2014\)](#) se define una estadística de prueba para la bondad de ajuste de la distribución normal sesgada, la cual, está basada en la razón de verosimilitudes (ecuación 1.15). En este trabajo proponemos una estadística de prueba basada en G_t^2 para la bondad de ajuste de la distribución Poisson, adaptándola al caso discreto. Sean o_j y p_j definidas como (1.2) y (1.4), respectivamente; y sea $v = \max\{x_1, x_2, \dots, x_n\}$ el máximo de la muestra observada, $l = \min\{j | o_j > 0\}$, $L_k = \sum_{j=0}^k o_j$ y $S_k = \sum_{j=0}^k p_j$. Definimos la estadística de prueba,

$$B = 2n \left\{ \sum_{k=l}^{v+2n} \left(\frac{L_k}{n} \right) \log \left[\frac{\frac{L_k}{n}}{S_k} \right] + \sum_{k=0}^{v-1} \left(1 - \frac{L_k}{n} \right) \log \left[\frac{1 - \frac{L_k}{n}}{1 - S_k} \right] \right\}.$$

Por otra parte, en cuanto a la variación del uso de la probabilidad condicional de la muestra observada para definir un criterio de rechazo tal como se utiliza en [González-Barríos et al. \(2006\)](#), es decir, después de enumerar las probabilidades condicionales de todas las posibles muestras (configuraciones) y eligiendo un nivel de significancia α , por ejemplo eligiendo $\alpha = 0.05$, se rechaza si la probabilidad condicional de la muestra observada está entre las 5% de las menos probables. La variación que se introduce aquí es un cociente, que denotaremos por Δ . El numerador del cociente es la probabilidad condicional de la muestra observada

dada la suma t_n , bajo la hipótesis de que la distribución es la Poisson. En el denominador, aparece la probabilidad condicional de la muestra observada, dada las estadísticas de orden; esto es dada la estadística suficiente minimal asumiendo que la distribución podría ser discreta sobre el rango $\{0, 1, \dots\}$.

La prueba de razón de probabilidades condicionales resultante, para el caso Poisson es:

$$\Delta = \frac{\frac{t_n!}{\prod_j (j!)^{o_j}} \left(\frac{1}{n}\right)^{t_n}}{\frac{\prod_j (o_j)!}{n!}}.$$

Por supuesto que si uno quiere probar bondad de ajuste de otra distribución discreta en el mismo rango de $\{0, 1, \dots\}$, por ejemplo, para la binomial o binomial negativa, como en [González-Barrios et al. \(2006\)](#), la correspondiente estadística de prueba puede ser construida simplemente cambiando la expresión en el numerador de Δ por la probabilidad condicional de la distribución correspondiente dada la estadística suficiente minimal.

La distribución condicional dada la estadística suficiente de la estadística de prueba, en el caso de la Poisson, bajo la hipótesis nula puede aproximarse con mucha precisión mediante la simulación de muchas muestras condicionalmente independientes dado t_n como se ve en la siguiente sección. Esta posibilidad de simular muestras condicionalmente independientes se ilustra en las otras dos distribuciones, es decir, en la binomial negativa y la binomial.

1.3. Muestras condicionales

Una posibilidad, si se quiere simular de la distribución conjunta de $(X_1, \dots, X_n)|T_n$, esto se puede hacer mediante la simulación de la distribución conjunta de (Z_1, \dots, Z_n) , donde $Z_n = X_n|T_n$, $Z_{n-1} = X_{n-1}|T_n, X_n$, $Z_{n-2} = X_{n-2}|T_n, X_n, X_{n-1}$, \dots , $Z_1 = X_1|T_n, X_n, X_{n-1}, \dots, X_2$.

[O'Reilly & Gracia-Medrano \(2006\)](#) utilizando el estimador Rao-Blackwell dada una estadística suficiente T_n (dada la muestra) y bajo algunos supuestos, se generan muestras condicionales, las cuales tendrán la misma distribución condicional que la muestra inicial (si la hipótesis nula es verdadera), y esto se debe a que las distribuciones condicionales de Z_r son iguales a $X_r|T_r$ con $r = 1, \dots, n$. A estas muestras simuladas los autores las refieren como muestras “look-alike” ó “*-muestra”.

A continuación, se dan las definiciones necesarias y el procedimiento para la generación de las llamadas muestras condicionales o muestras look-alike.

Una pregunta muy natural es, ¿Qué se obtiene si en lugar de estimar $F_0(\cdot; \theta)$ con $F_0(\cdot; \hat{\theta}_n)$, se utiliza otro tipo de estimador? Existe una teoría que proporciona un estimador insesgado y de varianza mínima de $F_0(\cdot; \theta)$, que se deduce a partir de la teoría de Rao-Blackwell, a dicho estimador se le conoce con el mismo nombre y se enuncia enseguida.

Sea X_1, \dots, X_n una muestra aleatoria de la función de distribución

$$F_0(\cdot; \theta); \theta \in \Theta_0 \quad \text{y sea } T_n = T_n(X_1, \dots, X_n),$$

una estadística suficiente minimal y completa, se define el estimador Rao-Blackwell de $F_0(x; \theta)$ basado en $T_n = t_n$ como:

$$\tilde{G}_n(x|t_n) := E[I_{[X_1 \leq x]} | T_n = t_n] = P[X_1 \leq x | T_n = t_n].$$

El estimador Rao-Blackwell es un estimador insesgado y de varianza mínima de $F_0(\cdot; \theta)$. A esta función también se le conoce como distribución Rao-Blackwell. En muestreo independiente T_n es simétrica, es decir, si para cualquier permutación π en $\{1, \dots, n\}$, se tiene que $T_n(X_1, \dots, X_n) = T_n(X_{\pi(1)}, \dots, X_{\pi(n)})$, por lo que las X_i dada T_n son intercambiables (ver O'Reilly 1971),

$$\tilde{G}_n(x|t_n) = P[X_i \leq x | T_n = t_n],$$

con $i = 1, \dots, n$.

La estadística T_n cumple la propiedad de doblemente transitividad, esto es, si el par (T_n, X_n) está en una correspondencia uno a uno con el par (T_{n-1}, X_n) . Lo cual significa que el conocimiento de T_n y X_n es equivalente al conocimiento de T_{n-1} y X_n . Este es el caso de las familias que pertenecen a la familia exponencial natural donde T_n es la suma de n elementos independientes e idénticamente distribuidos, es decir, $T_n = \sum_{i=1}^n X_i$. Por lo que en este caso, el conocimiento de $T_n = \sum_{i=1}^n X_i$ y X_n es equivalente al conocimiento de $T_{n-1} = \sum_{i=1}^{n-1} X_i$ y X_n . En muchas aplicaciones se tiene que la estadística T_n cumple la propiedad de doble transitividad.

Si la estadística T_n es doblemente transitiva y se quiere simular la muestra x_1^*, \dots, x_n^* de la distribución condicional de X_1, \dots, X_n dada $T_n = t_n$ (o lo que es lo mismo, dada T_n , en t_n), sólo se necesita la expresión $\tilde{G}_n(x|t_n)$.

Sea k el número de parámetros desconocidos de la familia paramétrica $F_0(\cdot; \theta)$, el siguiente resultado dice cómo simular las primeras $n - k$ componentes de la $*$ -muestra, y para obtener los restantes k términos se resuelve un sistema de k incógnitas.

Bajo doble transitividad, teniendo $n - k$ como el número máximo de términos de la muestra, para la distribución condicional de esos términos dada $T_n = t_n$, el procedimiento para obtener (por ejemplo los superiores) $n - k$ términos de la x^* -muestra (look-alike) que tengan exactamente la misma distribución condicional (como cualesquiera $n - k$ términos de la muestra), es la siguiente:

Dada $\tilde{G}_n(x|t_n)$, el estimador Rao-Blackwell basado en $t_n = T_n(x_1, \dots, x_n)$, se denota por $\tilde{G}_n^{-1}(u)$ su inversa.

Paso 1. Sea u_n una realización de una variable aleatoria $U(0, 1)$, se define $x_n^* = \tilde{G}_n^{-1}(u_n)$, y usando la doble transitividad, se calcula t_{n-1}^* de t_n y x_n^* , y se denota este nuevo valor por t_{n-1}^* .

Paso 2. Ahora con \tilde{G}_{n-1} , el estimador Rao-Blackwell basado en t_{n-1}^* y la obvia notación para su inversa, sea u_{n-1} otra realización de una variable aleatoria $U(0, 1)$ independiente de u_n ; se define el número $x_{n-1}^* = \tilde{G}_{n-1}^{-1}(u_{n-1})$ y del conocimiento de t_{n-1}^* y x_{n-1}^* , se calcula t_{n-2}^* ; se sigue el proceso hasta que:

⋮

Paso $(n - k)$. Usando \tilde{G}_{k+1}^{-1} generar x_{k+1}^* con u_{k+1} , seleccionando a u_{k+1} como una realización de una variable aleatoria $U(0, 1)$, independiente de las anteriores u_j 's.

Estos $n - k$ términos de la $*$ -muestra son una realización de la distribución condicional para cualesquiera $n - k$ términos de la muestra dada t_n .

Finalmente, se encuentran los restantes k términos x_1^*, \dots, x_k^* que pueden ser cualquier permutación de la única solución (ordenada) generada por la condición de que $T_n(x_1^*, \dots, x_n^*)$ sea igual a t_n . A la muestra obtenida x_1^*, \dots, x_n^* se le llamará muestra look-alike ó $*$ -muestra.

En caso discreto la obtención de las distribuciones condicionales es más directo. A continuación se ilustra como se obtienen las distribuciones condicionales en el caso discreto. Para obtener estas muestras condicionales (o look-alike), se inicia

con una muestra aleatoria X_1, \dots, X_n , asumiendo que \mathcal{H}_0 es cierta, y $T_n = \sum_{i=1}^n X_i$ es una estadística suficiente doblemente transitiva, por lo tanto, se obtiene que,

$$\begin{aligned}
 P(X_n = x_n | T_n = t_n) &= \frac{P(X_n = x_n, T_n = t_n; \theta)}{P(T_n = t_n; \theta)} \\
 &= \frac{P(X_n = x_n, T_{n-1} = t_n - x_n; \theta)}{P(T_n = t_n; \theta)} \\
 &= \frac{P(X_n = x_n; \theta)P(T_{n-1} = t_n - x_n; \theta)}{P(T_n = t_n; \theta)}. \quad (1.16)
 \end{aligned}$$

Similarmente, aplicando doble transitividad e independencia, se obtiene que la probabilidad condicional,

$$\begin{aligned}
 &P(X_{n-1} = x_{n-1} | T_n = t_n, X_n = x_n) \\
 &= \frac{P(X_{n-1} = x_{n-1}, T_n = t_n, X_n = x_n; \theta)}{P(T_n = t_n, X_n = x_n; \theta)} \\
 &= \frac{P(X_{n-1} = x_{n-1}, T_{n-1} = t_n - x_n, X_n = x_n; \theta)}{P(T_{n-1} = t_n - x_n, X_n = x_n; \theta)} \\
 &= \frac{P(X_{n-1} = x_{n-1}, T_{n-2} = t_n - x_n - x_{n-1}, X_n = x_n; \theta)}{P(T_{n-1} = t_n - x_n, X_n = x_n; \theta)} \\
 &= \frac{P(X_{n-1} = x_{n-1}, T_{n-2} = t_n - x_n - x_{n-1}; \theta)P(X_n = x_n; \theta)}{P(T_{n-1} = t_n - x_n; \theta)P(X_n = x_n; \theta)} \\
 &= \frac{P(X_{n-1} = x_{n-1}, T_{n-1} = t_n - x_n; \theta)}{P(T_{n-1} = t_n - x_n; \theta)} \\
 &= P(X_{n-1} = x_{n-1} | T_{n-1} = t_n - x_n).
 \end{aligned}$$

Es decir,

$$P(X_{n-1} = x_{n-1} | T_n = t_n, X_n = x_n) = P(X_{n-1} = x_{n-1} | T_{n-1} = t_n - x_n).$$

Análogamente, se obtiene,

$$\begin{aligned}
 &P(X_{n-2} = x_{n-2} | T_n = t_n, X_n = x_n, X_{n-1} = x_{n-1}) \\
 &= P(X_{n-2} = x_{n-2} | T_{n-2} = t_n - x_n - x_{n-1}) \\
 &\quad \vdots \\
 &P(X_1 = x_1 | T_n = t_n, X_n = x_n, \dots, X_2 = x_2) = P\left(X_1 = x_1 | T_1 = t_n - \sum_{i=2}^n x_i\right).
 \end{aligned}$$

Es decir, se tiene que,

$$\begin{aligned} P(X_{n-1} = x_{n-1} | T_n = t_n, X_n = x_n) &= P(X_{n-1} = x_{n-1} | T_{n-1} = t_{n-1}), \\ P(X_{n-2} = x_{n-2} | T_n = t_n, X_n = x_n, X_{n-1} = x_{n-1}) &= P(X_{n-2} = x_{n-2} | T_{n-2} = t_{n-2}), \\ &\vdots \\ P(X_1 = x_1 | T_n = t_n, X_n = x_n, \dots, X_2 = x_2) &= P(X_1 = x_1 | T_1 = t_1). \end{aligned}$$

Observe que la última probabilidad es igual a 0 ó a 1, es decir,

$$P(X_1 = x_1 | T_1 = t_1) = \begin{cases} 0 & \text{si } x_1 \neq t_1, \\ 1 & \text{si } x_1 = t_1. \end{cases}$$

El procedimiento para calcular el p -valor condicional mediante la generación de muestras look-alike es el siguiente: Sea $D(; \theta)$ una estadística de prueba, dada la muestra \underline{x} de tamaño n se calcula el estimador $\hat{\theta}$, luego se calcula $D(\underline{x}; \hat{\theta})$; se genera un número grande de muestras look-alike bajo \mathcal{H}_0 , digamos M muestras de tamaño n cada una, por lo que tendremos las muestras $\underline{y}_1, \dots, \underline{y}_M$; para cada muestra simulada se calcula el valor de la estadística de prueba D . Tendremos M observaciones simuladas de la estadística de prueba; es decir, $D(\underline{y}_1; \hat{\theta}), \dots, D(\underline{y}_M; \hat{\theta})$. El p -valor condicional, es la proporción de valores D_i 's que exceden a D , es decir

$$p \approx \frac{1}{M} \sum_{j=1}^M I\{D(\underline{y}_j; \hat{\theta}) > D(\underline{x}; \hat{\theta})\},$$

donde I es la función indicadora. Entre más grande sea M , mejor será la aproximación del p -valor condicional. Si se elige un nivel de significancia (nivel de rechazo) α , el criterio de rechazo será el siguiente:

- si $p < \alpha$, se rechaza \mathcal{H}_0 ;
- si $p \geq \alpha$, no se rechaza \mathcal{H}_0 .

1.3.1. Distribución Rao-Blackwell

La distribución Rao-Blackwell para la distribución Poisson resulta ser

$$\tilde{G}_n(x'_n | t_n) = \sum_{j=0}^{x'_n} \text{pbin}(j; t_n, \frac{1}{n}),$$

donde $pbm(j; t_n, \frac{1}{n})$ es la probabilidad de que una variable aleatoria binomial con parámetros t_n y $\frac{1}{n}$ tome el valor j .

Si se quiere simular una muestra look alike x_1^*, \dots, x_n^* , se procede como sigue. Primero, simulamos una variable aleatoria x_n^* de $Bin(t_n, 1/n)$, luego, se calcula $t_{n-1} = t_n - x_n^*$ y simulamos otra variable aleatoria x_{n-1}^* de $Bin(t_{n-1}, 1/(n-1))$, recalculamos $t_{n-2} = t_{n-1} - x_{n-1}^*$ y simulamos una variable aleatoria x_{n-2}^* de $Bin(t_{n-2}, 1/(n-2))$ etc. Finalmente, $x_1^* = t_1$ y esto termina la simulación.

La distribución Rao-Blackwell para la distribución binomial con m conocida es

$$\tilde{G}_n(x'_n | t_n) = \sum_{j=0}^{x'_n} \frac{\binom{m}{j} \binom{nm-m}{t_n-j}}{\binom{nm}{t_n}},$$

donde $\frac{\binom{m}{j} \binom{nm-m}{t_n-j}}{\binom{nm}{t_n}}$ es la probabilidad de que una variable aleatoria hipergeométrica con parámetros t_n , k y nk tome el valor j .

El procedimiento para la generación de una muestra condicional x_1^*, \dots, x_n^* de una distribución binomial es análogo del caso de la distribución Poisson.

Finalmente, la distribución Rao-Blackwell para la distribución binomial negativa con r conocida es

$$\tilde{G}_n(x'_n | t_n) = \sum_{j=0}^{x'_n} \frac{\binom{r+j-1}{j} \binom{(n-1)r+t_n-j-1}{t_n-j}}{\binom{nr+t_n-1}{t_n}}.$$

esta última distribución no es “muy conocida” (como lo es por ejemplo la distribución Poisson, binomial negativa, binomial, etc.), por lo que no está implementada en un paquete estadístico, por lo tanto, si queremos generar una muestra de esta distribución, debemos utilizar, por ejemplo, el método de la transformada inversa.

En la sección 1.4 se realiza un estudio de simulación para determinar la potencia de las diferentes estadísticas de prueba usando la distribución Rao-Blackwell para la generación de las muestras condicionales para las distribuciones Poisson, binomial negativa y binomial.

1.4. Estudio de potencia

La potencia de una estadística de prueba en bondad de ajuste tiene un papel muy importante similar a lo que el error cuadrático medio tiene en la teoría de la estimación. En las pruebas de hipótesis paramétricas, una estadística prueba Υ^* es uniformemente más potente con nivel de significancia α , si tiene un nivel de significancia α y entre todas las pruebas con nivel de significancia menor o igual a α es la que tiene mayor potencia para todos los valores alternativos de θ . En el caso de la familia exponencial, existe una estadística prueba Υ^* uniformemente más potente con nivel de significancia α (Lema de Neyman-Pearson). Desafortunadamente para el caso de bondad de ajuste (prueba de hipótesis no paramétrica) no existe un resultado similar donde se diga que una estadística de prueba es uniformemente más potente con nivel de significancia α . Generalmente, la potencia de una estadística de prueba depende de la familia paramétrica a probar. Es habitual, que para un amplio conjunto de distribuciones alternativas, no habrá una mejor prueba, es decir, con potencia más alta. Por lo que, se dirá que una estadística de prueba tiene una potencia aceptable o es una buena prueba, si la potencia de la estadística de prueba es muy parecida o superior a otra estadística de prueba ya reconocida en la literatura como una prueba recomendable.

Sea D una estadística de prueba para probar \mathcal{H}_0 . Dado un nivel de significancia α , la potencia de la estadística de prueba D de la distribución F_1 es la probabilidad de rechazar \mathcal{H}_0 dado que la verdadera distribución es F_1 , y se denota por $\Pi_D(F_1)$. El procedimiento para calcular la potencia de la estadística de prueba D de la función de distribución F_1 es la siguiente:

1. Dado el tamaño de la muestra n , se simula una muestra de tamaño n de la función de distribución F_1 .
2. Bajo \mathcal{H}_0 , se calcula el p -valor de la estadística de prueba D , se decide si se rechaza ó no la hipótesis nula.
3. Hacer por los pasos 1 y 2 un número grande de veces, se calcula la potencia de la estadística de prueba D de la distribución F_1 como el porcentaje de rechazos.

A la distribución F_1 también se le refiere como la distribución alternativa, es decir, a $\Pi_D(F_1)$ también se le llama la potencia de la estadística de prueba D cuando la alternativa es la distribución F_1 .

En esta sección se realiza un estudio de simulación para comparar la potencia de las estadísticas de prueba vistas anteriormente. Se simulan 1000 muestras de tamaño $n = 10, 20, 30, 40$ y 50 para diferentes distribuciones alternativas; para cada una de estas muestras, se generan 1000 muestras look-alike para la distribución Poisson, binomial negativa y binomial, respectivamente. Para ver si se rechaza \mathcal{H}_0 se eligió un nivel de significancia $\alpha = 0.10$ y $\alpha = 0.05$, respectivamente.

Las alternativas usadas en este estudio de simulación se tomaron de entre la distribución Poisson, binomial negativa, binomial, beta-binomial y uniforme discreta, denotadas, $Poisson(\lambda)$, $NB(r, p)$, $Bin(m, p)$, $BetaBin(m, \alpha, \lambda)$, y $Unif(a, b)$, respectivamente. La distribución beta binomial proviene de una composición de una distribución binomial con parámetros m , y p , donde p tiene una distribución beta con parámetros α y β . También, se utilizó como alternativa la distribución binomial negativa generalizada que recordamos a continuación.

Se dice que X tiene una distribución binomial negativa generalizada con parámetros $k > 0$, $b \geq 1$ y $0 \leq a \leq \frac{1}{b}$, denotado por $X \sim GNB(k, a, b)$ si y solo si (ver [Consult & Gupta 1980](#)),

$$P(X = x) = \frac{k}{k + bx} \binom{k + bx}{x} a^x (1 - a)^{k+bx-x} \quad \text{if } x = 0, 1, \dots$$

En el caso que $b = 1$, la distribución es la binomial negativa.

Por otra parte, para la estimación de los parámetros desconocidos usamos el estimador de máxima verosimilitud, por lo que, para la distribución Poisson su estimador es $\hat{\lambda} = \bar{x}$, para la binomial negativa es $\hat{p} = \frac{r}{r + \bar{x}}$, y finalmente, para la distribución binomial es $\hat{p} = \frac{\bar{x}}{m}$, con $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ la media muestral. En la sección [1.7](#), se pueden ver las tablas de este estudio de simulación.

1.4.1. Distribución Poisson

Las tablas [1.8-1.11](#) reportan la potencia de las pruebas para la distribución Poisson (λ desconocido) cuando la alternativa es binomial negativa, binomial, beta-binomial y uniforme discreta, respectivamente.

Se replicaron algunos escenarios usados en [Spinelli & Stephens \(1997\)](#) y posteriormente usados en [Frey \(2012\)](#). También replicamos algunos escenarios de [González-Barrios et al. \(2006\)](#), [Best & Rayner \(1999\)](#) y [Gurtler & Henze \(2000\)](#).

