



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS MATEMÁTICAS

LA DISTRIBUCIÓN FIDUCIAL EN MODELOS
QUE NO SON DE GRUPO.
CÓMO SIMULAR DE ELLA.

T E S I S

QUE PARA OPTAR POR EL GRADO ACADÉMICO DE
DOCTOR EN CIENCIAS

P R E S E N T A

EDILBERTO NÁJERA RANGEL

DIRECTOR DE TESIS: DR. FEDERICO O'REILLY TOGNO (IIMAS – UNAM).

MIEMBROS DEL COMITÉ TUTOR: DR. RAÚL RUEDA (IIMAS – UNAM),
DR. HUGO ARIZMENDI PEIMBERT (IMATE-UNAM).

MÉXICO, D. F. MARZO 2014.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

La distribución fiducial en modelos que no son de grupo.
Cómo simular de ella

Índice general

Introducción	6
1. Antecedentes de la distribución fiducial	7
1.1. La distribución fiducial de Fisher	7
1.2. Distribuciones fiduciales y el Teorema de Bayes	10
1.3. El método fiducial e invarianza	19
1.3.1. El modelo de las transformaciones	20
1.3.2. Una cantidad pivotal	23
1.3.3. La distribución fiducial	26
1.4. Ejemplos relevantes sobre la definición de probabilidad fiducial	29
1.4.1. El caso unidimensional	30
1.4.2. El caso multidimensional	30
1.4.3. Distribuciones fiduciales multivariadas y el Teorema de Bayes	34
2. La distribución fiducial	36
2.1. Contrastabilidad monótona y extrema	36
2.2. Ejemplos	37
2.3. Distribución fiducial	38
2.4. Familia exponencial	39
2.5. T discreta y θ continua	49
2.5.1. Análisis fiducial por medio de distribuciones de series de potencias	54
3. Generadores de la distribución fiducial	56
3.1. El generador inducido por la función de distribución	57
3.2. El generador de Engen y Lillegård	58
3.3. Distribución <i>Gamma</i> (α, β)	64

3.4. Distribución Exponencial Truncada 72

4. Conclusiones 80

Bibliografía 81

Introducción

Como un método alternativo al procedimiento Bayesiano, en el que se obtiene una distribución del parámetro θ a partir de una distribución inicial y la muestra (x_1, x_2, \dots, x_n) , en 1930 Fisher [9] introdujo el procedimiento fiducial para hacer inferencias sobre θ . El argumento fiducial no utiliza una distribución inicial, sino que se basa tan sólo en las observaciones disponibles, o sea (x_1, x_2, \dots, x_n) . En este planteamiento θ juega el papel lógico de una variable aleatoria, y el valor observado de la estadística suficiente minimal de θ “hace el papel” de un parámetro que define la distribución fiducial de θ .

Desde que fue expuesto por Fisher, el método fiducial ha sido muy controvertido: hubo quienes lo calificaron como inconsistente, pues argumentaron que en algunos casos hay diferentes estadísticos suficientes del parámetro, los cuales satisfacen las condiciones establecidas por Fisher, pero que sin embargo inducen distribuciones diferentes del parámetro; otros señalaron la falta de acuerdo entre los investigadores en cuanto a la derivación, manipulación e interpretación de la probabilidad fiducial.

A lo largo de los años, Fisher mismo fue cambiando su punto de vista del argumento fiducial, como puede verse por ejemplo en [9], [10], [11] y [12]. La idea inicial de Fisher fue obtener la distribución fiducial directamente de la función de distribución, pero este procedimiento enfrenta serias dificultades cuando se quiere aplicar al caso multiparamétrico. Como consecuencia, en esta situación propuso una variedad de métodos y condiciones para obtener la distribución fiducial (ver por ejemplo [3]).

No obstante lo controversial que ha resultado ser, el argumento fiducial ha sido históricamente influyente en el desarrollo de la estadística (ver [7]).

En 1957 Lindley [19] demostró que una condición necesaria y suficiente para que la distribución fiducial del parámetro simple θ (escalar) corresponda a una distribución final Bayesiana, es que existan transformaciones de x a u y de θ a τ , tal que τ sea un parámetro de localización para u , en donde x representa al observable (escalar) cuya distribución depende de θ , y u (escalar) es una transformación de x cuya distribución depende de τ . En este caso Lindley llama consistente al argumento fiducial. Posteriormente, en 1961, Fraser [13] demostró que si la distribución de la estadística suficiente minimal de θ tiene estructura de grupo, entonces el argumento fiducial es consistente, de acuerdo con Lindley, cuyo resultado es un caso particular del de Fraser.