Como era de esperarse, los resultados de [Frey \(2012\)](#) para la potencia de W^2 y A^2 son muy similares a los reportados en este estudio (ver Tablas 1.8-1.10), ya se utiliza las distribuciones condicionales de las estadísticas W^2 y A^2 . También, los resultados en [Gurtler & Henze \(2000\)](#) para las estadísticas d y Z son muy similares a lo que reportamos (ver Tablas 1.8, 1.9 y 1.11), sin embargo el enfoque utilizado en [Gurtler & Henze \(2000\)](#) para calcular la potencia fue por medio de remuestreo paramétrico.

Es interesante que en general, las estadísticas W^2 , A^2 , W_m^2 se comportan razonablemente bien, estando entre las mejores si la alternativa es beta binomial (Tabla 1.10). La prueba nueva B , nos sorprendió por tener muy buena potencia contra las tres familias de las alternativas, excepto en la binomial negativa que fue superada por Z . Las pruebas KL y d no fueron tan buenas en general. La prueba \tilde{f}_n es buena contra la binomial negativa pero tiene muy mala potencia contra la binomial, y en la mayoría de los casos tiene potencia inferior al nivel de significancia (Tabla 1.9). La prueba Δ no fue tan buena como las estadísticas de prueba de la familia Cramér-von Mises.

1.4.2. Distribución binomial negativa

Las tablas 1.12-1.16 reportan la potencia de las diferentes estadísticas de prueba para la distribución binomial negativa (r conocido y p desconocido) cuando las alternativas son la Poisson, binomial, beta-binomial, uniforme discreta y binomial negativa generalizada, respectivamente.

[Best & Rayner \(2003\)](#) realizan un estudio de potencia para probar bondad de ajuste para la distribución geométrica con parámetro $q = 1 - p$, es decir, una binomial negativa con parámetros $r = 1$ y $p = 1 - q$. Se replican algunos escenarios cuando la alternativa es binomial, beta-binomial y uniforme discreta (ver Tables 1.13, 1.14 y 1.15). Los resultados de nuestras tablas para A^2 son muy similares a los que reportan en [Best & Rayner \(2003\)](#).

Las estadísticas \tilde{f}_n y KL dan resultados muy pobres contra cuatro familias de alternativas pero tienen la mejor potencia contra la binomial negativa generalizada. En la mayoría de las situaciones la estadística Δ es mejor que \tilde{f}_n cuando la estadística \tilde{f}_n es sesgada (es decir, tiene potencia inferior al nivel de significancia). Las pruebas d y T_{nb} no fueron buenas si el tamaño de muestra es pequeño (con $n = 10$). Finalmente, las estadísticas Cramér-von Mises tienen buena potencia en general.

1.4.3. Distribución binomial

Las tablas 1.17-1.21 reportan la potencia de varias estadísticas de prueba para probar bondad de ajuste de la distribución binomial (m conocido y p desconocido) cuando la alternativa es Poisson, binomial negativa, beta-binomial, uniforme discreta y binomial negativa generalizada, respectivamente.

Se replicaron algunos escenarios de Best & Rayner (1997) cuando la alternativa es beta-binomial y uniforme discreta, los resultados obtenidos son muy similares a los que publicaron en Best & Rayner (1997) para W^2 y A^2 (ver Tablas 1.19 y 1.20).

En general, la prueba W^2 es menos potente que las demás estadísticas. La estadística T_b tiene buena potencia contra cuatro familias de alternativas pero es pobre contra la alternativa uniforme discreta. Mientras que las estadísticas KL , A^2 , W_m^2 , T_p y \tilde{f}_n tienen buena potencia en la mayoría de los escenarios; la prueba \tilde{f}_n tiene la mejor potencia.

Cuando el tamaño de muestra n es pequeño ($n = 10$) la prueba d no es recomendada como es mencionado en González-Barrios et al. (2006).

1.5. Ejemplos

1.5.1. Muertes por patadas de caballo en el ejército prusiano

La tabla 1.2 reporta los datos del número de muertes por patadas de caballos en 14 regimientos de la caballería prusiana durante los años 1875 a 1894 utilizados en Frey (2012) para probar bondad de ajuste para la distribución Poisson.

TABLA 1.2: Frecuencia de muertes por patadas de caballos en el ejército prusiano obtenido por (De [Bortkiewicz 1898](#)).

Año	Número de muertes	Año	Número de muertes
1875	0	1885	1
1876	0	1886	3
1877	0	1887	1
1878	1	1888	0
1879	1	1889	1
1880	2	1890	2
1881	2	1891	0
1882	1	1892	1
1883	1	1893	1
1884	2		

TABLA 1.3: Muertes por patadas de caballos en el ejército prusiano. Prueba de bondad de ajuste Poisson (λ desconocido). Valor de la estadística de prueba y su correspondiente p -valor condicional, para 10000 muestras look-alike.

KL	B	d	W^2	A^2	W_m^2	Z	Δ	uso de \tilde{f}_n
0.425	23.693	0.003	0.136	1.396	2.228	7.629	0.000	0.000
0.3542	0.0413	0.0177	0.0978	0.0167	0.0244	0.0073	0.5769	0.0589

Si se quiere probar si los datos siguen una distribución de Poisson con parámetro desconocido λ . La tabla 1.3 reporta los valores de las nueve estadísticas de prueba y sus correspondientes p -valores condicionales evaluados con 10000 muestras look-alike. Con siete estadísticas de prueba podemos rechazar \mathcal{H}_0 si elegimos un nivel de significancia de $\alpha = 0.1$, y con $\alpha = 0.05$ se puede rechazar con cinco estadísticas de prueba. Observe que utilizando las estadísticas KL y Δ , no se puede rechazar la hipótesis nula. En la literatura, este es un problema clásico y en la mayoría de los artículos se concluye que se debe rechazar \mathcal{H}_0 .

1.5.2. Ácaros rojos en las hojas de manzana

En [Bliss & Fisher \(1953\)](#) se muestra un ejemplo de el número de ácaros rojos en las hojas de manzana (ver Tabla 1.4), estos datos son usados en [Rueda & O'Reilly \(1999\)](#) para probar bondad de ajuste para la distribución binomial negativa.

Si se quiere probar que los datos provienen de una distribución binomial negativa con parámetro r conocido y p desconocido; el estimador de máxima verosimilitud es $(r, p) = (1.0246, 0.4719)$. La tabla 1.5 reporta el valor de varias estadísticas de prueba y su correspondiente p -valor condicional, basado en 10000 muestras look-alike. Observe que con $r = 0.8$ y $r = 1$, los p -valores de todas las estadísticas de prueba son más grandes que 0.6, por lo tanto, se concluye que el número de ácaros

TABLA 1.4: Frecuencia de ácaros rojos en las hojas de manzana (De [Bliss & Fisher 1953](#)).

j	0	1	2	3	4	5	6	7
O_j	70	38	17	10	9	3	2	1

TABLA 1.5: Ácaros rojos en las hojas de manzana. Prueba de bondad de ajuste binomial negativa (r conocido y p desconocido). Valor de la estadística de prueba y su correspondiente p -valor condicional, para 10000 muestras look-alike.

r	KL	d	W^2	A^2	W_m^2	T_{nb}	Δ	uso de \tilde{f}_n
0.8	0.01711	0.03383	0.04873	0.25874	0.19245	-1.10026	0.00000	0.00001
	0.9353	0.4051	0.5702	0.616	0.681	0.2371	0.9809	0.9811
1	0.01411	0.00017	0.00904	0.10147	0.10422	-0.50508	0.00000	0.00002
	0.9486	0.9834	0.9429	0.9094	0.8684	0.6032	0.95	0.95
1.5	0.01997	0.08149	0.13344	0.72031	0.49112	0.78002	0.00000	0.00001
	0.6968	0.1503	0.1844	0.1773	0.1844	0.4196	0.7547	0.502
2	0.03037	0.20027	0.29269	1.55719	1.01083	1.78101	0.00000	0.00000
	0.2992	0.0219	0.0304	0.0262	0.0256	0.0675	0.5482	0.1501

rojos en las hojas de manzana sigue una distribución binomial negativa con esos valores de r .

1.5.3. 10 muestras de cinco ratones

[Snedecor & Cochran \(1989\)](#) se reportan los datos de 10 muestras de cinco ratones del mismo laboratorio donde los especímenes fueron inyectados con la misma dosis de Typhimurium. La tabla 1.6 reporta la frecuencia de ratones muertos.

Si se quiere probar que los datos provienen de una distribución binomial con m conocido y p desconocido. En la tabla 1.7 reporta el valor de varias estadísticas de prueba y su correspondiente p -valor condicional, basado en 10000 muestras look-alike.

De la Tabla 1.7, se puede ver que el p -valor de d y A^2 es menor que 0.1. Por lo que, eligiendo estas dos estadísticas, se concluye que hay evidencia para rechazar \mathcal{H}_0 con un nivel de significancia 0.10. Sin embargo, con las otras estadísticas de prueba no podemos rechazar la hipótesis nula a menos que eligiéramos un nivel de significancia más grande que 0.1.

TABLA 1.6: Frecuencia de ratones muertos, inyectados con la misma dosis de Typhimurium (De [Snedecor & Cochran 1989](#), pag. 205).

j	0	1	2	3	4	5
O_j	0	1	2	3	1	3

TABLA 1.7: Ejemplo 10 muestras de cinco ratones. Prueba de bondad de ajuste binomial ($m = 5$ y p desconocido). Valor de la estadística de prueba y su correspondiente p -valor condicional, para 10000 muestras look-alike.

KL	d	W^2	A^2	W_m^2	T_b	Δ	uso de \tilde{f}_n
0.22721	0.23137	0.11828	1.01953	0.41911	1.53298	0.00018	0.0128
0.3149	0.0829	0.1108	0.0739	0.1334	0.1709	0.5126	0.1175

1.6. Comentarios finales

En el caso de la distribución de Poisson, existe una amplia literatura que trata con pruebas de bondad de ajuste, y en muchos estudios realizan comparaciones de potencia de las diferentes estadísticas de prueba. Mientras, en las distribuciones binomial y binomial negativa hay menos trabajo publicado. En este capítulo se espera haber dado una visión de las pruebas existentes en el caso discreto. Se propusieron dos estadísticas prueba, y se realizaron comparaciones de potencia para pruebas recientes y clásicas.

Para la distribución Poisson se recomienda usar B y las estadísticas de prueba de la familia Cramér-von Mises (W^2 , A^2 y W_m^2). La estadística de prueba B propuesta en este trabajo, tiene en general, buena potencia y es competitiva con cualquiera de las estadísticas de la familia Cramér-von Mises.

En cuanto a la prueba Δ , también propuesta en este manuscrito, su principal mérito es el de comportarse bien (no mostrar sesgo) en los casos que la prueba condicional (prueba \tilde{f}_n) en [González-Barríos et al. \(2006\)](#) reportó potencias inferiores al nivel de significancia α .

Para la binomial negativa se recomienda utilizar las estadísticas de prueba la familia Cramér-von Mises. Por último, para las pruebas de ajuste de la distribución binomial, se recomienda utilizar las pruebas KL , A^2 , W_m^2 , T_p y \tilde{f}_n , descartando W^2 ; la prueba \tilde{f}_n en este estudio, probó ser la más potente.

1.7. Tablas del estudio de simulación

TABLA 1.8: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución Poisson (λ desconocido). Número de rechazos en 1000 muestras de la distribución binomial negativa con parámetros r y p . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

p	r	n	KL	B	d	W^2	A^2	W_m^2	Z	Δ	uso de \tilde{f}_n
0.1667	1	10	899	914	933	797	913	898	938	867	944
			866	888	906	733	890	864	920	824	915
0.5	1	10	380	342	343	279	354	320	408	352	437
			280	278	272	193	288	245	338	237	321
		20	526	528	538	465	538	507	565	471	610
			441	459	434	345	441	405	510	364	502
		40	771	792	797	715	787	758	834	596	832
			700	733	732	620	704	662	773	473	770
0.625	1	10	256	209	221	186	219	207	263	221	289
			168	157	161	111	165	152	182	122	182
0.4	2	30	868	920	906	820	922	904	940	576	931
			819	881	880	734	885	856	921	455	902
		40	930	961	948	889	959	945	981	589	967
			884	934	926	823	932	920	962	489	942
2/3	2	20	300	303	301	231	286	266	329	259	369
			215	237	200	146	212	182	273	161	268
		40	456	496	475	384	471	430	537	244	549
			351	402	366	277	345	320	459	165	455
0.8	2	30	212	191	191	170	203	187	217	206	251
			134	129	136	108	128	117	158	137	168
		40	271	264	230	195	237	212	273	224	308
			181	188	164	128	157	136	203	138	214
0.75	3	20	243	209	207	163	206	186	228	192	284
			143	149	123	93	135	112	176	112	192
0.5	5	20	570	638	620	461	651	610	697	369	706
			467	553	505	368	545	501	634	264	616
		40	748	842	830	685	832	808	909	319	865
			673	798	768	590	781	744	858	228	802
0.9	9	50	167	155	137	125	142	139	161	87	202
			92	103	75	65	68	66	105	42	126
0.5	10	20	583	654	675	422	624	573	689	487	704
			498	573	606	312	524	482	633	369	612
2/3	10	20	316	336	367	239	340	308	382	199	428
			220	258	268	147	254	227	311	108	312
		40	430	541	507	360	500	468	592	120	594
			317	425	390	255	398	361	516	72	468
0.75	15	50	322	406	374	269	382	353	456	65	470
			215	297	284	173	279	251	361	31	335
0.9	45	50	143	151	146	116	128	126	158	80	192
			77	88	97	60	81	76	97	38	115

TABLA 1.9: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución Poisson (λ desconocido). Número de rechazos en 1000 muestras de la distribución binomial con parámetros m y p . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

p	m	n	KL	B	d	W^2	A^2	W_m^2	Z	Δ	uso de \tilde{f}_n
0.5	2	20	354	636	549	486	508	513	528	497	163
		40	219	400	409	351	342	364	291	340	70
	40	20	912	988	883	857	913	905	933	948	562
		40	784	934	804	740	816	794	819	861	361
1/3	3	20	120	238	250	193	196	201	195	194	49
		40	59	117	138	120	105	118	73	98	20
	40	20	278	456	469	401	412	414	457	427	107
		40	163	313	328	260	276	282	273	292	49
0.25	4	30	107	198	234	206	193	207	192	222	44
		50	62	106	146	131	111	126	89	141	21
	50	30	184	302	336	292	293	302	299	335	65
		50	98	190	222	186	191	197	179	204	25
0.1	10	30	83	102	118	115	108	117	78	116	56
		50	48	45	60	62	51	59	28	62	28
	50	30	85	101	121	125	121	126	95	125	55
		50	34	45	68	57	57	60	39	58	21
0.5	10	30	237	741	531	609	650	649	781	502	32
		40	132	604	218	488	521	528	606	353	11
	40	30	364	888	731	748	800	789	903	711	52
		40	213	759	422	614	687	674	791	572	18
0.5	12	20	92	555	381	444	451	453	573	462	12
		40	44	380	55	321	310	324	358	322	4
	40	20	314	873	668	737	806	795	912	642	38
		40	174	768	319	611	672	667	798	499	11
1/3	15	20	86	251	108	218	204	212	270	139	18
		40	37	138	11	129	117	122	116	74	7
	40	20	113	424	307	359	383	385	493	357	13
		40	58	301	110	250	263	275	312	234	5
0.25	20	30	96	223	155	201	188	193	247	160	30
		50	45	123	36	127	114	119	124	80	16
	50	30	109	335	245	255	272	272	364	291	24
		50	63	190	83	159	162	165	216	180	10
0.5	20	20	70	574	518	477	487	497	610	181	8
		40	30	390	216	333	324	340	370	101	0
0.3	25	30	68	279	238	245	240	251	333	173	16
		40	32	173	84	162	152	157	173	85	7
	40	30	105	386	331	323	335	338	431	246	22
		40	44	249	121	218	216	220	270	149	5
0.1	50	30	69	101	90	98	93	95	94	140	49
		50	31	44	31	53	45	46	39	64	17
	50	93	119	115	124	116	120	136	151	56	
			45	72	35	70	62	61	56	86	29

TABLA 1.10: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución Poisson (λ desconocido). Número de rechazos en 1000 muestras de la distribución beta-binomial con parámetros m , α y β . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

m	α	β	n	KL	B	d	W^2	A^2	W_m^2	Z	Δ	uso de \tilde{f}_n
3	1	2	20	96	76	62	129	109	114	27	71	88
				51	31	37	61	63	64	13	41	48
			40	176	147	107	205	187	186	20	52	129
				102	65	52	125	116	119	6	18	77
4	2	6	20	77	67	74	94	96	88	51	57	71
				45	26	35	42	38	38	23	26	40
			40	69	67	100	118	115	117	59	44	72
				34	24	50	48	48	53	19	15	26
6	4	2	20	221	633	191	478	490	483	428	355	58
				109	415	43	336	325	326	227	212	17
			40	734	978	305	799	881	858	734	867	247
				571	917	177	682	770	728	577	729	137
7	1	0.4	20	911	961	667	926	937	932	167	898	842
				837	884	535	872	885	876	105	827	747
			40	1000	1000	789	1000	1000	1000	163	999	1000
				1000	1000	723	999	1000	1000	103	920	992
7	3	2	20	146	250	141	243	230	229	98	158	96
				84	121	79	133	124	132	44	83	53
			40	268	629	162	422	446	425	205	262	133
				146	386	82	276	283	273	93	150	63
8	2	1.2	20	330	408	390	398	408	408	83	262	304
				204	264	283	273	293	293	44	157	203
			40	626	826	493	628	701	668	86	416	535
				495	660	411	502	555	541	41	291	403
13	2	0.6	20	760	901	491	850	864	852	209	803	633
				646	781	411	748	753	741	131	705	518

TABLA 1.11: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución Poisson (λ desconocido). Número de rechazos en 1000 muestras de la distribución uniforme con parámetros a y b . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

a	b	n	KL	B	d	W^2	A^2	W_m^2	Z	Δ	uso de \tilde{f}_n
0	2	20	367	483	142	320	356	341	164	367	175
			216	281	78	207	222	209	63	231	74
		30	632	789	173	405	569	526	306	654	310
			458	574	90	279	372	338	127	469	183
		50	970	989	250	692	890	830	551	957	749
			877	944	149	504	741	641	339	886	534
0	4	30	465	517	242	446	501	477	32	284	347
			328	302	166	307	348	330	14	166	243
		50	829	896	353	667	746	701	34	468	625
			695	715	261	537	610	589	16	312	490
0	5	30	536	552	519	518	599	591	124	261	498
			411	382	421	395	472	460	75	147	372
		50	858	897	708	794	849	839	183	377	766
			768	785	624	684	766	751	108	243	652
0	6	30	625	679	753	650	747	735	347	282	647
			513	514	696	520	624	616	254	184	535
		50	904	923	895	879	923	920	491	370	873
			822	861	848	791	873	865	381	251	781
2	12	30	558	642	555	692	751	747	461	328	621
			438	505	355	592	658	655	358	211	508
		40	714	788	646	816	864	859	582	361	747
			595	675	459	723	798	788	465	248	632
5	15	20	189	138	14	244	247	251	18	174	144
			111	49	3	146	148	153	10	100	76
7	13	20	211	1000	878	544	727	696	921	329	18
			93	995	382	345	452	444	637	177	6

TABLA 1.12: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial negativa (r conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución Poisson con parámetro λ . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

λ	r	n	KL	d	W^2	A^2	W_m^2	T_{nb}	Δ	uso de \tilde{f}_n
0.5	1	10	19	67	67	69	65	46	30	17
		30	13	40	52	49	50	8	10	12
	50	10	115	286	298	285	283	174	65	51
		30	57	166	176	172	170	45	32	23
	50	10	189	463	436	422	413	363	89	91
		30	109	329	321	316	314	150	47	49
0.5	5	10	33	29	30	31	31	24	38	30
		30	18	16	20	20	20	12	21	18
	50	10	72	83	86	84	85	38	63	53
		30	31	34	37	35	36	11	26	29
	50	10	65	91	93	86	87	72	86	45
		30	32	35	46	41	43	28	42	19
1	1	10	84	301	311	296	267	205	70	71
		30	59	174	193	184	182	66	53	54
	50	10	276	718	680	674	655	622	172	168
		30	179	555	538	532	514	373	112	95
	50	10	502	905	888	889	876	849	285	284
		30	351	830	805	801	795	625	175	166
1	5	10	60	74	77	80	80	50	80	49
		30	24	36	43	37	40	23	37	22
	50	10	78	146	145	128	132	123	115	45
		30	33	84	78	65	71	36	62	23
	50	10	94	182	171	161	161	147	177	48
		30	39	97	84	84	88	68	101	22
3	5	10	68	154	171	154	159	164	98	26
		30	33	39	93	68	75	42	43	10
	50	10	143	382	398	393	394	420	272	37
		30	70	241	281	276	276	224	145	13
	50	10	228	536	528	556	548	653	468	79
		30	132	374	399	438	425	427	329	32

TABLA 1.13: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial negativa (r conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución binomial con parámetros m y p^* . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

p^*	m	r	n	KL	d	W^2	A^2	W_m^2	T_{nb}	Δ	uso de \tilde{f}_n
0.17	4	1	25	286	644	607	624	608	543	152	171
				178	487	467	475	467	269	89	98
			50	593	896	875	884	878	856	358	372
				453	824	784	792	789	631	244	257
0.1	10	1	25	341	544	590	599	640	666	214	221
				228	475	505	521	543	375	120	142
			50	676	661	740	789	852	944	428	453
				537	615	696	730	788	791	294	312
0.2	10	5	10	98	215	232	210	216	199	149	37
				42	111	144	124	127	69	81	10
			30	217	533	483	512	505	575	381	71
				110	366	353	369	371	339	255	31
			50	385	727	683	722	716	814	603	169
				265	583	548	580	575	608	477	85
0.2	10	10	10	89	157	157	141	146	136	102	41
				50	64	92	71	73	40	52	16
			30	147	332	298	318	321	322	273	56
				85	228	195	195	199	170	172	21
			50	220	451	412	453	451	565	458	80
				131	332	289	310	309	356	321	34
				141	615	768	753	753	823	229	27
				71	111	638	586	600	491	130	11
0.5	10	5	10	141	615	768	753	753	823	229	27
				71	111	638	586	600	491	130	11
			30	819	989	997	1000	999	1000	951	305
				662	956	994	996	996	998	885	177
			50	993	1000	1000	1000	1000	1000	1000	840
				976	998	1000	1000	1000	1000	1000	711
0.5	10	10	10	93	142	518	498	511	602	157	5
				40	2	369	329	346	273	90	2
			30	546	921	969	977	976	993	835	112
				396	755	911	938	938	963	701	49
			50	917	983	999	1000	1000	1000	995	453
				832	960	998	999	999	1000	978	289

TABLA 1.14: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial negativa (r conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución beta-binomial con parámetros m , α y β . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

m	α	β	r	n	KL	d	W^2	A^2	W_m^2	T_{nb}	Δ	uso de \tilde{f}_n
3	1	2	1	25	171	466	442	472	490	468	74	83
				50	91	307	305	324	342	209	42	36
				50	570	770	769	822	849	872	211	214
9	0.5	1	1	25	147	77	169	176	290	97	128	125
				50	86	46	90	96	165	17	78	84
				50	260	96	308	394	613	226	188	191
9	1	2	1	25	89	373	463	466	511	511	44	54
				50	42	232	322	318	369	234	17	32
				50	138	573	710	739	805	852	56	58
3	2	3	3	25	161	378	382	404	412	412	187	62
				50	79	234	237	263	270	174	105	28
				50	536	599	624	691	693	755	568	149
5	0.9	6	5	25	150	137	136	149	147	127	135	190
				50	80	75	70	77	78	77	73	96
				50	181	212	216	220	218	158	136	232
					104	121	132	139	134	109	74	139

TABLA 1.15: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial negativa (r conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución uniforme con parámetros a y b . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

a	b	r	n	KL	d	W^2	A^2	W_m^2	T_{nb}	Δ	uso de \tilde{f}_n
0	2	1	10	296	511	539	532	531	442	185	192
			30	172	327	351	351	405	215	162	165
			50	955	932	950	978	983	964	804	812
				874	857	899	934	952	833	647	651
				999	996	998	1000	1000	1000	989	989
				997	979	989	998	999	987	969	961
0	3	5	10	136	93	213	195	209	101	144	94
			30	82	64	115	89	100	31	77	40
			50	602	148	513	607	606	339	530	275
				428	76	355	427	431	157	360	157
				934	183	727	856	846	591	904	636
				825	108	564	722	715	333	768	444
0	4	1	10	228	360	492	491	555	486	205	189
			30	142	243	326	324	403	238	109	115
			50	807	838	932	956	974	961	587	556
				646	718	858	901	949	839	429	415
				995	951	995	999	999	1000	929	937
				981	913	982	994	999	988	842	843
2	9	5	10	74	54	251	213	238	250	108	15
			30	35	0	139	104	125	64	51	4
			50	423	822	622	818	797	900	579	68
				248	391	457	598	608	680	397	29
				997	994	908	1000	999	1000	1000	367
				946	940	792	989	968	971	1000	206
2	11	5	10	88	22	199	168	188	160	77	30
			30	49	1	118	81	100	36	39	10
			50	233	309	444	543	577	612	330	45
				118	71	286	339	375	336	194	22
				739	715	711	896	876	905	825	188
				516	355	518	724	738	702	665	101

TABLA 1.16: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial negativa (r conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución binomial negativa generalizada con parámetros k , a y b . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

a	b	k	r	n	KL	d	W^2	A^2	W_m^2	T_{nb}	Δ	uso de \tilde{f}_n
0.1	2	1	1	10	22	14	14	18	18	19	31	21
				30	9	9	9	9	9	12	19	10
				50	159	149	145	145	145	154	144	157
					99	88	92	93	93	89	76	87
					178	175	159	159	159	172	182	189
					129	132	116	120	120	130	121	134
0.1	4	1	5	10	158	159	153	154	154	165	141	162
				30	137	129	127	132	127	136	120	138
				50	428	436	425	423	423	434	430	434
					388	402	381	382	381	395	400	397
					639	654	653	658	656	649	630	652
					590	604	575	576	575	595	576	603
0.3	2	1	1	10	248	191	176	200	210	218	266	247
				30	164	137	116	134	141	181	197	161
				50	612	610	571	593	595	614	680	667
					518	517	467	485	489	560	588	575
					766	796	770	779	783	795	811	837
					679	725	691	702	700	742	746	745
0.3	3	1	5	10	746	741	733	740	740	756	780	759
				30	709	699	693	702	699	721	743	721
				50	989	991	989	990	989	991	993	993
					985	987	987	988	987	986	993	988
					1000	1000	1000	1000	1000	1000	1000	1000
					1000	1000	1000	1000	999	1000	1000	1000

TABLA 1.17: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial (m conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución Poisson con parámetro λ . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

λ	m	n	KL	d	W^2	A^2	W_m^2	T_b	Δ	uso de \tilde{f}_n	
4	7	10	586	669	520	634	571	629	534	676	
			533	582	475	570	523	597	484	597	
		30	909	951	908	942	924	945	798	944	
			890	925	882	933	906	931	791	924	
		50	985	994	984	993	988	995	930	990	
			979	988	971	990	984	992	930	987	
	4	9	10	405	447	278	434	366	420	311	485
				320	325	195	364	287	371	226	390
			30	656	765	553	736	649	774	334	767
				574	673	473	654	570	707	285	692
		50	802	905	737	881	833	901	388	894	
			733	849	660	829	761	876	369	842	
7	15	10	442	36	274	456	386	459	356	541	
			345	31	180	381	292	411	266	420	
		30	663	813	530	748	668	786	324	808	
			572	719	427	671	582	731	252	709	
		50	836	953	727	911	865	940	245	921	
			775	902	629	865	791	903	196	875	
7	18	10	312	14	203	342	290	342	268	395	
			234	4	123	263	211	300	191	303	
		30	522	697	370	590	510	656	219	677	
			409	570	272	483	401	584	142	567	
		50	634	836	512	735	661	817	125	775	
			526	756	390	634	569	737	60	680	
10	20	10	470	18	286	482	420	500	435	548	
			356	13	184	390	318	434	340	447	
		30	702	874	565	783	724	832	410	820	
			622	801	463	726	633	791	301	761	
		50	877	974	792	939	910	965	370	958	
			811	947	704	904	855	946	276	914	
10	25	10	321	25	172	318	272	346	294	397	
			227	12	113	242	189	287	209	287	
		30	504	729	402	608	544	666	288	688	
			403	602	287	500	428	588	182	559	

TABLA 1.18: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial (m conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución binomial negativa con parámetros r y p^* . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

p^*	r	m	n	KL	d	W^2	A^2	W_m^2	T_b	Δ	uso de \tilde{f}_n				
0.5	1	10	10	415	411	316	413	379	449	362	472				
			30	322	339	241	340	304	391	275	379				
				771	862	696	776	740	814	613	825				
				678	798	594	699	639	770	516	752				
			50	906	967	887	924	905	953	756	941				
				875	934	828	894	861	925	675	909				
				0.8	1	10	10	90	89	79	84	84	90	102	90
							30	41	40	39	40	39	41	43	42
			232					253	227	241	227	244	247	250	
			172					199	156	167	160	186	168	190	
50	314	358	295				316	296	329	293	342				
	224	266	211				225	212	252	203	266				
	0.5	5	30				10	560	5	374	585	527	627	487	677
							30	470	4	283	512	447	578	376	577
854								892	760	905	865	942	535	936	
798								823	676	860	818	921	448	894	
50				961	992	922	984	977	995	407	857				
				937	982	864	973	960	991	324	849				
				0.8	5	30	10	169	28	120	157	135	168	139	201
							30	104	8	62	108	92	115	73	125
248								237	186	244	214	306	146	329	
155								162	122	169	144	223	78	210	
50	317	449	236				304	276	381	97	403				
	225	309	148				207	180	300	50	286				

TABLA 1.19: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial (m conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución beta-binomial con parámetros k , α y β . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

k	α	β	m	n	KL	d	W^2	A^2	W_m^2	T_b	Δ	uso de \tilde{f}_n
5	0.3	2.7	5	20	640	709	586	674	617	684	580	695
				40	564	650	502	596	523	625	505	627
					862	913	864	881	872	889	808	893
					814	873	802	840	809	859	749	862
5	0.7	6.3	5	20	366	464	332	411	362	421	305	436
				40	273	349	249	312	270	321	219	343
					525	649	494	547	513	574	410	601
					434	543	394	455	404	492	319	491
5	1.5	1.5	5	20	675	894	651	802	751	844	256	843
				40	577	823	532	737	665	803	160	751
					933	993	922	967	964	988	188	977
					883	976	862	946	928	975	123	951
5	1.6667	10.5	5	20	226	317	158	232	196	250	154	289
				40	144	221	100	152	123	172	91	196
					350	467	294	361	311	361	151	407
					250	340	197	253	215	284	86	302
5	3.5	3.5	5	20	310	585	296	436	383	467	63	493
				40	202	443	197	326	265	374	29	351
					485	779	477	618	572	691	16	641
					370	648	358	512	451	602	6	522
5	3.7222	33.5	5	20	97	136	80	105	89	109	82	124
				40	53	69	47	57	48	64	47	66
					151	182	110	139	124	140	92	176
					78	111	63	76	65	95	52	104
5	5.8333	5.8333	5	20	201	367	189	250	224	257	53	305
				40	121	253	128	175	152	190	27	208
					283	494	294	390	348	410	19	412
					207	393	200	298	245	327	7	302
5	18.5	18.5	5	20	118	186	112	142	123	114	77	167
				40	61	116	59	78	65	73	37	100
					107	218	93	120	110	127	63	141
					56	127	52	68	58	78	29	75
20	1.6667	10.5	20	20	754	812	648	828	782	871	507	868
				675	759	532	767	715	822	404	795	
20	3.7222	33.5	20	20	306	298	227	341	298	365	157	411
				207	228	148	249	213	306	79	298	

TABLA 1.20: Estimación de la potencia de diferentes estadísticas de prueba para probar bondad de ajuste de la distribución binomial (m conocido y p desconocido). Número de rechazos en 1000 muestras de la distribución uniforme con parámetros a y b . Con un nivel de significancia de $\alpha = 0.1$ y $\alpha = 0.05$, respectivamente. n denota el tamaño de la muestra.

a	b	m	n	KL	d	W^2	A^2	W_m^2	T_b	Δ	uso de \tilde{f}_n
0	3	3	10	316	553	372	398	404	366	182	407
				221	365	207	329	288	329	124	310
			30	779	926	793	847	836	875	151	856
				671	862	699	773	750	812	84	776
			50	959	989	950	967	968	980	157	970
				912	975	891	939	928	960	85	940
0	3	5	10	156	261	201	179	200	62	133	173
				103	128	120	101	113	23	89	102
			30	563	526	497	534	517	152	200	483
				421	386	374	401	404	86	113	343
			50	838	704	730	782	744	217	290	737
				738	549	630	679	651	131	166	635
0	4	5	20	570	821	592	661	664	617	211	667
				462	718	464	558	562	528	139	559
			30	800	938	786	851	853	823	235	849
				693	883	692	771	772	739	154	753
0	4	20	20	351	283	374	403	402	72	155	330
				232	204	262	285	288	40	105	219
0	5	5	20	890	976	852	937	920	967	448	955
				838	962	778	899	880	948	338	920
0	8	15	10	659	31	564	716	699	732	566	795
				561	27	427	629	613	683	445	698
			30	965	998	959	992	988	990	654	987
				933	993	920	978	977	983	555	972
			50	998	1000	998	999	999	999	693	999
				998	999	992	998	999	998	611	998
2	9	9	10	702	790	608	797	735	783	575	818
				586	674	498	729	637	746	479	744
			30	982	992	975	997	986	998	643	995
				959	981	952	992	980	990	531	986
			50	1000	1000	999	1000	1000	1000	674	1000
				999	999	998	1000	1000	1000	588	1000
6	14	20	20	340	148	410	454	449	214	240	361
				242	83	308	345	354	129	151	241

Capítulo 2

Remuestreo paramétrico vs remuestreo condicional

Sea X_1, \dots, X_n es una muestra aleatoria de una función de distribución F , si se quiere probar $\mathcal{H}_0 : F \in \mathcal{F}_0$ vs. $\mathcal{H}_1 : F \in (\mathcal{F} - \mathcal{F}_0)$, donde \mathcal{F}_0 es una familia paramétrica de funciones de distribuciones continuas y \mathcal{F} , es la familia no paramétrica de todas las distribuciones continuas (por ejemplo en \mathbb{R} o \mathbb{R}^+).

En el caso que la familia paramétrica \mathcal{F}_0 es de localización y/o escala (como por ejemplo la distribución normal, exponencial, etc.), la estadística de prueba es invariante ante transformaciones, es decir, suponga que la familia tiene localización μ y escala σ , entonces, se tiene que $D(\underline{x}; \mu, \sigma) = D(\frac{\underline{x}-\mu}{\sigma}; 0, 1)$. Lo que significa que la estadística de prueba no depende del valor del parámetro desconocido; por lo que por simulación, se puede generar un gran número muestras de la distribución F_0 (conocida) con escala 0 y escala 1, y con ello calcular p -valor de la estadística de prueba.

Sin embargo, para las distribuciones cuyos parámetros no se pueden considerar de localización y/o escala, como lo es el caso discreto o en el caso continuo para las distribuciones Gamma, Gaussiana-inversa, exponencial truncada, von Mises, etc., se tiene que la estadística de prueba depende del valor del parámetro desconocido, por lo que existen diversos procedimientos para abordar el problema de bondad de ajuste. En este capítulo se analizan dos enfoques, uno que es una aproximación (remuestreo paramétrico) y otro método que se ilustró para el caso discreto en el capítulo 1 y que produce muestras condicionales (o “look-alike”). Dicho método en general da una solución exacta mediante el uso de la distribución Rao-Blackwell y para usar un lenguaje similar lo referiremos también como remuestreo condicional.

En [Lockhart et al. \(2012\)](#) se estudia las distribuciones asintóticas de las estadísticas de prueba de la familia Crámer-von Mises para las distribuciones que pertenecen a la familia exponencial mediante remuestreo condicional y remuestro paramétrico. Para muestras grandes, se demuestra que el remuestreo condicional y remuestro paramétrico son aproximados.

En este capítulo comparamos los dos enfoques para tamaños de muestras pequeñas ($n = 10$ y $n = 20$). Particularmente, se estudiará la distribución exponencial trunca en 1 y la distribución Gaussiana-inversa. Se hará un estudio de simulación para estimar la potencia de diferentes estadísticas. Para la primer familia mencionada no se tiene conocimiento que se haya realizado un estudio comparativo de potencias.

El primer enfoque que se analiza es el remuestreo paramétrico, este método estima la función de distribución $F_0(\cdot; \theta)$ mediante $F_0(\cdot; \hat{\theta})$, donde $\hat{\theta}$ es un estimador de θ . Este enfoque es una aproximación ya que usa un estimador en lugar del parámetro desconocido. El procedimiento para generar el p-valor mediante remuestreo paramétrico es el siguiente: Sea $D(\cdot; \theta)$ una estadística de prueba, dada la muestra \underline{x} se calcula el estimador $\hat{\theta}$, luego se calcula $D(\underline{x}; \hat{\theta})$; se genera un número grande de muestras de la distribución $F_0(\cdot; \hat{\theta})$, digamos M muestras, por lo que tendremos las muestras $\underline{y}_1, \dots, \underline{y}_M$; para cada muestra se estima θ y se calcula el valor de la estadística de prueba D . Tendremos M observaciones simuladas de la estadística de prueba; es decir, $D(\underline{y}_1; \hat{\theta}_1), \dots, D(\underline{y}_M; \hat{\theta}_M)$. El p-valor (aproximado), es la proporción de valores D_i 's que exceden a D , es decir

$$p \approx \frac{1}{M} \sum_{j=1}^M I\{D(\underline{y}_j; \hat{\theta}_j) > D(\underline{x}; \hat{\theta})\},$$

donde I es la función indicadora. Entre más grande sea M , mejor será la aproximación del p-valor.

Por otro lado, para el segundo enfoque, discutido en el capítulo 1, la familia de distribuciones debe admitir una estadística suficiente minimal y se debe conocer la función de distribución Rao-Blackwell (estimadora de la distribución). Por lo cual, en [O'Reilly & Gracia-Medrano \(2006\)](#) se proporciona un método para encontrar una solución exacta; ellos en particular analizan la distribución Gaussiana-inversa, y posteriormente en [González-Barríos et al. \(2010\)](#), se estudia la exponencial trunca. En el caso, que la familia admita una estadística suficiente minimal y se desconozca la distribución Rao-Blackwell, existen procedimientos más generales,

donde se puede usar el muestreador de Gibbs (Gibbs sampler), como por ejemplo lo que hacen en [Lockhart et al. \(2007\)](#) al estudiar bondad de ajuste de la distribución Gamma; y en [Lockhart et al. \(2009\)](#) para la distribución von Mises. Mediante estos procedimientos se da una solución exacta al problema de bondad de ajuste, identificando el p-valor condicional de la estadística de prueba.

En la siguiente sección se mencionan algunas estadísticas de prueba utilizadas en el caso continuo. En la secciones [2.2](#) y [2.3](#) se abordan la distribución exponencial truncada y la distribución Gaussiana-inversa, donde se realizará un pequeño estudio de simulación con los enfoques mencionados anteriormente.

2.1. Estadísticas de prueba

Suponga que se tiene una muestra aleatoria observada, es decir, x_1, \dots, x_n y sea $u_{(i)} = F_0(x_{(i)}; \hat{\theta})$ con $i = 1, \dots, n$. Sin duda una de las estadísticas más utilizada en la literatura clásica para bondad de ajuste es la estadística de prueba Anderson-Darling y se define como:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\log(u_{(i)}) + \log(1 - u_{(n+1-i)})].$$

Esta estadística pertenece a las estadísticas de la familia Cramér-von Mises, y es muy utilizada ya que en general ha demostrado ser muy potente.

[Alizadeh Noughabi & Arghami \(2013\)](#) proponen una estadística basada en la información Kullback-Leibler, que en casos particulares de la función de distribución a probar, resultó ser equivalente a otras estadísticas de prueba también basadas en la información Kullback-Leibler ya existentes en la literatura. Esta estadística resulta tener buena potencia para varias familias y es competitiva con la estadística Anderson-Darling. No se tiene conocimiento que hayan explorado la potencia de esta prueba en el caso de la distribución Gaussiana-Inversa. A continuación se dan algunos conceptos de la información Kullback-Leibler y se ve la estadística de prueba propuesta en [Alizadeh Noughabi & Arghami \(2013\)](#).

Suponga que X_1, \dots, X_n es una muestra aleatoria de una función de densidad desconocida f . Sea $f_0(x; \theta)$ una familia paramétrica de funciones de densidad con espacio parametral Θ . En el caso compuesto, la hipótesis de interés es,

$$\mathcal{H}_0 : f(x) = f_0(x; \theta), \text{ para algún } \theta \in \Theta.$$

Y la hipótesis alternativa es,

$$\mathcal{H}_1 : f(x) \neq f_0(x; \theta), \text{ para cualquier } \theta \in \Theta.$$

Para discriminar entre \mathcal{H}_0 y \mathcal{H}_1 , se usa la información Kullback-Leibler que es dada por:

$$\begin{aligned} R(f, f_0; \theta) &= \int_{-\infty}^{\infty} f(x) \log \left(\frac{f(x)}{f_0(x; \theta)} \right) dx \\ &= -H(f) - \int_{-\infty}^{\infty} f(x) \log f_0(x; \theta) dx \end{aligned}$$

donde $H(f)$ es la entropía de f . La información Kullback-Leibler siempre es mayor o igual a cero y es igual a cero si y sólo si $f(x) = f_0(x; \theta)$. Se define la estadística de prueba,

$$R_{inf} = \inf_{\theta \in \Theta} R(f, f_0; \theta)$$

se rechaza \mathcal{H}_0 para valores “grandes” de la estadística de prueba. De forma directa se tiene que,

$$R_{inf} = -H(f) - I,$$

donde,

$$I = \sup_{\theta \in \Theta} \int_{-\infty}^{\infty} f(x) \log f_0(x; \theta) dx.$$

[Alizadeh Noughabi & Arghami \(2013\)](#) para estimar $H(f)$ usan el estimador de Vasicek H_{mn} , el cual implica una conceptualización de la f empírica. Dicho estimador de $H(f)$ es,

$$H_{mn} = \frac{1}{n} \sum_{i=1}^n \log \left\{ \frac{n}{2m} (X_{(i+m)} - X_{(i-m)}) \right\},$$

donde m es un entero positivo más pequeño que $n/2$, $X_{(i)} = X_{(1)}$ si $i < 1$, $X_{(i)} = X_{(n)}$ si $i > n$, y $X_{(1)}, \dots, X_{(n)}$ son las estadísticas de orden. Para estimar I se propone,

$$\hat{I} = \sup_{\theta \in \Theta} \sum_{i=1}^n \frac{1}{n} \log f_0(x_i; \theta),$$

es decir, el estimador de máxima verosimilitud. Por lo que estadística de prueba es (ver [Alizadeh Noughabi & Arghami 2013](#)),

$$\hat{R}_{inf} = -H_{mn} - \hat{I},$$

esta estadística es consistente y es invariante bajo familias de localización y/o escala. Como se discute en [Alizadeh Noughabi & Arghami \(2013\)](#), para elegir el

entero m donde la prueba \hat{R}_{inf} sea más potente depende de la distribución a probar. En general, no se tiene un procedimiento para elegir m donde la potencia sea óptima.

2.2. Distribución exponencial truncada

Se dice que una variable aleatoria X tiene una distribución exponencial truncada en 1 con parámetro $\theta \in \mathbb{R}$, se denota por $ExpT(\theta)$, si tiene función de densidad,

$$f(x; \theta) = \frac{\theta e^{\theta x}}{e^{\theta} - 1}, \quad 0 \leq x \leq 1, \quad \theta \in \mathbb{R}. \quad (2.1)$$

La función de distribución de la exponencial truncada en 1 es,

$$F(x; \theta) = \frac{e^{\theta x} - 1}{e^{\theta} - 1}, \quad 0 \leq x \leq 1,$$

cuando $\theta = 0$ se convierte en la distribución uniforme en $(0, 1)$. Sea x_1, x_2, \dots, x_n una muestra aleatoria de la distribución exponencial truncada en 1; en este caso $t_n = \sum_{i=1}^n x_i$ es suficiente minimal.

La obtención del estimador de máxima verosimilitud se obtiene resolviendo numéricamente la siguiente igualdad,

$$\frac{1}{\theta} - \frac{1}{1 - e^{-\theta}} + \frac{t_n}{n} = 0 \quad (2.2)$$

En el artículo de [González-Barrios et al. \(2010\)](#) se proporciona el estimador Rao-Blackwell basado en t_n para la exponencial truncada en 1,

$$\tilde{G}_n(x|t_n) = (n-1) \frac{\sum_{j=0}^{n-1} (-1)^j \binom{n-1}{j} \int_0^x (t_n - j - u)_+^{n-2} du}{\sum_{j=0}^n (-1)^j \binom{n}{j} (t_n - j)_+^{n-1}}, \quad (2.3)$$

donde $(\cdot)_+$ es la parte positiva del argumento, y con unos cálculos se tiene que,

$$\tilde{G}_n(x|t_n) = \frac{\sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \{(t_n - j)_+^{n-1} - (t_n - j - x)_+^{n-1}\}}{\sum_{j=0}^n \binom{n}{j} (-1)^j (t_n - j)_+^{n-1}}, \quad (2.4)$$

Observe que la función 2.4 está definida para $n \geq 2$.

Una vez que se tiene la distribución Rao-Blackwell, para generar la muestra condicional se procede como se vió en el capítulo 1, es decir, se generan primero los $(n-1)$ elementos de la muestra; x_2^*, \dots, x_n^* , y por último, se obtiene $x_1^* = t_n - \sum_{j=2}^n x_j^*$. La muestra-* que se acaba de obtener es una muestra condicional dada t_n .

En la siguiente sección se hace un estudio de simulación para estimar la potencia de la estadística Anderson-Darling.

2.2.1. Estudio de potencia

Se estudiará la potencia de la estadística Anderson-Darling. El cálculo de las potencias de las pruebas se hizo de la siguiente forma: de la distribución alternativa elegida se genera una muestra, bajo \mathcal{H}_0 , se generan 1000 muestras look-alike y 1000 muestras por remuestreo paramétrico ($\hat{\theta}$ por máxima verosimilitud), y con ellas se calcula los correspondientes p -valores de las estadísticas de prueba, se eligió un nivel de significancia $\alpha = 0.05$, si el p -valor es menor que α se rechaza \mathcal{H}_0 . Lo anterior se hizo 1000 veces, y se calculó la potencia como el porcentaje de rechazos de las respectivas estadísticas de prueba.

Es este estudio se usarán como alternativas, la distribución exponencial truncada en 1, la distribución uniforme, la distribución beta, y por último la distribución normal con media μ y desviación estándar σ truncada de 0 a 1 que se denotará por $TrN(\mu, \sigma)$.

Para la estimación de la estadística de prueba Anderson-Darling, se usaron dos métodos de estimación, el primero por medio del estimador de máxima verosimilitud, y el segundo mediante el estimador Rao-Blackwell (es decir se estima $F(\cdot; \theta)$ con $\tilde{G}_n(\cdot|t_n)$). La potencia de la estadística de prueba por dos métodos de estimación son muy similares. Por lo que, aquí solo se presenta la estadística estimada por el método de estimación usual (máxima verosimilitud). En la tabla 2.1 se reportan

TABLA 2.1: Potencia de la estadística de prueba Anderson-Darling para probar bondad de ajuste de la distribución exponencial truncada (caso compuesto) para diferentes alternativas. Tamaño de muestra $n = 10$ y $\alpha = 0.05$. $\rho(\cdot)$ denota la correlación de los p-valores de la estadística Anderson-Darling mediante look-alike y remuestreo.

Alternativas	look-alike	remuestreo	$\rho(A^2)$
<i>ExpT</i> (1)	0.049	0.049	0.9978794
<i>Unif</i> (0, 1)	0.052	0.050	0.9980927
<i>Beta</i> (0.5, 0.5)	0.54	0.537	0.9987123
<i>Beta</i> (1, 3)	0.055	0.057	0.9978939
<i>Beta</i> (2, 5)	0.292	0.289	0.9978624
<i>TrN</i> (0.2, 0.3)	0.089	0.082	0.9980035
<i>TrN</i> (0.5, 0.3)	0.123	0.113	0.9981478

las potencias de la estadística de prueba Anderson-Darling para diferentes alternativas mediante la generación de muestras look-alike y por remuestreo paramétrico. También se reportan la correlación de los p-valores de la respectiva estadística de prueba mediante muestras look-alike y por remuestreo paramétrico.

En las figuras 2.1 a 2.7, se gráfica los p-valores con remuestreo condicional contra los p-valores con remuestreo paramétrico para la estadística de prueba A^2 para las diferentes altertativas con $n = 10$. $\rho(A^2)$ denota el coeficiente de correlación de los p-valores obtenidos por los dos enfoques. Se puede observar que existe un gran parecido entre los dos métodos de remuestreo (el condicional y el paramétrico).

2.3. Distribución Gaussiana-inversa

Se dice que una variable aleatoria X tiene una distribución Gaussiana-inversa con parámetros $\mu > 0$ y $\lambda > 0$, y se denota por $X \sim InvGau(\mu, \lambda)$, si la función de densidad de X es

$$f(x; \mu, \lambda) = \begin{cases} [\lambda/2\pi x^3]^{\frac{1}{2}} \exp[-\lambda(x - \mu)^2/2\mu^2 x], & \text{si } x > 0, \\ 0, & \text{si } x \leq 0. \end{cases}$$

La función de distribución de la Gaussiana-inversa puede ser escrita como:

$$F(x; \mu, \lambda) = \Phi(R) + \Phi(L) \exp\left\{\frac{2\lambda}{\mu}\right\}, \quad \text{con } x > 0,$$

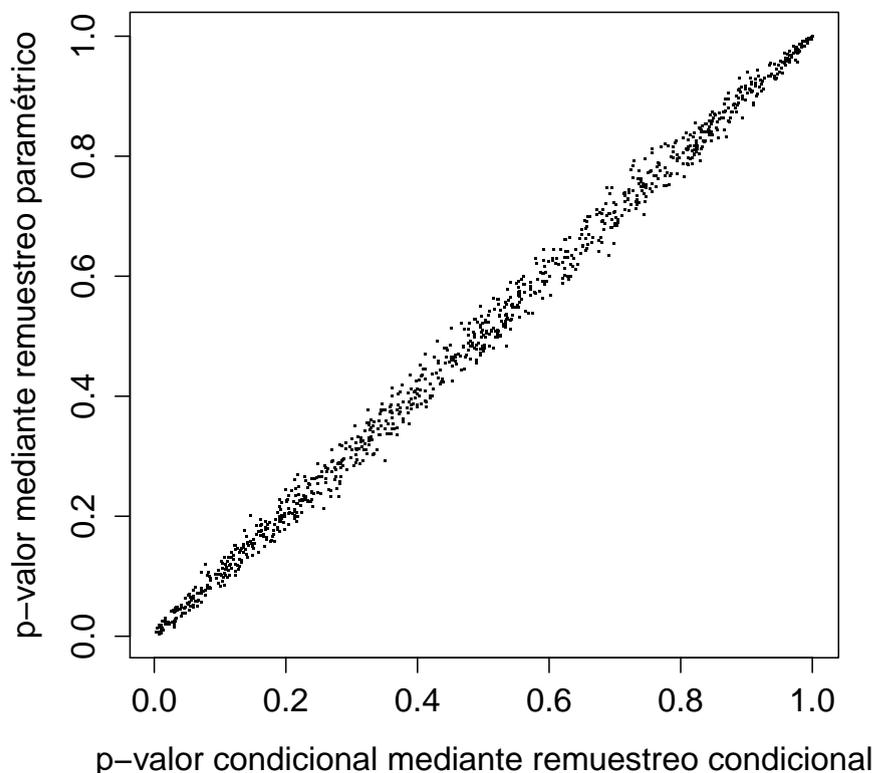


FIGURA 2.1: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la alternativa es $ExpT(1)$ con $n = 10$.
 $\rho(A^2) = 0.9978794$.

donde $R = R(x; \mu, \lambda) = -(\frac{\lambda}{x})^{\frac{1}{2}} + (\frac{\lambda x}{\mu^2})^{\frac{1}{2}}$, $L = L(x; \mu, \lambda) = -(\frac{\lambda}{x})^{\frac{1}{2}} - (\frac{\lambda x}{\mu^2})^{\frac{1}{2}}$ y Φ es la función de distribución normal estándar. La estadística suficiente minimal para esta familia es

$$T_n = t_n = (t_{n_1}, t_{n_2}) = \left(\sum_{i=1}^n x_i, \sum_{i=1}^n \frac{1}{x_i} \right).$$

El correspondiente estimador Rao-Blackwell es (ver [Raj S. Chhikara 1974](#)):

$$\tilde{G}_n(x|t_n) = \begin{cases} 0 & \text{si } x \leq l, \\ 1 & \text{si } x \geq u, \\ H_{n-2}(W) + \frac{n-2}{n} \left[1 + \frac{4(n-1)\hat{\lambda}}{n^2\hat{\mu}} \right]^{\frac{(n-3)}{2}} H_{n-2}(-W') & \text{en otro caso,} \end{cases}$$

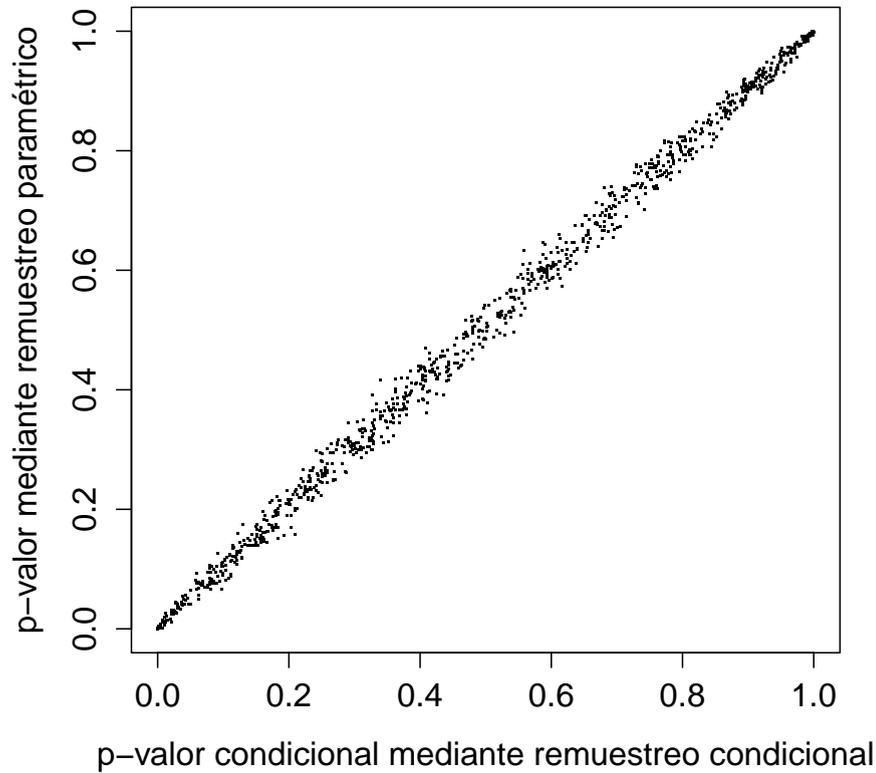


FIGURA 2.2: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la altertativa es $Unif(0,1)$ con $n = 10$. $\rho(A^2) = 0.9980927$.

con $\hat{\mu}$ y $\hat{\lambda}$ los estimadores de máxima verosimilitud, es decir,

$$\hat{\mu} = \bar{X} \quad \text{y} \quad \hat{\lambda} = \frac{n}{\sum_{i=1}^n \left(\frac{1}{X_i} - \frac{1}{\bar{X}} \right)},$$

se tiene que $\hat{\mu}$ y $\hat{\lambda}$ son funciones de la estadística suficiente minimal T_n , donde H_{n-2} es la función de distribución t Student con $n - 2$ grados de libertad, y

$$W = \frac{1}{C} \sqrt{n(n-2)} \left(\frac{x}{\hat{\mu}} - 1 \right), \quad W' = \frac{1}{C} \sqrt{n(n-2)} \left(1 + \frac{n-2}{n} \frac{x}{\hat{\mu}} \right),$$

$$\text{y} \quad C = \sqrt{\frac{n}{\hat{\lambda}} \left(n - \frac{x}{\hat{\mu}} \right) x - n \left(1 - \frac{x}{\hat{\mu}} \right)^2}.$$

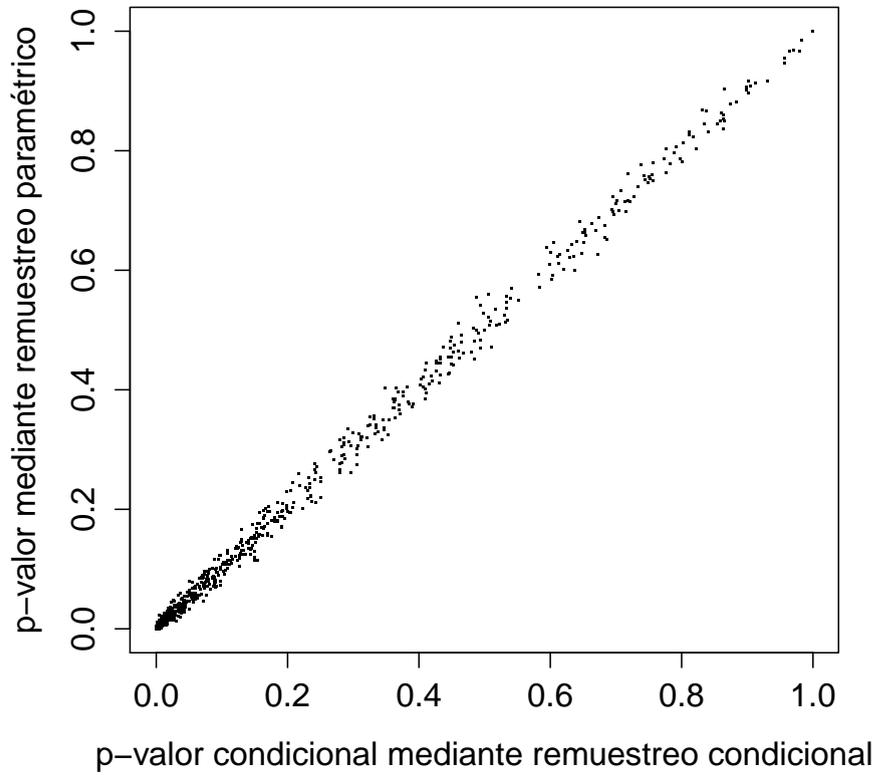


FIGURA 2.3: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la alternativa es $Beta(0.5, 0.5)$ con $n = 10$. $\rho(A^2) = 0.9987123$.

Por último,

$$l = \frac{\bar{X}}{2(n + V\bar{X})} \left[n(2 + V\bar{X}) - \sqrt{4n(n-1)V\bar{X} + n^2V^2\bar{X}^2} \right],$$

$$u = \frac{\bar{X}}{2(n + V\bar{X})} \left[n(2 + V\bar{X}) + \sqrt{4n(n-1)V\bar{X} + n^2V^2\bar{X}^2} \right],$$

$$y \quad V = \sum_{i=1}^n \left(\frac{1}{X_i} - \frac{1}{\bar{X}} \right).$$

Para generar la muestra condicional se procede como se vio en el capítulo 1, es decir, se generan primero los $(n-2)$ elementos de la muestra; x_3^*, \dots, x_n^* , y por último, sea $c_1 = t_{n_1} - \sum_{j=3}^n x_j^*$ y $c_2 = t_{n_2} - \sum_{j=3}^n \frac{1}{x_j^*}$, por lo que x_1^* y x_2^* se obtienen resolviendo la ecuación cuadrática $x^2 - c_1x + \frac{c_1}{c_2} = 0$.

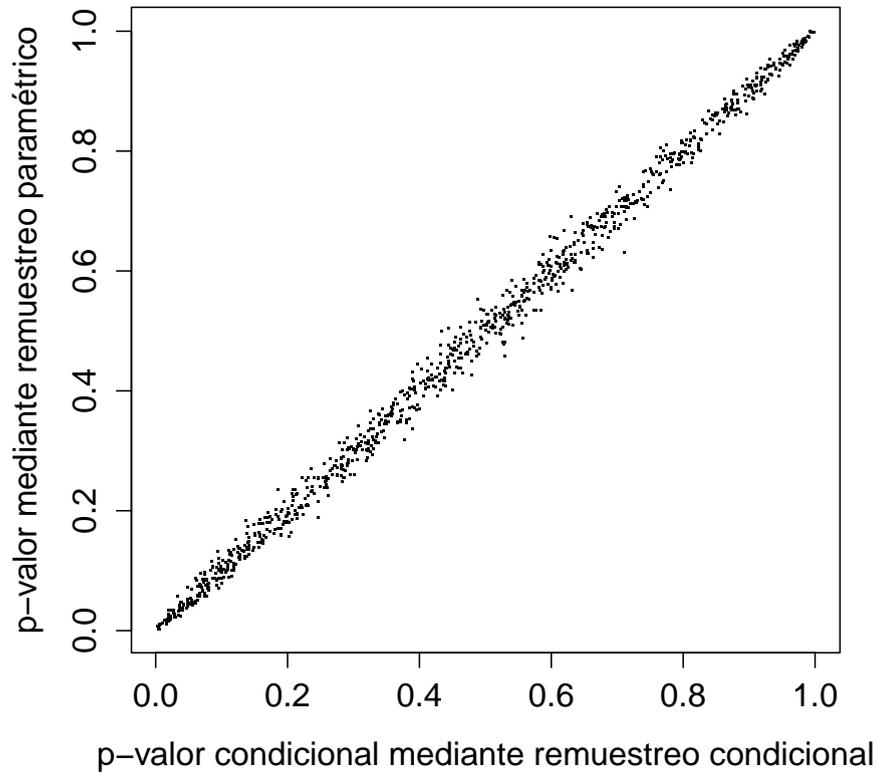


FIGURA 2.4: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la alternativa es $Beta(1,3)$ con $n = 10$. $\rho(A^2) = 0.9978939$.

En la siguiente sección se hace un estudio de simulación para estimar la potencia de la estadística \hat{R}_{inf} y de la estadística Anderson-Darling.

2.3.1. Estudio de potencia

Para calcular la potencia de las estadísticas de prueba se hizo análogamente como se describe en el estudio de potencia de la distribución exponencial truncada. Se usaron las alternativas utilizadas en O'Reilly & Gracia-Medrano (2006), adicionalmente se incluyeron la distribución beta, gamma, y la distribución χ^2 .

Para la estimación de las estadísticas de prueba, se usaron dos métodos de estimación, el primero por medio del estimador de máxima verosimilitud, y el segundo mediante el estimador Rao-Blackwell. La potencia de las estadísticas de prueba por

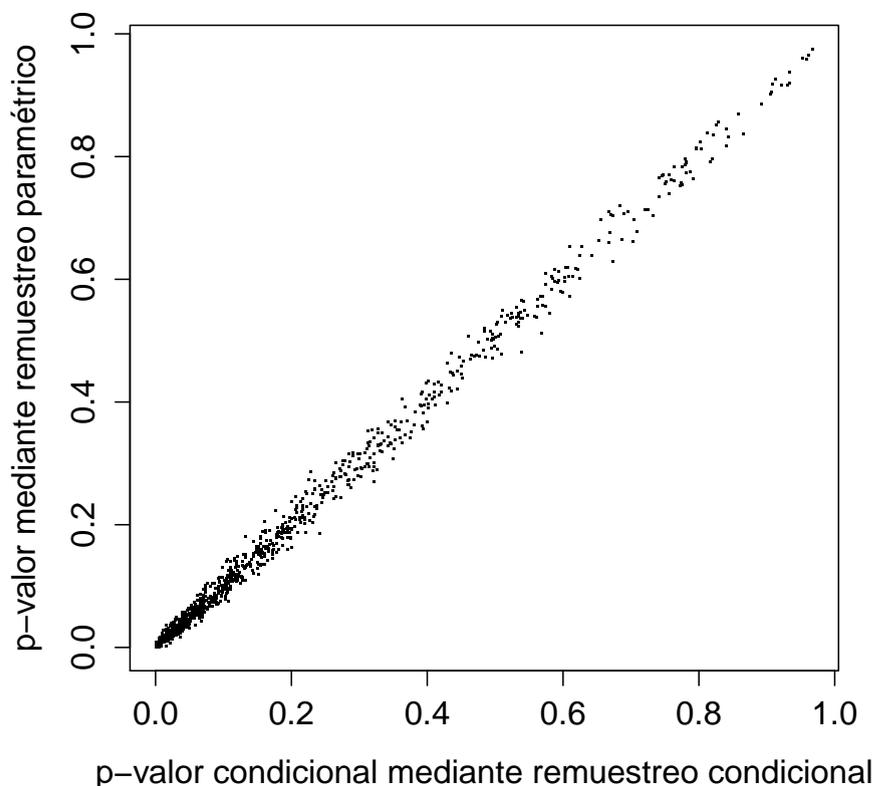


FIGURA 2.5: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la alternativa es $Beta(2, 5)$ con $n = 10$. $\rho(A^2) = 0.9978624$.

dos métodos de estimación fueron muy similares. Por lo que, aquí solo se presenta las estadísticas estimadas por el método de máxima verosimilitud.

En las tablas 2.2 y 2.3, se reporta la potencia para las diferentes estadísticas de prueba, mediante la generación de muestras look-alike y por medio de remuestreo paramétrico con $n = 10$ y $n = 20$, respectivamente. Se calculó la correlación de los p-valores de las correspondientes estadísticas de prueba por los dos procedimientos utilizados. Siguiendo Alizadeh Noughabi & Arghami (2013), se eligió el entero m donde la estadística de prueba \hat{R}_{inf} tuviera mayor potencia. Por lo que, para $n = 10$, se eligió $m = 4$, y con $n = 20$, se eligió $m = 7$.

De las tablas 2.2 y 2.3, se puede ver que existe una fuerte correlación entre los procedimientos usando la generación de muestras mediante muestreo condicional (look-alike) y por medio de remuestreo paramétrico. En general, la estadística \hat{R}_{inf}

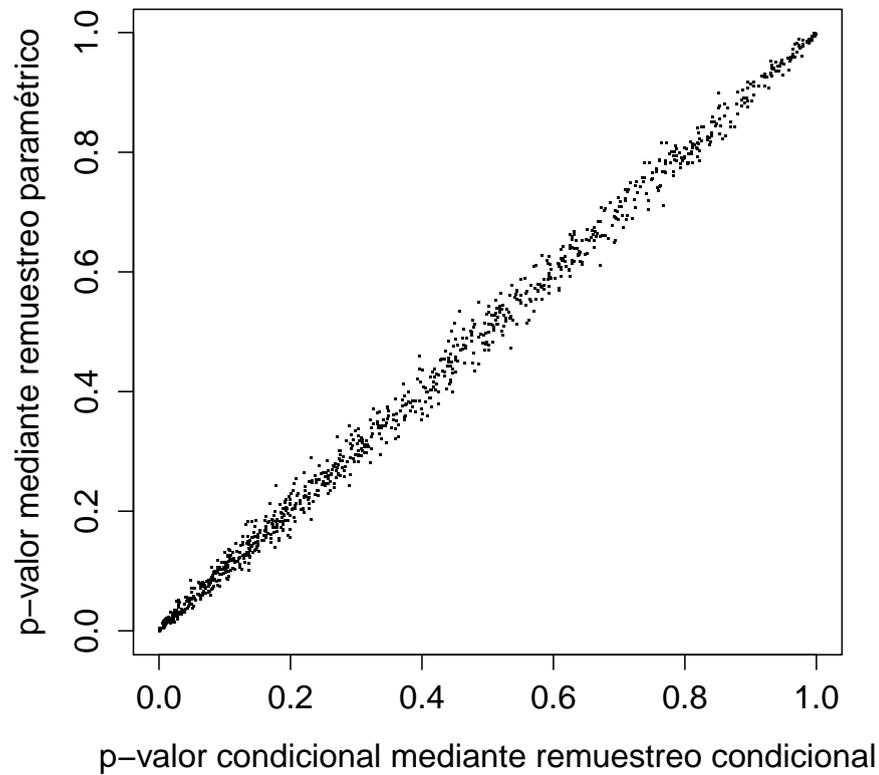


FIGURA 2.6: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la alternativa es $TrN(0.2, 0.3)$ con $n = 10$. $\rho(A^2) = 0.9980035$.

resultó tener mejor potencia que la estadística de prueba Anderson-Darling. Como era de esperarse, para $n = 20$, los resultados para la potencia de la estadística Anderson-Darling mediante la generación de muestras look-alike son muy similares a lo que se reporta en [O'Reilly & Gracia-Medrano \(2006\)](#).

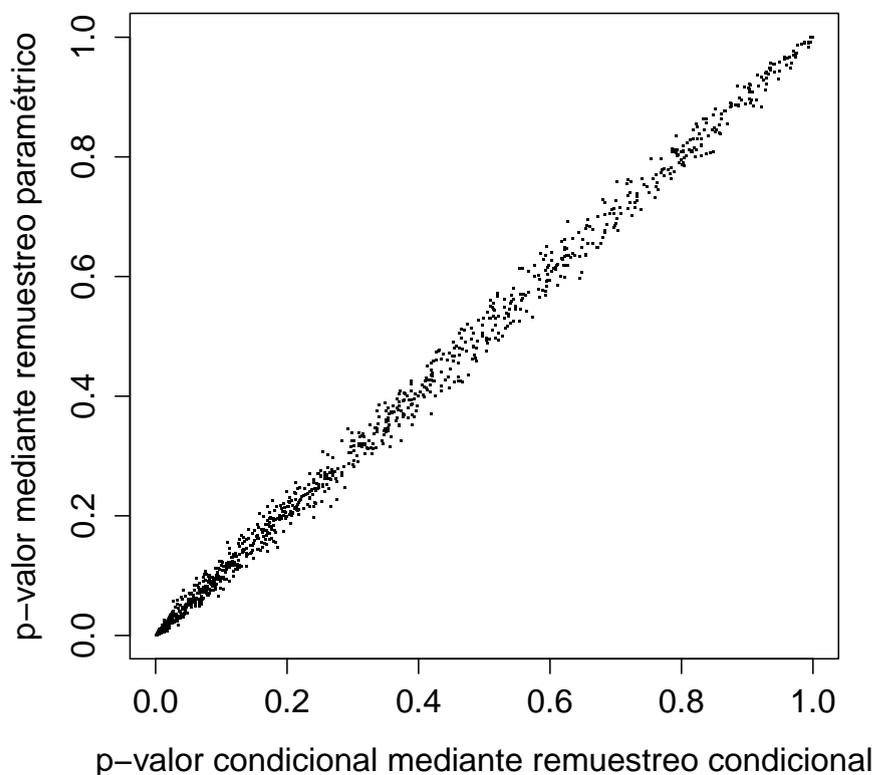


FIGURA 2.7: Bondad de ajuste para probar la exponencial truncada en 1 (caso compuesto). p-valores mediante remuestreo condicional y remuestreo paramétrico para la estadística de prueba A^2 cuando la alternativa es $TrN(0.5, 0.3)$ con $n = 10$. $\rho(A^2) = 0.9981478$.

TABLA 2.2: Potencia de bondad de ajuste de la distribución Gaussiana-inversa (caso compuesto) para diferentes alternativas. Tamaño de muestra $n = 10$ y $\alpha = 0.05$. $\rho(\cdot)$ denota la correlación de los p-valores de las estadísticas mediante look-alike y remuestreo.

Alternativas	look-alike		remuestreo		correlación	
	A	\hat{R}_{inf}	A	\hat{R}_{inf}	$\rho(A)$	$\rho(\hat{R}_{inf})$
$InvGau(1, 1)$	0.056	0.058	0.055	0.044	0.9980022	0.9883003
$Exp(1)$	0.387	0.447	0.393	0.357	0.9986623	0.9743089
$Lognor(0.5, 1)$	0.087	0.096	0.096	0.073	0.9980509	0.9818908
$Beta(2, 5)$	0.189	0.281	0.192	0.255	0.9983146	0.9944399
$Weibull(1, 2)$	0.209	0.259	0.212	0.253	0.9983251	0.9958822
$Gamma(1, 9)$	0.062	0.080	0.066	0.081	0.9976899	0.9976412
$Unif(0, 1)$	0.531	0.689	0.539	0.667	0.9983342	0.9907283
$\chi^2(3)$	0.243	0.294	0.246	0.259	0.9984521	0.9866859

TABLA 2.3: Potencia de bondad de ajuste de la distribución Gaussiana-inversa (caso compuesto) para diferentes alternativas. Tamaño de muestra $n = 20$ y $\alpha = 0.05$. $\rho(\cdot)$ denota la correlación de los p-valores de las estadísticas mediante look-alike y remuestreo.

Alternativas	look-alike		remuestreo		correlación	
	A	\hat{R}_{inf}	A	\hat{R}_{inf}	$\rho(A)$	$\rho(\hat{R}_{inf})$
<i>InvGau</i> (1, 1)	0.037	0.053	0.044	0.039	0.9977807	0.9955343
<i>Exp</i> (1)	0.632	0.661	0.629	0.589	0.9988188	0.9885381
<i>Lognor</i> (0.5, 1)	0.114	0.130	0.117	0.095	0.9983583	0.9939851
<i>Beta</i> (2, 5)	0.412	0.527	0.414	0.506	0.9984390	0.9974811
<i>Weibull</i> (1, 2)	0.429	0.494	0.429	0.476	0.9986409	0.9981379
<i>Gamma</i> (1, 9)	0.069	0.085	0.073	0.081	0.9982374	0.9981969
<i>Unif</i> (0, 1)	0.883	0.965	0.887	0.963	0.9972613	0.9911532
$\chi^2(3)$	0.435	0.498	0.442	0.443	0.9989052	0.9951717

Capítulo 3

Distribución exponencial con datos censurados tipo II

3.1. Introducción

Se dice que una variable aleatoria X tiene distribución exponencial con parámetro (de escala) θ , si,

$$f(x; \theta) = \frac{1}{\theta} \exp\left(\frac{-x}{\theta}\right), \quad x > 0 \text{ y } \theta > 0.$$

Sea X_1, \dots, X_n una muestra aleatoria de una distribución continua F . Las estadísticas de orden se denotan por $X_{(1)} < \dots < X_{(n)}$. En presencia de datos censurados tipo II por la derecha, las observaciones disponibles, son $X_{(1)} < \dots < X_{(r)}$; o sea las r estadísticas de orden menores.

En [O'Reilly & Stephens \(1982\)](#) se hace un estudio de bondad de ajuste de la distribución exponencial. En ese artículo, cuando se está en presencia de datos censurados se proponen algunas transformaciones de los datos con lo que se pasa a un caso sin censura, para posteriormente probar bondad de ajuste para la distribución exponencial. Recientemente en [Zardasht et al. \(2014\)](#) y [Park & Lim \(2015\)](#) se hace un estudio también para probar bondad de ajuste de la distribución exponencial. En presencia de datos censurados en [Park & Lim \(2015\)](#) utilizando los conceptos de entropía y de la información de Kullback-Leibler, proponen una estadística de prueba. En [Zardasht et al. \(2014\)](#), basados en la entropía residual

acumulada (CRE) también definen una estadística de prueba. La idea en el presente trabajo es comparar, los procedimientos propuestos en los artículos mencionados anteriormente.

En la siguiente sección se describe como mediante unas transformaciones en presencia de datos censurados tipo II, se puede “eliminar” la censura. En la sección 3.3 se menciona las estadísticas de pruebas utilizadas para el estudio de simulación que se realizará en la sección 3.4.

3.2. La transformación J y la transformación N

A continuación veremos dos transformaciones de una muestra aleatoria, que nos servirán cuando se está en presencia de datos censurados tipo II y se quiere probar si los datos provienen de una distribución exponencial.

La transformación J muy conocida, es un mapeo de n variables i.i.d exponenciales a $(n - 1)$ variables aleatorias uniformes ya ordenadas (ver por ejemplo O’Reilly & Stephens 1982). De hecho el resultado es parte de una caracterización. Defina

$$Z_{(j)} = \frac{\sum_{i=1}^j X_i}{\sum_{i=1}^n X_i}, \quad j = 1, \dots, n - 1,$$

la caracterización mencionada aparece en el siguiente teorema.

Teorema 3.1. *Sea X_1, \dots, X_n una muestra aleatoria de una distribución continua, entonces dicha muestra proviene de una distribución exponencial con parámetro θ si y sólo si $Z_{(1)}, \dots, Z_{(n-1)}$ están distribuidas como una muestra ordenada de tamaño $(n - 1)$ de una distribución $Unif(0, 1)$.*

Existe otra transformación, también ampliamente conocida debida a Sukhatme (1937) que se denota por N . La transformación N mapea una muestra i.i.d exponencial ya ordenada de tamaño n en una muestra exponencial i.i.d de tamaño n con el mismo parámetro. Definimos,

$$Y_j = (n - (j - 1))(X_{(j)} - X_{(j-1)}), \quad j = 1, \dots, n \text{ y sea } X_{(0)} = 0.$$

Teorema 3.2 (Sukhatme (1937)). *Sea X_1, \dots, X_n una muestra aleatoria de una distribución continua entonces X_1, \dots, X_n tiene una distribución exponencial con parámetro θ si y sólo si Y_1, \dots, Y_n son exponenciales independientes con parámetro θ .*

Por lo anterior, si se quiere probar que los datos siguen una distribución exponencial, primero se aplica la transformación N y posteriormente la transformación J, por lo que se tendrá una muestra aleatoria uniforme ordenada de tamaño $(n - 1)$. Siendo caracterización, ahora lo que se tiene que probar es que los datos finales provienen una distribución uniforme. Uno podría pensar sólo en aplicar la transformación J a los datos, pero como es discutido en [O'Reilly & Stephens \(1982\)](#) se viola el principio de verosimilitud de Fisher.

Si se está en presencia de datos censurados tipo II, se hace algo análogo a lo anterior, es decir si se tiene $X_{(1)} < \dots < X_{(r)}$. Primero a la muestra ordenada se le aplica la transformación N para $j = 1, \dots, r$ y luego a los Y_j 's resultantes, se les aplica la transformación J. Observe que con estas transformaciones, se convierte el problema original de bondad de ajuste compuesto y con censura (tipo II por la derecha) a un problema simple de bondad de ajuste sin censura; de hecho uno de probar uniformidad con una muestra ordenada de tamaño $(r - 1)$.

3.3. Estadísticas de prueba

En esta sección se dan algunas estadísticas de prueba que posteriormente serán utilizadas en la siguiente sección para calcular la potencia de estas estadísticas mediante un estudio de simulación. Se expone una estadística de prueba clásica y algunas que son más recientes y que están basadas en la entropía.

Suponga que se tiene una muestra aleatoria observada, es decir, x_1, \dots, x_n y sea $u_{(i)} = F(x_{(i)}; \hat{\theta})$ con $i = 1, \dots, n$. En [Zhang \(2002\)](#) basados en la estadística de razón de verosimilitudes y en la familia de estadísticas Cramér-von Mises, definen la estadística de prueba,

$$Z = - \sum_{i=1}^n \left[\frac{\log(u_{(i)})}{n - i + \frac{1}{2}} + \frac{\log(1 - u_{(i)})}{i - \frac{1}{2}} \right],$$

que en algunos casos resulta ser más potente que las estadísticas de la familia Cramér-von Mises.

Por otro lado, si se quiere probar que los datos siguen una distribución exponencial, recientemente en [Zardasht et al. \(2014\)](#) basados en la entropía residual acumulada

definen la estadística de prueba,

$$C = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\bar{x}} \exp\left(-\frac{x_i}{\bar{x}}\right),$$

donde $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$. Bajo \mathcal{H}_0 , se demuestra que $R = \sqrt{\frac{382n}{5}}|C - 1/4|$ converge en distribución a una distribución normal estándar. En [Zardasht et al. \(2014\)](#) eligiendo un nivel de significancia α , se rechaza la hipótesis nula, si $R > Z_{1-\frac{\alpha}{2}}$, donde $Z_{1-\frac{\alpha}{2}}$ es el $(1 - \frac{\alpha}{2})$ -percentil de la distribución normal estándar.

En [Park & Lim \(2015\)](#) se aborda el problema de probar exponencialidad con datos censurados tipo II y allí se define la estadística,

$$\begin{aligned} T &= \frac{1}{\hat{\theta}} \sum_{i=0}^{r-1} (x_{(i+1)} - x_{(i)}) \frac{n-i}{r} \log \frac{n-i}{n} + \frac{1}{2\hat{\theta}^2} \sum_{i=0}^{r-1} \frac{n-i}{r} (x_{(i+1)}^2 - x_{(i)}^2) + \\ &+ \frac{n}{r} \left(1 - \exp\left(\frac{x_{(r)}}{\hat{\theta}}\right)\right) - 1, \end{aligned} \tag{3.1}$$

con $\hat{\theta} = \frac{x_{(1)} + \dots + x_{(r)} + (n-r)x_{(r)}}{r}$. Esta estadística esta basada en la teoría de la entropía y la información de Kullback-Leibler. En la siguiente sección se hace un estudio comparativo de simulación, para evaluar potencias de las estadísticas mencionadas junto con la estadística de prueba Anderson-Darling.

3.4. Estudio de potencia

Como se comentó en la sección 3.2, si se está en el caso sin censura o con censura para las estadísticas A^2 y Z , se usaran las transformaciones J y N para probar uniformidad en lugar de probar exponencialidad, y estas pruebas se compararán con estadísticas recientes para probar exponencialidad. Se replicaron algunos escenarios de [Zardasht et al. \(2014\)](#) y [Park & Lim \(2015\)](#). Para el cálculo de los valores críticos de las estadísticas R , A^2 y Z se realizaron 100000 replicas. Para calcular la potencia de las diferentes pruebas se simulan 50000 muestras de las diferentes alternativas.

Las alternativas usadas en este estudio de potencia son la distribución gamma, lognormal, Weibull y Gaussiana-inversa. A continuación, se describen de estas distribuciones con el objeto de identificar la parametrización utilizada.

Se dice que X tiene una distribución gamma con parámetros $\alpha > 0$ y $\beta > 0$, se denota por $X \sim \text{Gamma}(\alpha, \beta)$ si (ver [Forbes et al. 2011](#)),

$$f(x; \alpha, \beta) = \frac{x^{\beta-1} e^{-x/\beta}}{\alpha^\beta \Gamma(\beta)} \quad \text{para } x > 0.$$

Se dice que X tiene una distribución lognormal con parámetros $\mu > 0$ y $\sigma > 0$, se denota por $X \sim \text{Lognor}(\mu, \sigma)$ si,

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2 / 2\sigma^2} \quad \text{para } x > 0.$$

Se dice que X tiene una distribución Weibull con parámetros $\alpha > 0$ y $\beta > 0$, se denota por $X \sim \text{Weibull}(\alpha, \beta)$ si,

$$f(x; \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\} \quad \text{para } x > 0.$$

Se dice que X tiene una distribución Gaussiana-inversa con parámetros $\mu > 0$ y $\lambda > 0$, se denota por $X \sim \text{InvGau}(\mu, \lambda)$ si,

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right\} \quad \text{para } x > 0.$$

Se sabe que si X es una variable aleatoria exponencial, se cumple que $\frac{E(X^2)}{2E(X)} = 1$. Para los escenarios usados en [Zardasht et al. \(2014\)](#), la selección de los parámetros en las distribuciones alternativas fueron elegidos de tal forma que $\frac{E(X^2)}{2E(X)} = 1$. Por lo que, para la distribución gamma se tiene $\alpha^* = \frac{2}{1+\beta}$, para la distribución lognormal $\sigma^* = \sqrt{\frac{2}{3} \ln(2 - \mu)}$, para la distribución Weibull $\alpha^* = \frac{2\Gamma(1+\frac{1}{\beta})}{\Gamma(1+\frac{2}{\beta})}$, por último para la distribución Gaussiana-inversa $\lambda^* = \frac{\mu^2}{2-\mu}$ (ver tabla 3.1). En presencia de datos censurados se usaron los escenarios de [Park & Lim \(2015\)](#) (ver 3.2–3.5).

Si no se tiene censura de la tabla 3.1, se puede ver que la potencia para R , A^2 y Z son muy similares. En el caso que la alternativa es la Gaussiana-inversa, uniforme y χ^2 , las estadísticas A^2 y Z superan a R . En conclusión, se puede decir que es recomendable usar cualesquiera de éstas estadísticas de prueba. Como era de esperarse, los resultados obtenidos para la potencia de la estadística R son muy similares a lo que se reporta en [Zardasht et al. \(2014\)](#).

Por otro lado en presencia de datos censurados tipo II de las tablas 3.2–3.5. En algunos escenarios la prueba T es sesgada, es decir, la potencia está por debajo del

TABLA 3.1: Potencia de bondad de ajuste de la distribución exponencial (λ desconocida) para diferentes alternativas. $n = 10$. $\alpha = 0.05$.

Alternativas	R	A^2	Z
<i>Exp</i> (1)	0.047	0.051	0.049
<i>Gamma</i> (1, 0.5)	0.339	0.419	0.405
<i>Gamma</i> (1, 1.5)	0.119	0.091	0.099
<i>Gamma</i> (1, 2)	0.271	0.197	0.217
<i>Gamma</i> (α^* , 5)	0.929	0.835	0.870
<i>Gamma</i> (α^* , 7)	0.993	0.963	0.976
<i>Lognor</i> (0, 0.6)	0.587	0.422	0.531
<i>Lognor</i> (0, 1)	0.074	0.087	0.119
<i>Lognor</i> (0, 2)	0.666	0.717	0.673
<i>Lognor</i> (0.4, σ^*)	0.948	0.854	0.922
<i>Lognor</i> (0.5, σ^*)	0.997	0.980	0.993
<i>Lognor</i> (0.6, σ^*)	1	1	1
<i>Weibull</i> (1, 0.5)	0.664	0.743	0.720
<i>Weibull</i> (1, 0.8)	0.126	0.161	0.145
<i>Weibull</i> (1, 2)	0.650	0.558	0.554
<i>Weibull</i> (α^* , 2)	0.651	0.563	0.555
<i>Weibull</i> (α^* , 3)	0.976	0.960	0.950
<i>Weibull</i> (α^* , 4)	0.999	0.999	0.998
<i>InvGau</i> (0.1, λ^*)	0.660	0.716	0.680
<i>InvGau</i> (0.2, λ^*)	0.516	0.588	0.549
<i>InvGau</i> (0.3, λ^*)	0.376	0.462	0.427
<i>Unif</i> (0, 1)	0.304	0.409	0.342
χ_1^2	0.342	0.424	0.408

nivel de significancia α . Otra observación interesante de la estadística de prueba T , es que tiene un comportamiento muy extraño o inusual, ya que se supone, que en general, cuando se disminuye la censura debe aumentar la potencia de la prueba, y en algunos escenarios se tiene que en la potencia de T ocurre todo lo contrario, es decir, cuando se disminuye la censura la potencia disminuye. Lo anterior sugiere que conceptualmente hay algo que debe estudiarse en la propuesta de la estadística T , que simplemente referimos aquí como un comportamiento inconsistente. Los resultados obtenidos para la potencia de la estadística T son muy similares a lo que reportan en [Park & Lim \(2015\)](#), y no hacen comentario alguno de estas inconsistencias de la estadística T .

TABLA 3.2: Potencia de bondad de ajuste de la distribución exponencial (λ desconocida) para diferentes alternativas. $\alpha = 0.05$. $n = 10$. Datos censurados.

Puntos de censura Alternativas	$r = 4$			$r = 6$		
	T	A^2	Z	T	A^2	Z
<i>Exp</i> (1)	0.053	0.051	0.052	0.051	0.051	0.051
<i>Gamma</i> (1, 0.5)	0.065	0.245	0.246	0.176	0.312	0.309
<i>Gamma</i> (1, 1.5)	0.094	0.052	0.055	0.078	0.064	0.069
<i>Gamma</i> (1, 2)	0.146	0.078	0.080	0.136	0.112	0.123
<i>Gamma</i> (α^* , 5)	0.431	0.258	0.271	0.527	0.476	0.521
<i>Gamma</i> (α^* , 7)	0.577	0.369	0.389	0.712	0.661	0.708
<i>Lognor</i> (0, 0.6)	0.339	0.192	0.204	0.349	0.306	0.349
<i>Lognor</i> (0, 1)	0.117	0.061	0.064	0.078	0.064	0.073
<i>Lognor</i> (0, 2)	0.037	0.082	0.082	0.153	0.192	0.179
<i>Lognor</i> (0.4, σ^*)	0.556	0.347	0.366	0.649	0.593	0.656
<i>Lognor</i> (0.5, σ^*)	0.715	0.487	0.513	0.838	0.795	0.849
<i>Lognor</i> (0.6, σ^*)	0.928	0.757	0.785	0.989	0.980	0.991
<i>Weibull</i> (1, 0.5)	0.087	0.303	0.303	0.274	0.441	0.432
<i>Weibull</i> (1, 0.8)	0.038	0.084	0.086	0.075	0.103	0.100
<i>Weibull</i> (1, 2)	0.220	0.122	0.127	0.259	0.223	0.237
<i>Weibull</i> (α^* , 2)	0.218	0.124	0.129	0.262	0.226	0.240
<i>Weibull</i> (α^* , 3)	0.437	0.270	0.282	0.601	0.551	0.575
<i>Weibull</i> (α^* , 4)	0.635	0.431	0.449	0.835	0.799	0.818
<i>InvGau</i> (0.1, λ^*)	0.072	0.043	0.046	0.093	0.086	0.094
<i>InvGau</i> (0.2, λ^*)	0.078	0.043	0.046	0.077	0.069	0.077
<i>InvGau</i> (0.3, λ^*)	0.090	0.048	0.051	0.068	0.056	0.067
<i>Unif</i> (0, 1)	0.074	0.054	0.056	0.092	0.077	0.076
χ_1^2	0.066	0.244	0.245	0.183	0.317	0.313

TABLA 3.3: Potencia de bondad de ajuste de la distribución exponencial (λ desconocida) para diferentes alternativas. $\alpha = 0.05$. $n = 10$. Datos censurados.

Puntos de censura Alternativas	$r = 8$			$r = 9$		
	T	A^2	Z	T	A^2	Z
<i>Exp</i> (1)	0.050	0.050	0.049	0.049	0.050	0.049
<i>Gamma</i> (1, 0.5)	0.296	0.363	0.355	0.324	0.393	0.380
<i>Gamma</i> (1, 1.5)	0.034	0.079	0.085	0.012	0.084	0.091
<i>Gamma</i> (1, 2)	0.058	0.151	0.166	0.009	0.176	0.193
<i>Gamma</i> (α^* , 5)	0.387	0.674	0.727	0.105	0.759	0.806
<i>Gamma</i> (α^* , 7)	0.604	0.860	0.898	0.226	0.923	0.948
<i>Lognor</i> (0, 0.6)	0.173	0.383	0.457	0.029	0.407	0.499
<i>Lognor</i> (0, 1)	0.043	0.060	0.079	0.063	0.063	0.086
<i>Lognor</i> (0, 2)	0.422	0.401	0.369	0.573	0.553	0.505
<i>Lognor</i> (0.4, σ^*)	0.467	0.750	0.833	0.135	0.812	0.884
<i>Lognor</i> (0.5, σ^*)	0.729	0.931	0.967	0.315	0.962	0.985
<i>Lognor</i> (0.6, σ^*)	0.981	0.999	1	0.786	1	1
<i>Weibull</i> (1, 0.5)	0.491	0.578	0.564	0.579	0.659	0.637
<i>Weibull</i> (1, 0.8)	0.132	0.124	0.118	0.160	0.139	0.129
<i>Weibull</i> (1, 2)	0.166	0.358	0.380	0.034	0.451	0.460
<i>Weibull</i> (α^* , 2)	0.168	0.364	0.385	0.034	0.452	0.457
<i>Weibull</i> (α^* , 3)	0.536	0.798	0.811	0.208	0.895	0.890
<i>Weibull</i> (α^* , 4)	0.833	0.966	0.967	0.505	0.992	0.990
<i>InvGau</i> (0.1, λ^*)	0.362	0.295	0.284	0.570	0.493	0.464
<i>InvGau</i> (0.2, λ^*)	0.285	0.213	0.211	0.467	0.375	0.352
<i>InvGau</i> (0.3, λ^*)	0.220	0.154	0.160	0.370	0.271	0.259
<i>Unif</i> (0, 1)	0.068	0.152	0.137	0.015	0.237	0.199
χ_1^2	0.296	0.362	0.356	0.323	0.393	0.379

TABLA 3.4: Potencia de bondad de ajuste de la distribución exponencial (λ desconocida) para diferentes alternativas. $\alpha = 0.05$. $n = 20$. Datos censurados.

Puntos de censura Alternativas	$r = 8$			$r = 12$		
	T	A^2	Z	T	A^2	Z
<i>Exp</i> (1)	0.049	0.051	0.053	0.051	0.050	0.051
<i>Gamma</i> (1, 0.5)	0.176	0.398	0.387	0.327	0.502	0.485
<i>Gamma</i> (1, 1.5)	0.125	0.085	0.096	0.111	0.110	0.124
<i>Gamma</i> (1, 2)	0.234	0.169	0.188	0.243	0.250	0.288
<i>Gamma</i> (α^* , 5)	0.811	0.721	0.778	0.906	0.925	0.954
<i>Gamma</i> (α^* , 7)	0.940	0.891	0.926	0.985	0.990	0.996
<i>Lognor</i> (0, 0.6)	0.680	0.574	0.660	0.697	0.735	0.843
<i>Lognor</i> (0, 1)	0.174	0.121	0.151	0.108	0.112	0.167
<i>Lognor</i> (0, 2)	0.070	0.128	0.115	0.293	0.344	0.290
<i>Lognor</i> (0.4, σ^*)	0.934	0.879	0.932	0.969	0.979	0.995
<i>Lognor</i> (0.5, σ^*)	0.990	0.974	0.989	0.998	0.999	1
<i>Lognor</i> (0.6, σ^*)	1	1	1	1	1	1
<i>Weibull</i> (1, 0.5)	0.250	0.505	0.493	0.513	0.693	0.667
<i>Weibull</i> (1, 0.8)	0.055	0.112	0.107	0.104	0.143	0.127
<i>Weibull</i> (1, 2)	0.402	0.311	0.334	0.524	0.544	0.555
<i>Weibull</i> (α^* , 2)	0.399	0.311	0.336	0.523	0.544	0.558
<i>Weibull</i> (α^* , 3)	0.803	0.717	0.744	0.938	0.950	0.952
<i>Weibull</i> (α^* , 4)	0.959	0.924	0.936	0.997	0.998	0.998
<i>InvGau</i> (0.1, λ^*)	0.082	0.062	0.090	0.150	0.134	0.157
<i>InvGau</i> (0.2, λ^*)	0.092	0.066	0.096	0.114	0.100	0.132
<i>InvGau</i> (0.3, λ^*)	0.108	0.077	0.109	0.091	0.077	0.121
<i>Unif</i> (0, 1)	0.085	0.063	0.063	0.127	0.119	0.102
χ_1^2	0.175	0.397	0.391	0.327	0.499	0.480

TABLA 3.5: Potencia de bondad de ajuste de la distribución exponencial (λ desconocida) para diferentes alternativas. $\alpha = 0.05$. $n = 20$. Datos censurados.

Puntos de censura Alternativas	$r = 16$			$r = 18$		
	T	A^2	Z	T	A^2	Z
<i>Exp</i> (1)	0.051	0.051	0.050	0.051	0.050	0.049
<i>Gamma</i> (1, 0.5)	0.462	0.586	0.565	0.495	0.616	0.599
<i>Gamma</i> (1, 1.5)	0.063	0.142	0.159	0.022	0.150	0.172
<i>Gamma</i> (1, 2)	0.157	0.341	0.384	0.060	0.386	0.435
<i>Gamma</i> (α^* , 5)	0.889	0.986	0.992	0.703	0.995	0.998
<i>Gamma</i> (α^* , 7)	0.986	1	1	0.93	1	1
<i>Lognor</i> (0, 0.6)	0.536	0.817	0.921	0.266	0.825	0.939
<i>Lognor</i> (0, 1)	0.060	0.090	0.162	0.085	0.081	0.168
<i>Lognor</i> (0, 2)	0.662	0.670	0.588	0.821	0.827	0.769
<i>Lognor</i> (0.4, σ^*)	0.945	0.996	1	0.792	0.998	1
<i>Lognor</i> (0.5, σ^*)	0.998	1	1	0.976	1	1
<i>Lognor</i> (0.6, σ^*)	1	1	1	1	1	1
<i>Weibull</i> (1, 0.5)	0.741	0.841	0.820	0.823	0.897	0.884
<i>Weibull</i> (1, 0.8)	0.177	0.179	0.156	0.224	0.211	0.182
<i>Weibull</i> (1, 2)	0.504	0.765	0.763	0.300	0.858	0.846
<i>Weibull</i> (α^* , 2)	0.503	0.765	0.762	0.303	0.854	0.842
<i>Weibull</i> (α^* , 3)	0.963	0.997	0.996	0.895	1	0.999
<i>Weibull</i> (α^* , 4)	0.999	1	1	0.997	1	1
<i>InvGau</i> (0.1, λ^*)	0.587	0.518	0.481	0.817	0.763	0.717
<i>InvGau</i> (0.2, λ^*)	0.469	0.387	0.365	0.718	0.638	0.593
<i>InvGau</i> (0.3, λ^*)	0.363	0.277	0.279	0.595	0.486	0.451
<i>Unif</i> (0, 1)	0.167	0.305	0.239	0.100	0.504	0.403
χ_1^2	0.461	0.580	0.564	0.498	0.620	0.604

Capítulo 4

Enfoque Bayesiano

En este capítulo se exploran dos artículos recientes donde utilizan un enfoque de la teoría estadística Bayesiana para probar bondad de ajuste. [He et al. \(2013\)](#) usa el enfoque paramétrico, es decir, utiliza una distribución inicial en los parámetros y se basa en la noción del “p-valor predictivo posterior”. Mientras que [Al Labadi & Zarepour \(2014\)](#) utiliza un enfoque de estadística Bayesiana no paramétrica, es decir, se utiliza una distribución inicial en un espacio no paramétrico; particularmente utiliza el proceso de Dirichlet. En las siguientes dos secciones se ven los dos enfoques mencionados.

4.1. Uso del p-valor predictivo posterior

En esta sección se verá lo publicado en [He et al. \(2013\)](#). En dicho trabajo se utiliza un enfoque Bayesiano y se ve la definición del p-valor predictivo posterior. Posteriormente, se aborda el problema de bondad de ajuste para la distribución normal y hacen un pequeño estudio de simulación. Aquí se replica lo anterior, se incorpora la estadística de Anderson-Darling y se incorporan otras alternativas (distribuciones) en la simulación.

Desde el contexto Bayesiano, del primer autor que se tenga conocimiento que uso la distribución predictiva final (es decir, la observación futura) para probar bondad de ajuste fue [Guttman \(1967\)](#). Ahí se propone una estadística tipo ji-cuadrada, la cual mide la discrepancia entre las frecuencias observadas y la distribución predictiva final (o a posteriori). En [O’Reilly & Villegas \(1987\)](#) estudian dos transformaciones; una desde el contexto clásico y la otra desde el contexto Bayesiano;

bajo algunos supuestos, estas transformaciones resultan ser equivalentes; estos supuestos se cumplen en el caso de normalidad y cuando la distribución inicial es la de Jeffreys. La primera transformación es la Rao-Blackwell y la segunda es una nueva transformación y es llamada PPIT (por sus siglas en inglés, predictive probability integral transformation), la cual, es una ligera modificación de la distribución predictiva final. Ésta nueva transformación, resulta ser equivalente a la transformación Rao-Blackwell. Consecuentemente, si ahora, se interpreta una observación futura si proviene de la transformación PPIT, se tendrá que el p-valor predictivo posterior y el p-valor condicional son equivalentes. Aquí, no se usará la transformación PPIT, pero se considera que es importante mencionar esta relación. A continuación se dará la definición de p-valor predictivo posterior tal como se ha utilizado en artículos Bayesianos.

El p-valor predictivo posterior se basa en el uso del concepto de p-valor clásico. Fue introducido por [Guttman \(1967\)](#) y [Rubin \(1984\)](#); y fue definido formalmente por [Meng \(1994\)](#) y [Gelman et al. \(1996\)](#).

Sea D una estadística de prueba, la cual, es función de la muestra \underline{x} de tamaño n ; y cuya distribución depende del parámetro θ , por conveniencia utilizando notación de los autores, se denotará por $D(\underline{x}, \theta)$. En [Gelman et al. \(1996\)](#), el p-valor predictivo posterior es definido como:

$$p_b(\underline{x}) = P(D(\underline{x}^{rep}, \theta) > D(\underline{x}, \theta) | \underline{x}),$$

donde \underline{x}^{rep} se entiende como una observación futura de tamaño n , replica de \underline{x} , y la probabilidad es tomada sobre la distribución posterior conjunta de $(\underline{x}^{rep}, \theta)$, es decir,

$$f(\underline{x}^{rep} | \theta) \pi(\theta | \underline{x})$$

donde la distribución

$$\int f(\underline{x}^{rep} | \theta) \pi(\theta | \underline{x}) d\theta$$

es la llamada distribución predictiva posterior.

Análogamente al cálculo del p-valor clásico, el p-valor predictivo posterior puede ser calculado vía simulación de Monte Carlo. Específicamente, se simula θ_j , $j = 1, \dots, M$ de la distribución posterior de θ . Dado θ_j , se simula una replica del conjunto de datos, por decir \underline{y}_j de la distribución $f(\cdot | \theta_j)$. Se puede aproximar por

la frecuencia relativa del evento $\{D(\underline{x}^{rep}, \theta) > D(\underline{x}, \theta)\}$, es decir,

$$p_b(\underline{x}) \approx \frac{1}{M} \sum_{j=1}^M I\{D(\underline{y}_j, \theta_j) > D(\underline{x}, \theta_j)\} \quad (4.1)$$

donde I es la función indicadora. Entre más grande sea el valor de M , mejor será la aproximación al p-valor predictivo posterior.

4.1.1. Distribución normal

Suponga que se quiere probar normalidad. Para utilizar el enfoque Bayesiano objetivo en He et al. (2013) se eligió una distribución no informativa, es este caso, la a priori de Jeffreys para (μ, σ) es,

$$\pi(\mu, \sigma) \propto \frac{1}{\sigma}.$$

Condicionando en la muestra observada $\underline{x} = (x_1, \dots, x_n)$, la distribución posterior de (μ, σ) es

$$[\sigma | \underline{x}] \sim \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\chi_{n-1}^2}}, \quad [\mu | \sigma, \underline{x}] \sim N(\bar{x}, \frac{\sigma^2}{n}),$$

en donde \bar{x} es la media muestral y χ_{n-1}^2 es una variable aleatoria con distribución ji-cuadrada con $n - 1$ grados de libertad.

En He et al. (2013) usan la estadística clásica de Kolmogorov-Smirnov que está dada por:

$$K(\underline{x}, \theta) = \sqrt{n} \max_{1 \leq i \leq n} \left\{ \left| \frac{i}{n} - F_0(x_{(i)}) \right|, \left| \frac{i-1}{n} - F_0(x_{(i)}) \right| \right\},$$

y utilizan otra estadística de prueba que en general resulta ser aún menos potente, la cual, no se incluye en el pequeño estudio de simulación que aparece mas adelante. Es bien sabido que para probar normalidad, una prueba más potente que la estadística Kolmogorov-Smirnov es la estadística Anderson-Darling que se vió en la sección 2.1, la cual, se incluirá en este estudio. Los autores comentan que el uso del p-valor predictivo posterior para la estadística Anderson-Darling no fue satisfactorio, por lo que no se incluyó en su estudio. Es importante notar que los autores no incluyeron la estadística de Anderson-Darling utilizada en el contexto clásico; la cual como se verá en este estudio resulta ser más potente que el uso del p-valor predictivo posterior para la estadística Kolmogorov-Smirnov.

El p-valor predictivo posterior para la estadística Kolmogorov-Smirnov se denotará por PK , es decir,

$$PK(\underline{x}) = P(K(\underline{x}^{rep}, \theta) > K(\underline{x}, \theta) | \underline{x}).$$

En [He et al. \(2013\)](#) se demuestra que la distribución de $PK(\underline{X})$ es invariante bajo transformaciones de localización y/o escala, y también se demuestra que es consistente.

4.1.2. Estudio de potencia

En la tabla [4.1](#) se reportan los valores críticos de la estadística PK bajo el supuesto de normalidad a un nivel de significancia del 5% obtenido de [He et al. \(2013\)](#). A continuación se da el algoritmo que se usa para calcular la potencia de la estadística PK .

1. Generar una muestra $\underline{x} = (x_1, \dots, x_n)$ de la distribución alternativa, calcular la media \bar{x} y la varianza s^2 .
2. Dada la muestra \underline{x} , bajo \mathcal{H}_0 , generar una variable aleatoria σ^2 de la distribución de $\frac{(n-1)s^2}{\chi_{n-1}^2}$; luego generar una variable aleatoria μ de $N(\bar{x}, \frac{\sigma^2}{n})$. Y por último, se genera la muestra $\underline{y} = (y_1, \dots, y_n)$ de $N(\mu, \sigma)$. Calcular $K(\underline{x}, \mu, \sigma^2)$ y $K(\underline{y}, \mu, \sigma^2)$.
3. Repetir el paso (2) 1000 veces; se tendrán $K(\underline{x}, \mu^j, \sigma^{2,j})$ y $K(\underline{y}^j, \mu^j, \sigma^{2,j})$ con $j = 1, \dots, 1000$. Calcular el p-valor predictivo posterior PK como en ecuación [\(4.1\)](#).
4. Si $PK < \text{valor crítico}$ (ver tabla [4.1](#)), se rechaza \mathcal{H}_0 .
5. Repetir del paso (1) a (4) 1000 veces, y calcular la proporción de rechazos.

Es este estudio se usarán algunas alternativas utilizadas en [He et al. \(2013\)](#) y adicionalmente se utilizan como alternativas la distribución normal estándar, la gamma, la Weibull, la Gaussiana-inversa, la t-Student con k grados de libertad que se denota por $t(k)$, y por último, la distribución χ^2 con k grados de libertad, se denota por $\chi^2(k)$. En [He et al. \(2013\)](#) se usa la distribución normal estándar truncada en a y b que se denota por $TrN(a, b)$.

TABLA 4.1: Valores críticos para la estadística PK bajo normalidad con $\alpha = 0.05$.

n	10	20	30	40	50
PK	0.2930	0.2853	0.2815	0.2841	0.2808

TABLA 4.2: Potencia de bondad de ajuste de la distribución normal (caso compuesto) para diferentes alternativas. $n = 10$. $\alpha = 0.05$.

Alternativas	K	PK	A^2
$N(0, 1)$	0.043	0.047	0.046
$Unif(0, 1)$	0.068	0.064	0.081
$Unif(0, 4)$	0.064	0.063	0.065
$t(1)$	0.574	0.600	0.608
$t(2)$	0.269	0.286	0.298
$t(7)$	0.079	0.088	0.095
$\chi^2(1)$	0.548	0.668	0.700
$\chi^2(2)$	0.282	0.346	0.392
$\chi^2(4)$	0.170	0.178	0.218
$Beta(1, 1)$	0.066	0.064	0.079
$Beta(4, 1.5)$	0.089	0.088	0.097
$TrN(-1, 1)$	0.042	0.038	0.044
$TrN(-1.5, 1.5)$	0.031	0.032	0.034
$Gamma(5, 1)$	0.082	0.100	0.100
$Gamma(9, 0.5)$	0.076	0.082	0.086
$Weibull(1, 1.5)$	0.309	0.371	0.425
$Weibull(1, 5)$	0.284	0.360	0.413
$InvGau(1, 3)$	0.212	0.224	0.261
$InvGau(3, 1)$	0.635	0.722	0.759

En las tablas 4.2 y 4.3 se reporta la potencia de las estadísticas K , PK y A^2 para diferentes alternativas con $n = 10$ y $n = 20$ a un nivel de significancia del 5%. Al igual que He et al. (2013), en éstas tablas, se puede ver que para la mayoría de las alternativas la estadística PK es mejor que K . Mientras que la estadística Anderson-Darling A^2 tiene mejor potencia que PK .

En conclusión, para probar normalidad se recomienda usar la estadística de prueba A^2 en lugar de la estadística PK . Otra ventaja de usar A^2 es que no necesita tantos recursos computacionales, como es el caso de PK , que requiere un mayor esfuerzo.

TABLA 4.3: Potencia de bondad de ajuste de la distribución normal (caso compuesto) para diferentes alternativas. $n = 20$. $\alpha = 0.05$.

Alternativas	K	PK	A^2
$N(0, 1)$	0.051	0.049	0.048
$Unif(0, 1)$	0.097	0.107	0.184
$Unif(0, 4)$	0.099	0.106	0.171
$t(1)$	0.837	0.868	0.876
$t(2)$	0.484	0.521	0.562
$t(7)$	0.090	0.106	0.120
$\chi^2(1)$	0.891	0.951	0.972
$\chi^2(2)$	0.607	0.691	0.790
$\chi^2(4)$	0.346	0.370	0.488
$Beta(1, 1)$	0.089	0.092	0.172
$Beta(4, 1.5)$	0.153	0.150	0.227
$TrN(-1, 1)$	0.059	0.060	0.081
$TrN(-1.5, 1.5)$	0.044	0.037	0.056
$Gamma(5, 1)$	0.161	0.164	0.210
$Gamma(9, 0.5)$	0.100	0.099	0.128
$Weibull(1, 1.5)$	0.611	0.718	0.807
$Weibull(1, 5)$	0.582	0.669	0.773
$InvGau(1, 3)$	0.397	0.415	0.546
$InvGau(3, 1)$	0.946	0.973	0.982

4.2. Proceso de Dirichlet y su medida base

En esta sección se expone un enfoque de la estadística Bayesiana no paramétrica utilizada para probar bondad de ajuste. Se utiliza una distribución inicial en un espacio no paramétrico; particularmente se usa el proceso de Dirichlet que fue definido en [Ferguson \(1973\)](#). El proceso de Dirichlet es una familia de procesos estocásticos cuyas realizaciones son distribuciones de probabilidad. Este proceso es muy usado en inferencia Bayesiana no paramétrica para describir el conocimiento a priori acerca de la distribución de las variables aleatorias. Una propiedad muy importante del proceso de Dirichlet, es que es una familia conjugada, es decir, la distribución final pertenece a la misma clase que la inicial. Hay pocos trabajos donde se usa el proceso de Dirichlet para probar bondad de ajuste. Dos artículos para probar bondad de ajuste en el caso continuo donde se utiliza el proceso de Dirichlet es en [Muliere & Tardella \(1998\)](#) y [Swartz \(1999\)](#). En estos trabajos se usa con la

distancia de Kolmogorov, y se mide la distancia entre la distribución a posteriori y la distribución inicial propuesta. En particular, se usa el proceso de Dirichlet como a priori para la verdadera distribución. Sin embargo, expresiones explícitas para calcular la distancia no son proporcionadas. Recientemente en [Al Labadi & Zarepour \(2014\)](#) se dan expresiones explícitas para calcular la distancia entre el proceso de Dirichlet y su medida base, usando la distancia de Kolmogorov y la distancia Cramér-von Mises. Estos autores mediante el cálculo de la distancia del proceso de Dirichlet y su correspondiente medida base proponen un nuevo procedimiento para probar bondad de ajuste. En la siguiente sección se ve la definición de un proceso de Dirichlet. En la sección 4.2.2 se aborda como calcular la distancia de Kolmogorov y la distancia Cramér-von Mises para el proceso de Dirichlet y su medida base. Y finalmente, en la sección 4.2.3 se aborda el problema de bondad de ajuste.

4.2.1. Medida de probabilidad aleatoria y el proceso de Dirichlet

Sea $(\mathfrak{X}, \mathcal{A})$ un espacio medible. Se dice que P es una medida de probabilidad aleatoria en $(\mathfrak{X}, \mathcal{A})$, si para cada $A \in \mathcal{A}$, $P(A)$ toma valores sólo en $[0, 1]$, y $P(\mathfrak{X})$ es degenerado a 1. Además, si A_1, \dots, A_l y B_1, \dots, B_k son particiones medibles de \mathfrak{X} , y si A_1, \dots, A_l es un refinamiento de B_1, \dots, B_k con $B_1 = \cup_{j=1}^{r_1} A_j$, $B_2 = \cup_{j=r_1+1}^{r_2} A_j$, \dots , $B_k = \cup_{j=r_{k-1}+1}^l A_j$, entonces la distribución de

$$\left(\sum_{j=1}^{r_1} P(A_j), \sum_{j=r_1+1}^{r_2} P(A_j), \dots, \sum_{j=r_{k-1}+1}^l P(A_j) \right)$$

es igual a la distribución de $(P(B_1), \dots, P(B_k))$.

En [Ferguson \(1973\)](#) se demuestra el siguiente resultado. Sea P una medida de probabilidad aleatoria, y sean las distribuciones conjuntas de las variables aleatorias $(P(A_1), \dots, P(A_m))$ para cada m y para cualesquiera conjuntos medibles A_1, \dots, A_m entonces existe una medida de probabilidad \mathcal{P} en $([0, 1]^{\mathcal{A}}, \mathcal{BF}^{\mathcal{A}})$ que producen éstas distribuciones conjuntas (teorema de consistencia de Kolmogorov). Donde, $[0, 1]^{\mathcal{A}}$ representa el espacio de todas las funciones de \mathcal{A} en $[0, 1]$, y $\mathcal{BF}^{\mathcal{A}}$ representa la σ -álgebra generada por el álgebra de los conjuntos cilindros.

Sea P es una medida de probabilidad aleatoria en $(\mathfrak{X}, \mathcal{A})$. Se dice que X_1, \dots, X_m es una muestra de tamaño m de P , si para cualquier $n = 1, 2, \dots$ y conjuntos

medibles $A_1, \dots, A_n, B_1, \dots, B_m$,

$$\mathcal{P}(X_1 \in B_1, \dots, X_m \in B_m | P(A_1), \dots, P(A_n), P(B_1), \dots, P(B_m)) = \prod_{j=1}^m P(B_j),$$

esta igualdad es casi seguramente (a.s).

Se dice que X_1, \dots, X_m es una muestra de tamaño m de P significa (aproximadamente), si dadas $P(B_1), \dots, P(B_m)$, los eventos $\{X_1 \in B_1\}, \dots, \{X_m \in B_m\}$ son independientes del resto del proceso, y son independientes entre ellos, con $\mathcal{P}(X_j \in B_j | P(B_1), \dots, P(B_m)) = P(B_j)$ casi seguramente para $j = 1, \dots, m$ (ver [Ferguson 1973](#)).

A continuación se ve la distribución de Dirichlet para posteriormente definir el proceso de Dirichlet.

Se dice que un vector aleatorio $X = (X_1, \dots, X_k)$ de orden $k \geq 2$, tiene una distribución de Dirichlet con parámetros $\alpha_1, \dots, \alpha_k > 0$, denotado por $X \sim Dir(\alpha_1, \dots, \alpha_k)$, si

$$f(x_1, \dots, x_k; \alpha_1, \dots, \alpha_k) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k x_i^{\alpha_i-1},$$

con $0 < x_i < 1$, $i = 1, \dots, k$ y $\sum_{i=1}^k x_i = 1$.

Una posibilidad, si se quiere generar un vector aleatorio $X = (X_1, \dots, X_k)$ de una distribución Dirichlet con parámetros $\alpha_1, \dots, \alpha_k > 0$. Primero se generan $Y_i \sim Gamma(\alpha_i, 1)$, $i = 1, \dots, k$, independientes. Entonces

$$X = (X_1, \dots, X_k) = \left(\frac{Y_1}{\sum_{i=1}^k Y_i}, \dots, \frac{Y_k}{\sum_{i=1}^k Y_i} \right) \sim Dir(\alpha_1, \dots, \alpha_k).$$

Como antes, sea $(\mathfrak{X}, \mathcal{A})$ un espacio medible y H una medida de probabilidad en $(\mathfrak{X}, \mathcal{A})$. Sea $a > 0$ arbitrario, una medida de probabilidad aleatoria $P = \{P(A)\}_{A \in \mathcal{A}}$ es llamado un proceso de Dirichlet con parámetros a y H en $(\mathfrak{X}, \mathcal{A})$, si para cualquier partición finita $\{A_1, \dots, A_k\}$ de \mathfrak{X} , el vector aleatorio $(P(A_1), \dots, P(A_k))$ tiene una distribución Dirichlet con parámetros $(aH(A_1), \dots, aH(A_k))$, donde $k \geq 2$. El parámetro a es conocido como el parámetro de concentración y la medida de probabilidad H es conocida como la medida base de P . Si P es un proceso de Dirichlet con parámetro de concentración a y medida base H , se denotará por

$P \sim DP(a, H)$. Se asume que si $H(A_j) = 0$, entonces $P(A_j) = 0$ con probabilidad 1.

Ferguson (1973) demuestra que el proceso Dirichlet tiene la propiedad de conjugación (o sea la final está en la misma clase que la inicial), es decir, si X_1, \dots, X_m es una muestra de $P \sim DP(a, H)$, entonces la distribución posterior de P dado X_1, \dots, X_m es un proceso de Dirichlet con parámetro de concentración a^* y medida base H^* , con

$$a^* = a + m, \quad H^* = \frac{a}{a + m}H + \frac{m}{a + m} \frac{\sum_{i=1}^m \delta_{X_i}}{m}, \quad (4.2)$$

donde δ_X denota la medida de Dirac, es decir, $\delta_X(A) = 1$ si $X \in A$ y 0 en otro caso. Se usará la notación “*” en el superíndice para denotar cantidades de la distribución posterior. En la estadística clásica, se define el proceso empírico y se demuestra que dicho proceso converge a un Browniano atado. Lo (1987), similarmente demostró, cuando $m \rightarrow \infty$, que el proceso de Dirichlet escalado y centrado $\sqrt{m}(P^* - H^*)$ converge al Browniano atado en $D[0, 1]$ con respecto a la topología de Skorohod.

Una aproximación al proceso de Dirichlet puede ser obtenido mediante modelos de mezclas finitas. En particular, Ishwaran & Zarepour (2002) prueban que, cuando $n \rightarrow \infty$, el proceso finito dimensional

$$P_n^{F.D.}(\cdot) = \sum_{i=1}^n p_{i,n} \delta_{\theta_i}(\cdot) \quad (4.3)$$

converge en distribución a $P \sim DP(a, H)$, donde $(\theta_i)_{1 \leq i \leq n}$ es una sucesión de variables aleatorias con distribución común H y $(p_{1,n}, \dots, p_{n,n})$ tiene una distribución Dirichlet con parámetros $(a/n, \dots, a/n)$ e independiente de $(\theta_i)_{1 \leq i \leq n}$. Para generar $(p_{1,n}, \dots, p_{n,n})$ en la ecuación 4.3, se puede definir $p_{i,n}$ como:

$$p_{i,n} = \frac{G_{i,n}}{\sum_{i=1}^n G_{i,n}}$$

donde $(G_{i,n})_{1 \leq i \leq n}$ son variables aleatorias i.i.d con distribución $Gamma(a/n, 1)$ e independientes de $(\theta_i)_{1 \leq i \leq n}$. Existen otras aproximaciones al proceso de Dirichlet, pero aquí solo se usará la aproximación mencionada anteriormente.

4.2.2. Distancia entre el proceso de Dirichlet y su medida base

En [Al Labadi & Zarepour \(2014\)](#) se derivan expresiones de la distancia de Kolmogorov y Cramér-von Mises, las cuales miden la discrepancia entre el proceso de Dirichlet y su medida base, a continuación se exponen dichas distancias.

Se usará la notación de medida base y su correspondiente función de distribución, es decir, $P(x) = P((-\infty, x])$, $H(x) = H((-\infty, x])$ y $\mathfrak{X} = \mathbb{R}$. Sea H una distribución continua, y $(\theta_k)_{1 \leq k \leq n}$ variables aleatorias independientes con distribución común H . Sea $P_n = \sum_{k=1}^n J_k \delta_{\theta_k}$ la medida de probabilidad discreta definida como en [4.3](#). Se denotará por $\theta_{(1)}, \dots, \theta_{(n)}$ las estadísticas de orden y se define $\theta_{(0)} = -\infty$. La distancia de Kolmogorov entre la medida de probabilidad y la medida base es,

$$K(P_n, H) = \max\{K^{(1)}, K^{(2)}\},$$

donde

$$K^{(1)} = \max_{1 \leq i \leq n} |P_n(\theta_{(i-1)}) - H(\theta_{(i)})|, \quad K^{(2)} = \max_{1 \leq i \leq n} |P_n(\theta_{(i)}) - H(\theta_{(i)})|.$$

De forma directa se puede ver que

$$K^{(1)} = \max_{1 \leq i \leq n} \left| \sum_{k=1}^{i-1} J'_k - H(\theta_{(i)}) \right|, \quad K^{(2)} = \max_{1 \leq i \leq n} \left| \sum_{k=1}^i J'_k - H(\theta_{(i)}) \right|,$$

donde $(J'_i)_{1 \leq i \leq n}$ son los pesos reordenados asociados con $(\theta_{(i)})_{1 \leq i \leq n}$ tal que $\sum_{i=1}^n J'_i \delta_{\theta_{(i)}} = \sum_{i=1}^n J_i \delta_{\theta_i}$. Con la convención que $\sum_{k=1}^0 J'_k = 0$.

Por otro lado, la distancia Cramér-von Mises entre dos funciones de distribución F y G está dado por,

$$CvM(F, G) = \int_{-\infty}^{\infty} (F(x) - G(x))^2 g(x) dx,$$

donde $g(x) = G'(x)$ es la función de densidad con respecto a la medida de Lebesgue.

En [Al Labadi & Zarepour \(2014\)](#) se deriva una forma explícita de la distancia de Cramér-von Mises entre el proceso de Dirichlet y su medida base que a continuación enunciamos.

Lema 4.1. *Sea H una función de distribución continua, y $(\theta_k)_{1 \leq k \leq n}$ variables aleatorias independientes con distribución común H . Sea $P_n = \sum_{k=1}^n J_k \delta_{\theta_k}$ una medida*

de probabilidad discreta. Sea $\theta_{(1)}, \dots, \theta_{(n)}$ las estadísticas de orden de $\theta_1, \dots, \theta_n$, y J_1', \dots, J_n' los correspondientes tamaños de saltos. Entonces

$$CvM(P_n, H) = \frac{1}{3} + \sum_{i=1}^n J_i' H^2(\theta_{(i)}) - \sum_{i=1}^n J_i'^2 H(\theta_{(i)}) - 2 \sum_{i=1}^n \left[J_i' H(\theta_{(i)}) \sum_{k=1}^{i-1} J_k' \right].$$

En Al Labadi & Zarepour (2014) se demostró que para cualquier métrica de probabilidad d , la distancia $d(P_n^{F.D.}, H)$ converge en distribución a la distancia $d(P, H)$, donde $P_n^{F.D.}$ como en la ecuación 4.3, P es el proceso de Dirichlet y H su correspondiente medida base.

4.2.3. Bondad de ajuste

Se sabe que el enfoque clásico y Bayesiano no paramétrico tienen fundamentos diferentes. A continuación se discuten los dos enfoques para probar bondad de ajuste continuo en el caso simple.

En las pruebas de bondad de ajuste desde el perspectiva clásica, se supone que se tiene una muestra aleatoria X_1, X_2, \dots, X_m con distribución común F , y se está interesado en probar $\mathcal{H}_0 : F = F_0$ vs. $\mathcal{H}_1 : F \in (\mathcal{F} - \{F_0\})$, donde F_0 , es una función de distribución totalmente conocida y \mathcal{F} , es la familia no paramétrica de todas las distribuciones continuas (por ejemplo en \mathbb{R} o \mathbb{R}^+). Observe que en la alternativa \mathcal{H}_1 se tiene una familia no paramétrica muy grande (es decir, $\mathcal{F} - \{F_0\}$).

Por otra lado, desde la perspectiva de la Bayesiana no paramétrica (donde se usa el proceso de Dirichlet), se supone que se tiene una muestra X_1, X_2, \dots, X_m de una distribución P ; es decir, se está interesado en probar $\mathcal{H}_0 : P = H$ vs. $\mathcal{H}_1 : P \neq H$. Con H una función de distribución continua conocida (caso simple). En la alternativa \mathcal{H}_1 , se tiene identificado al proceso estocástico, es decir, es un proceso de Dirichlet. Esto es una diferencia notable a las pruebas de bondad de ajuste clásicas, en donde la alternativa \mathcal{H}_1 no es especificada “completamente”. Tenga en cuenta que, aunque el proceso de Dirichlet es discreto con probabilidad uno, ésta discretización no es mas problema que el carácter discreto del proceso empírico. Al imponer la topología débil, el soporte del proceso de Dirichlet es bastante grande. En concreto, el soporte para el proceso de Dirichlet es el conjunto de todas las medidas de probabilidad cuyo soporte está contenido en el soporte de la medida base. Esto significa que si el soporte de la medida base es \mathfrak{X} , entonces el espacio de todas las medidas de probabilidad es el soporte del proceso de Dirichlet.

Por ejemplo, si tenemos una medida base de una distribución normal, entonces el proceso de Dirichlet puede elegir cualquier medida de probabilidad (ver [Al Labadi & Zarepour 2014](#)).

En esencia lo que se hace en el procedimiento que veremos a continuación, es rechazar la hipótesis nula, si la probabilidad posterior estimada del proceso de Dirichlet sea menor a la probabilidad de la creencia previa de que la distribución P es aproximadamente equivalente a H . Para ello, es necesario calcular la distancia posterior. Sea P_n^* una aproximación para el proceso de Dirichlet posterior, dado los datos. La aproximación posterior puede ser obtenida del proceso de Dirichlet con parámetro de concentración a^* y medida base H^* (ecuación 4.2). Sea $(\theta_k^*)_{1 \leq k \leq n}$ los saltos de P_n^* . Bajo la hipótesis nula \mathcal{H}_0 , $(\theta_k^*)_{1 \leq k \leq n}$ son variables aleatorias i.i.d con distribución común H . Por lo que las estadísticas de Kolmogorov y Cramér-von Mises vistas en la sección anterior pueden ser utilizadas para calcular la distancia entre P_n^* y H .

Algoritmo 1 (caso simple)

1. Dada una función de distribución H , encontrar un valor aproximado de ϵ como sigue (ver [Swartz 1999](#)):
 - (a) De los datos, encontrar un valor máximo x_0 de tal manera que la medición $x \pm x_0$ es considerada cerca de x .
 - (b) Se calcula ϵ como:

$$\epsilon = \max_{x \in \mathbb{R}} \{H(x + x_0) - H(x)\} \quad (4.4)$$

2. Encontrar el valor apropiado del parámetro de concentración tal que

$$P\{K(P_n, H) \leq \epsilon\} = q,$$

donde $0 < q < 1$ (para detalles ver [Al Labadi & Zarepour \(2014\)](#)). Aquí q representa la creencia previa que la distribución P es aproximadamente equivalente a H . Usualmente se toma $q = 0.5$.

3. Generar una muestra aleatoria de P_n^* .
4. Calcular $K(P_n^*, H)$.
5. Repetir los pasos 3 y 4 para obtener r muestras i.i.d de $K(P_n^*, H)$. Para n y r grande la distribución empírica de estos valores es una aproximación a

la distribución de $K(P^*, H)$. Por lo que la probabilidad $P\{K(P^*, H) \leq \epsilon\}$ es estimada como la proporción de $K(P_n^*, H)$ que son menores o iguales a ϵ . Basados en esta estimación, se decide si la distribución P es cercana a la distribución hipotética H . Es decir, si el valor de la probabilidad posterior estimada es más pequeña que la probabilidad a priori q , entonces existe suficiente evidencia para rechazar \mathcal{H}_0 .

De igual forma, para calcular la distancia Cramér-von Mises se utiliza el algoritmo 1, solo que ϵ se calcula como:

$$\epsilon = \int_{-\infty}^{\infty} (H(x + x_0) - H(x))^2 h(x) dx. \quad (4.5)$$

El enfoque para probar bondad de ajuste en el caso simple se puede generalizar al caso compuesto. El problema a considerar es $\mathcal{H}_0 : P = H_\theta$, donde H_θ es una familia de distribuciones con parámetro de dimensión finita θ , y distribución a priori $\pi(\theta)$. El algoritmo propuesto en [Al Labadi & Zarepour \(2014\)](#) es el siguiente:

Algoritmo 2 (caso compuesto)

1. Dada una función de distribución H_θ , encontrar un valor aproximado de ϵ como sigue (ver [Swartz 1999](#)):

Se calcula ϵ como:

$$\epsilon = \max_{x \in \mathbb{R}} \{H_{\hat{\theta}}(x + x_0) - H_{\hat{\theta}}(x)\},$$

donde $\hat{\theta} = E(\theta)$.

2. Encontrar el valor apropiado del parámetro de concentración tal que

$$P\{K(P_n, H_{\hat{\theta}}) \leq \epsilon\} = q,$$

donde $0 < q < 1$. De igual forma que en algoritmo 1, aquí q representa la creencia previa que la distribución P es aproximadamente equivalente a $H_{\hat{\theta}}$. Usualmente se toma $q = 0.5$.

3. Generar R realizaciones $(\theta_0^{(i)})_{1 \leq i \leq R}$ de la distribución final θ dada X_1, \dots, X_m .
4. Encontrar $\theta_{opt} = \arg \min_{\theta_0^{(i)}} K(P_n^*, H_{\theta_0^{(i)}})$ para $i = 1, \dots, R$
5. Calcular $K(P_n^*, H_{\theta_{opt}})$ como en algoritmo 1.

De igual forma, para calcular la distancia Cramér-von Mises se utiliza el algoritmo 2, solo que ϵ se calcula como:

$$\epsilon = \int_{-\infty}^{\infty} (H_{\hat{\theta}}(x + x_0) - H_{\hat{\theta}}(x))^2 h(x) dx.$$

4.2.4. Ejemplos y un estudio de potencia

A continuación reproducimos los ejemplos vistos en [Al Labadi & Zarepour \(2014\)](#). Posteriormente, se realiza un estudio de simulación para calcular la potencia de las distancias Kolmogorov y Cramér-von Mises vistas en la sección 4.2.2, también se incorporan dos estadísticas de prueba que se utilizan en la estadística clásica, que son la Kolmogorov y la Anderson-Darling.

Ejemplo 4.1. *Generamos una muestra de tamaño $m = 200$ de la distribución normal estándar. Se quiere ver si los datos siguen una distribución normal estándar. De los datos, se determina que la medida de precisión $x_0 = 0.2$. Para la distancia Kolmogorov, aplicando ecuación 4.4, se tiene que $\epsilon = 0.08$. Se determina el parámetro de concentración a , tal que cumple,*

$$P(K(P_n^{F.D}, H) \leq \epsilon) = 0.5$$

Para $\epsilon = 0.08$, se tiene que $a = 106$. Aplicando algoritmo 1, con $n = 2000$ y $r = 3000$, se tiene que la correspondiente probabilidad posterior $P(K(P_n^, H) \leq \epsilon)$ es igual a 0.89. Por lo que no se puede rechazar \mathcal{H}_0 . De igual forma, para la distancia Cramér-von Mises, aplicando ecuación 4.5, se tiene que $\epsilon = 0.004$. Se determina el parámetro de concentración a , tal que cumple que,*

$$P(CvM(P_n^{F.D}, H) \leq \epsilon) = 0.5$$

Para $\epsilon = 0.004$, se tiene que $a = 30$. Aplicando algoritmo 1, se tiene que la correspondiente probabilidad posterior $P(CvM(P_n^, H) \leq \epsilon)$ es igual a 0.9883. Por lo que no se puede rechazar \mathcal{H}_0 . En las figuras 4.1 y 4.2 se gráfica la distribución normal estándar acumulada y dos recorridos del proceso de Dirichlet a priori y a posteriori, respectivamente. En la primera gráfica, el ajuste fue mediante la distancia de Kolmogorov, y en la segunda fue por la distancia de Cramér-von Mises.*

Ejemplo 4.2. *Generamos una muestra de tamaño $m = 200$ de la distribución Cauchy. Se quiere ver si los datos siguen una distribución normal estándar. Se elige la medida de precisión del ejemplo anterior, por lo que $x_0 = 0.2$, y para la distancia Kolmogorov $\epsilon = 0.08$ y $a = 106$. Aplicando algoritmo 1, con $n = 2000$ y $r = 3000$, se tiene que la correspondiente probabilidad posterior $P(K(P_n^*, H) \leq \epsilon)$ es igual a 0.0233. Por lo que existe evidencia para rechazar \mathcal{H}_0 . De igual forma, para la distancia Cramér-von Mises, del ejemplo anterior, se tiene que $\epsilon = 0.004$ y $a = 30$. Aplicando algoritmo 1, se tiene que la correspondiente probabilidad posterior $P(CvM(P_n^*, H) \leq \epsilon)$ es igual a 0.243. Por lo que existe evidencia para rechazar*

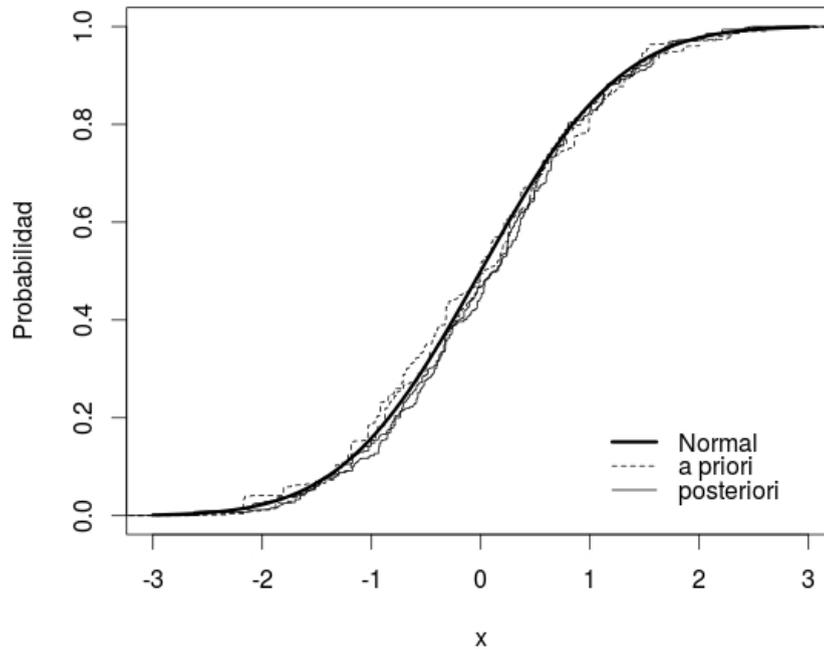


FIGURA 4.1: Ejemplo 4.1. Distribución normal estándar. Aproximación al proceso de Dirichlet a priori y a posteriori ajustado mediante la distancia de Kolmogorov. $n = 2000$, $m = 200$ y $a = 106$.

\mathcal{H}_0 . En las figuras 4.3 y 4.4 se gráfica la distribución normal estándar acumulada y dos recorridos del proceso de Dirichlet a priori y a posteriori, respectivamente. En la primera gráfica, el ajuste fue mediante la distancia de Kolmogorov, mientras que en la segunda fue por la distancia de Cramér-von Mises.

Estudio de potencia

Se realizó un estudio de simulación para calcular el número de rechazos de varias estadísticas de prueba en el caso de normalidad. Se eligió un tamaño de muestra m igual a 200. En la tabla 4.4 se reporta la frecuencia de rechazos con 1000 iteraciones. Para cada iteración se eligió $n = 2000$ y $r = 3000$ para calcular $P(K(P_n^*, H) \leq \epsilon)$ y $P(CvM(P_n^*, H) \leq \epsilon)$, con los parámetros de concentración y ϵ vistos en ejemplo 4.1. Por otra parte, también se incorporaron dos estadísticas de prueba utilizadas en la estadística clásica, es decir, la estadística Kolmogorov y la estadística Anderson-Darling, que se denotan por K y A^2 , respectivamente. Para éstas estadísticas se eligió un nivel de significancia $\alpha = 0.05$. Las alternativas elegidas son la distribución normal, t-student, normal estándar truncada, y la normal estándar sesgada. Denotadas por $N(\cdot, \cdot)$, $t(\cdot)$, $TrN(\cdot, \cdot)$ y $SN(\cdot)$, respectivamente.

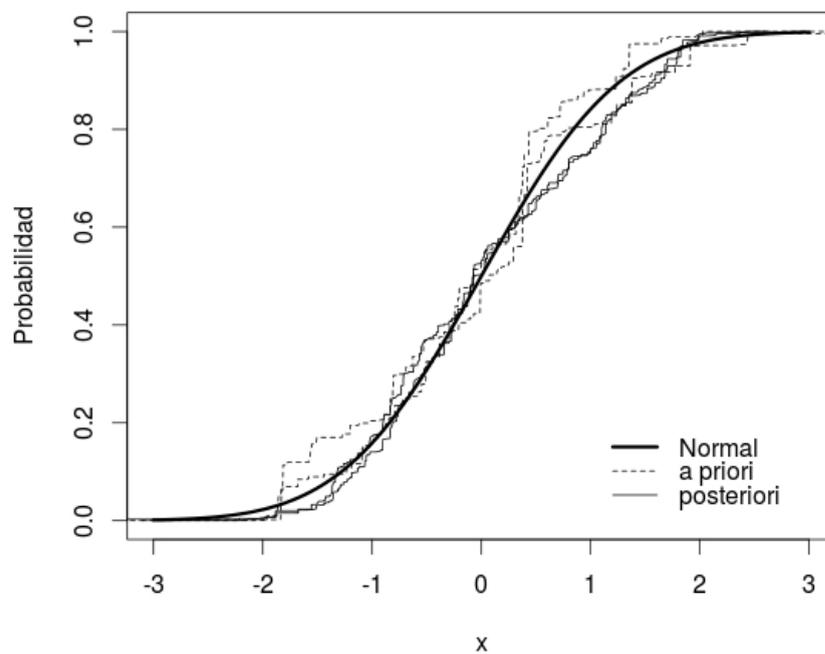


FIGURA 4.2: Ejemplo 4.1. Distribución normal estándar. Aproximación al proceso de Dirichlet a priori y a posteriori ajustado mediante la distancia de Cramér-von Mises, $n = 2000$, $m = 200$ y $a = 30$.

De la tabla 4.4 se puede ver que las estadísticas K y $K(P_n^*, H)$ tienen un potencia muy parecida, $K(P_n^*, H)$ supera ligeramente a K . La distancia $CvM(P_n^*, H)$ resultó tener potencia muy baja. Mientras que la estadística Anderson-Darling A^2 resultó ser la que tiene mayor número de rechazos.

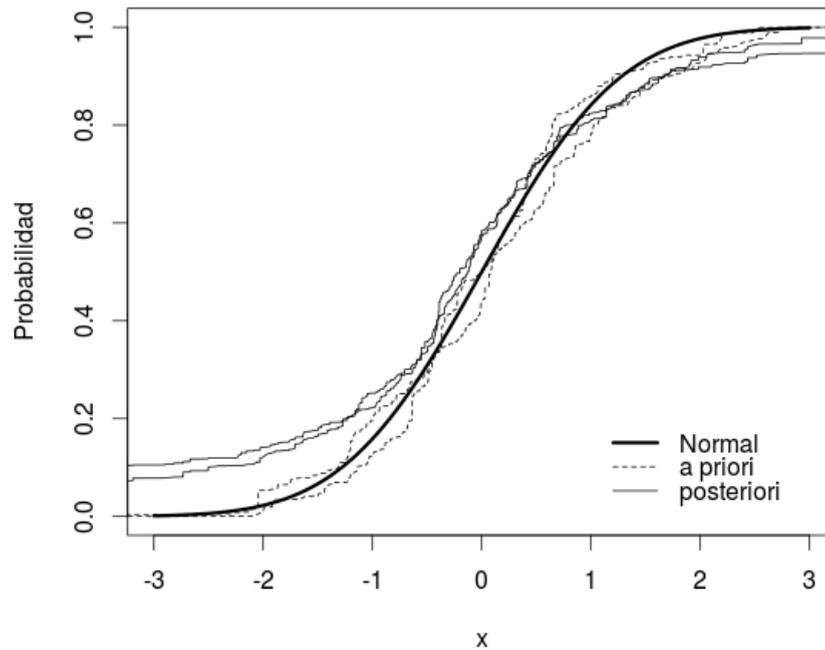


FIGURA 4.3: Ejemplo 4.2. Distribución normal estándar. Aproximación al proceso de Dirichlet a priori y a posteriori ajustado mediante la distancia de Kolmogorov. $n = 2000$, $m = 200$ y $a = 106$.

TABLA 4.4: Potencia de bondad de ajuste de la distribución normal estándar (caso simple) para diferentes alternativas. Tamaño de muestra $m = 200$. $n = 2000$, $r = 3000$ y $\alpha = 0.05$.

Alternativas	K	A^2	$K(P_n^*, H)$	$CvM(P_n^*, H)$
$N(0, 1)$	0.043	0.043	0.033	0.004
$N(0.2, 1)$	0.695	0.788	0.652	0.384
$t(1)$	1	1	0.998	0.893
$t(2)$	0.565	1	0.566	0.109
$t(6)$	0.089	0.384	0.050	0.006
$TrN(-1.5, 1.5)$	0.25	0.934	0.269	0.01
$TrN(-2, 2)$	0.055	0.082	0.044	0.002
$SN(0.1)$	0.151	0.175	0.143	0.033
$SN(0.3)$	0.806	0.878	0.852	0.571

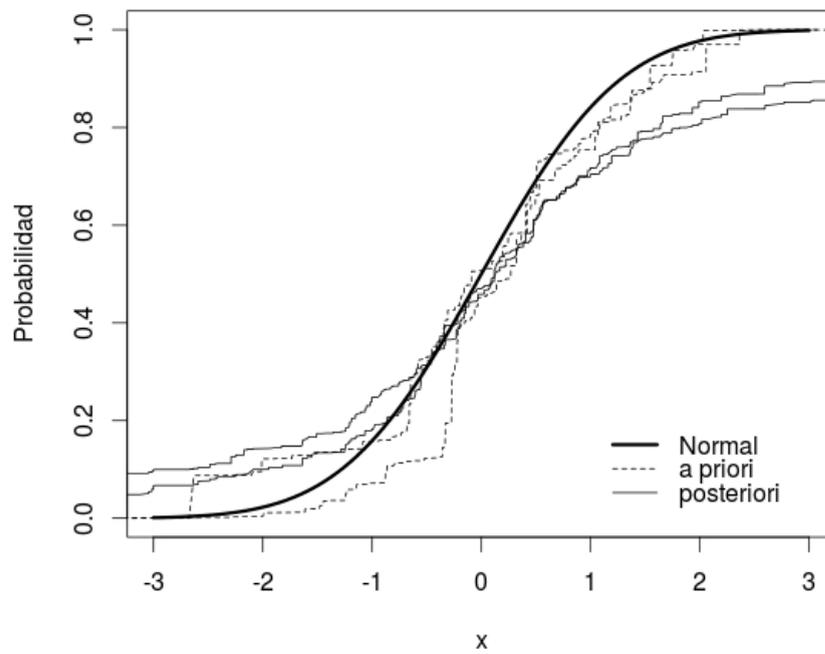


FIGURA 4.4: Ejemplo 4.2. Distribución normal estándar. Aproximación al proceso de Dirichlet a priori y a posteriori ajustado mediante la distancia de Cramér-von Mises. $n = 2000$, $m = 200$ y $a = 30$.

Capítulo 5

Conclusiones y perspectivas

El interés en el estudio de bondad de ajuste ha resurgido ya que han proliferado propuestas recientes tanto en las estadísticas, como en los procedimientos para calibrarlas; esto es importante, ya que el nivel de la prueba debe estar bien calculado. En particular, se analizaron dos enfoques, uno que es una aproximación (remuestro paramétrico) y otro que da una solución exacta al problema (remuestro condicional).

Para el caso discreto se han realizado estudios comparativos de diferentes pruebas. En el caso continuo, se estudia la distribución exponencial truncada y la Gaussiana-inversa, mediante los dos procedimientos mencionados, es decir, el remuestro paramétrico y el remuestro condicional. También, se aborda el problema de bondad de ajuste para la distribución exponencial (sin censura, con censura) para dos pruebas muy recientes, comparándolas con dos pruebas clásicas. Por último, se exploran dos propuestas recientes basadas en la estadística Bayesiana, uno paramétrico y otro no paramétrico.

En el caso discreto, mediante la teoría de muestreo condicional se hace un estudio exhaustivo de la familia Poisson, binomial negativa y binomial, respectivamente, para comparar la potencia entre varias estadísticas de prueba y se proponen dos nuevas estadísticas, una de las cuales resultó ser muy potente e igual de competitiva a las tradicionales estadísticas de prueba de la familia Cramér-von Mises.

Para el estudio las distribuciones exponencial truncada y la Gaussiana-inversa, en las simulaciones realizadas se ve que existe una muy fuerte correlación, entre los dos procedimientos usados para calcular por simulación el p-valor, estos métodos

son el remuestreo paramétrico y el remuestreo condicional. Para el caso de probar de ajuste para la distribución Gaussiana-inversa, la estadística basada en la información de Kullback-Leibler resultó tener mejor potencia que la estadística de prueba de Anderson-Darling. Por otro lado, para la distribución exponencial truncada no se tiene conocimiento que anteriormente se hubiera realizado un estudio de potencia, teniendo la estadística de prueba de Anderson-Darling una potencia aceptable.

También para las pruebas de bondad de ajuste se exploraron dos enfoques Bayesianos. En el primero, los autores [He et al. \(2013\)](#), utilizan la estadística Bayesiana paramétrica; en este trabajo, se concluye que para probar normalidad se recomienda usar la estadística de prueba Anderson-Darling en lugar de la estadística que ellos proponen, ya que resultó ser más potente. Otra ventaja de usar la estadística de Anderson-Darling es que no se necesita tanto recurso computacional, al contrario de la prueba propuesta en [He et al. \(2013\)](#), en la cual, se requiere un mayor esfuerzo computacional. De igual forma, se tiene que el enfoque de la estadística Bayesiana no paramétrica, donde se usa el proceso de Dirichlet, en el pequeño estudio de potencia que se realizó, la estadística de prueba Anderson-Darling resultó potente.

Para el caso la distribución exponencial se analizaron dos casos posibles, es decir, en presencia de datos sin censura y con datos censurados tipo II. Se realizó un estudio de potencia para dos estadísticas de prueba clásicas y dos estadísticas de prueba muy recientes, una basada en la entropía residual acumulada y la otra en la información de Kullback-Leibler. Se encontraron inconsistencias en la estadística de prueba basada en la información de Kullback-Leibler, ya que en algunas familias cuando se disminuye la censura, la potencia disminuye, y debería ocurrir lo contrario. Quedó pendiente analizar con más detalle esta estadística y ver si se puede encontrar un resultado teórico, que respalde porqué esta estadística tiene este comportamiento inconsistente.

Bibliografía

- Al Labadi, L. & Zarepour, M. (2014), ‘Goodness-of-fit tests based on the distance between the dirichlet process and its base measure’, *Journal of Nonparametric Statistics* **26**(2), 341–357.
- Alizadeh Noughabi, H. & Arghami, N. R. (2013), ‘General treatment of goodness-of-fit tests based on Kullback-Leibler information’, *J. Stat. Comput. Simul.* **83**(8), 1556–1569.
- Best, D. J. & Rayner, J. C. W. (1997), ‘Goodness of fit for the binomial distribution’, *Austral. J. Statist.* **39**(3), 355–364.
- Best, D. J. & Rayner, J. C. W. (1999), ‘Goodness of fit for the poisson distribution’, *Statistics and probability letters* **44**(3), 259–265.
- Best, D. J. & Rayner, J. C. W. (2003), ‘Tests of fit for the geometric distribution’, *Comm. Statist. Simulation Comput.* **32**(4), 1065–1078.
- Bliss, C. & Fisher, R. (1953), ‘Fitting the negative binomial distribution to biological data and note on the efficient fitting of the negative binomial’, *Biometrics* **9**, 176–200.
- Bortkiewicz, L. (1898), *Das Gesetz der kleinen Zahlen*, Teubner, Leipzig .
- Consult, P. & Gupta, H. (1980), ‘The generalized negative binomial distribution and its characterization by zero regression’, *SIAM J. APPL. MATH.* **39**(2).
- Cressie, N. & Read, T. R. C. (1984), ‘Multinomial goodness-of-fit tests’, *J. Roy. Statist. Soc. Ser. B* **46**(3), 440–464.
- Engen, S. & Lillegard, M. (1997), ‘Stochastic simulations conditioned on sufficient statistics’, *Biometrika* **84**, 235–240.
- Ferguson, T. S. (1973), ‘A bayesian analysis of some nonparametric problems’, *The annals of statistics* pp. 209–230.

- Fisher, R. (1950), 'The significance of deviations from expectations in a poisson series', *Biometrics* **6**, 17–24.
- Forbes, C., Evans, M., Hastings, N. & Peacock, B. (2011), *Statistical distributions*, John Wiley & Sons.
- Frey, J. (2012), 'An exact kolmogorov-smirnov test for the poisson distribution with unknown mean', *Comm. Statist. Simulation Comput.* **82**(7), 1023–1033.
- Gelman, A., li Meng, X. & Stern, H. (1996), 'Posterior predictive assessment of model fitness via realized discrepancies', *Statistica Sinica* pp. 733–807.
- González-Barrios, J. M., O'Reilly, F. & Rueda, R. (2010), 'Durbin's random substitution and conditional monte carlo', *Metrika* **72**(3), 369–383.
- González-Barrios, J., O'Reilly, F. & Rueda, R. (2006), 'Goodness of fit for discrete random variables using the conditional density', *Metrika* **64**, 77–94.
- Gurtler, N. & Henze, N. (2000), 'Recent and classical goodness-of-fit tests for the poisson distribution', *Journal of Statistical Planning and Inference* **90**(2), 207–225.
- Guttman, I. (1967), 'The use of the concept of a future observation in goodness-of-fit problems', *J. Roy. Statist. Soc. Ser. B* **29**, 83–100.
- He, D., Xu, X. & Liu, X. (2013), 'The use of posterior predictive p-values in testing goodness-of-fit', *Communications in Statistics - Theory and Methods* **42**(23), 4287–4297.
- Ishwaran, H. & Zarepour, M. (2002), 'Exact and approximate sum representations for the dirichlet process', *The Canadian Journal of Statistics/La Revue Canadienne de Statistique* pp. 269–283.
- Kyriakoussis, A., Gang, L. & A., P. (1998), 'On characterization and goodness-of-fit test of some discrete distribution families', *Journal of Statistical Planning and Inference* **74**, 215–228.
- Lindqvist, B. H. & Taraldsen, G. (2005), 'Monte Carlo conditioning on a sufficient statistic', *Biometrika* **92**(2), 451–464.
- Lo, A. Y. (1987), 'A large sample study of the bayesian bootstrap', *The Annals of Statistics* pp. 360–375.

- Lockhart, R. A., O'Reilly, F. J. & Stephens, M. A. (2007), 'Use of the gibbs sampler to obtain conditional tests, with applications', *Biometrika* **94**(4), 992–998.
- Lockhart, R. A., O'Reilly, F. & Stephens, M. (2009), 'Exact conditional tests and approximate bootstrap tests for the von mises distribution', *Journal of Statistical Theory and Practice* **3**(3), 543–554.
- Lockhart, R. A. et al. (2012), 'Conditional limit laws for goodness-of-fit tests', *Bernoulli* **18**(3), 857–882.
- Meng, X.-L. (1994), 'Posterior predictive p-values', *The Annals of Statistics* **22**(3), pp. 1142–1160.
- Muliere, P. & Tardella, L. (1998), 'Approximating distributions of random functionals of ferguson-dirichlet priors', *Canadian Journal of Statistics* **26**(2), 283–297.
- O'Reilly, F. & Gracia-Medrano, L. (2006), 'On the conditional distribution of goodness of fit tests', *Communications in Statistics. (Theory and Methods)* **35**(3), 541–549.
- O'Reilly, F. J. (1971), On goodness-of-fit tests based on Rao-Blackwell distribution function estimators., PhD thesis, Citeseer.
- O'Reilly, F. J. & Stephens, M. A. (1982), 'Characterizations and goodness of fit tests', *J. Roy. Statist. Soc. Ser. B* **44**(3), 353–360.
- O'Reilly, F. J. & Villegas, C. (1987), 'Conditional and predictive probability integral transforms', *Canadian Journal of Statistics* **15**(2), 151–158.
- Park, S. & Lim, J. (2015), 'On censored cumulative residual kullback-leibler information and goodness-of-fit test with type ii censored data', *Statistical Papers* **56**(1), 247–256.
- Potas, N., Saridoy, E. & Kara, M. (2014), 'A modified goodness-of-fit test based on likelihood ratio for the skew-normal distribution', *Applied Mathematical Sciences* **8**(78), 3869–3887.
- Raj S. Chhikara, J. L. F. (1974), 'Estimation of the inverse gaussian distribution function', *Journal of the American Statistical Association* **69**(345), 250–254.
- Rayner, J. & Best, D. (1989), 'Smooth tests of goodness of fit', *Oxford University Press, New York* .

- Rubin, D. B. (1984), 'Bayesianly justifiable and relevant frequency calculations for the applied statistician', *Ann. Statist.* **12**(4), 1151–1172.
- Rueda, R. & O'Reilly, F. (1999), 'Tests of fit for discrete distributions based on the probability generating function', *Comm. Statist. Simulation Comput.* **28**(1), 259–274.
- Rueda, R., Pérez-Abreu, V. & O'Reilly, F. (1991), 'Goodness of fit for the poisson distribution based on the probability generating function', *Comm. Statist. Theo. Meth.* **20**, 3093–3110.
- Snedecor, G. & Cochran, W. (1989), 'Statistical methods (8th edn.)', *Ames: Iowa State University Press.* .
- Spinelli, J. & Stephens, M. (1997), 'Cramér-von mises tests of fit for the poisson distribution', *Can. Jour. Statisti.* **25**(2), 257–268.
- Sukhatme, P. (1937), 'Tests of significance for sample of the χ^2 population with two degrees of freedom', *Ann Eugenics* **8**.
- Swartz, T. (1999), 'Nonparametric goodness-of-fit', *Communications in Statistics-Theory and Methods* **28**(12), 2821–2841.
- Zardasht, V., Parsi, S. & Mousazadeh, M. (2014), 'On empirical cumulative residual entropy and a goodness-of-fit test for exponentiality', *Statistical Papers* pp. 1–12.
- Zhang, J. (2002), 'Powerful goodness-of-fit tests based on the likelihood ratio', *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**(2), 281–294.