En parte debido al trabajo citado de Lindley y en parte por las controversias en torno suyo, aproximadamente desde fines de la década de los 60's y hasta principios de los 90's del siglo pasado, la inferencia fiducial prácticamente fue olvidada, salvo por esporádicas publicaciones en las que se intentaba resolver o explicar algunas de las controversias, o bien se reportaban algunas ventajas del argumento fiducial en ciertas aplicaciones. No obstante, como lo consignan Hannig [17] y Taraldsen y Lindqvist [27], en los últimos 20 años ha surgido un renovado interés por la inferencia fiducial, aunque a veces bajo la etiqueta de “inferencia generalizada”, en la que, como lo señala Hannig [17], se usa el argumento fiducial.

En 2007 O'Reilly y Rueda [23] obtuvieron la distribución fiducial del parámetro θ de la distribución exponencial truncada. Este es un caso de un parámetro unidimensional que no es de localización ni de escala, y por lo tanto su distribución fiducial no es igual a una distribución final Bayesiana. Sin embargo, mostraron numéricamente que ambas prácticamente coinciden cuando la distribución inicial de θ es la de Jeffreys.

En 1997, Engen y Lillegård [8] propusieron un método, revisado posteriormente por Lindqvist y Taraldsen [20] en 2005, para realizar simulaciones de Monte Carlo condicionando sobre una estadística suficiente. La idea básica consiste en ajustar los valores del parámetro en la correspondiente simulación no condicionada, de modo tal que se obtenga el valor observado de la estadística suficiente. De este modo, aparentemente sin ser el objetivo, terminan proponiendo un simulador de una posible distribución fiducial del parámetro. En este trabajo aplicamos esta idea para el caso de las distribuciones exponencial truncada y $gamma(\alpha, \beta)$. En cada caso, numéricamente vemos que las distribuciones fiduciales obtenidas son muy parecidas a las correspondientes distribuciones finales Bayesianas, cuando la distribución inicial es la no informativa de Jeffreys.

En 2009 Hannig [17] extendió el argumento fiducial de Fisher y obtuvo un método fiducial generalizado, basado en la solución de una “ecuación estructural”, el cual aplicó a varios ejemplos de complejidad diversa. Observó que algunas propiedades de los procedimientos estadísticos obtenidos con el método fiducial generalizado tuvieron muy buen desempeño, incluso cuando fueron comparados con los correspondientes obtenidos con la Inferencia Bayesiana.

En 2013 Taraldsen y Lindqvist [27] desarrollaron el argumento fiducial en un escenario basado en la solución de una “ecuación fiducial”, en este sentido similar al enfoque utilizado por Hannig. Afirman que la distribución fiducial así obtenida determina procedimientos de inferencia frecuentista equivariantes optimales, no sólo en modelos de grupo, sino también en modelos de cuasigrupo. Asimismo sugieren que el argumento fiducial se puede considerar como una alternativa al argumento Bayesiano en la construcción de buenos procedimientos de inferencia frecuentista.

Sin embargo, las dos propuestas anteriores son muy diferentes de la que aquí se hace en el sentido de que son de carácter muy general, mientras que los procedimientos que aquí presentamos conducen a la elección de una única distribución fiducial.

Este trabajo está desarrollado como sigue: En el capítulo 1 se ve el argumento fiducial expuesto por Fisher en 1930. También se ven algunos de los resultados que creemos son importantes, tales como los ya referidos de Lindley y Fraser, para entender tanto las ventajas como las desventajas que tradicionalmente se le han atribuido al método fiducial. En el capítulo 2, esencialmente se hace una revisión de un enfoque que utiliza la significación en pruebas de hipótesis [22]. En él se analizan las familias exponencial y exponencial natural (en cuyo caso se da una demostración de que siempre es posible la construcción tradicional de la distribución fiducial). En el caso de las distribuciones asociadas a la Bernoulli (binomial, binomial negativa y Poisson), presenta una distribución fiducial única invariante, la invarianza en cuanto al intercambio entre “éxito” o “fracaso”. Además, la densidad fiducial resulta ser la misma en los casos binomial y binomial negativa. Así, la distribución fiducial propuesta satisface el principio fuerte de verosimilitud. Por último, para dichas distribuciones en esta tesis se propone hacer el análisis fiducial correspondiente utilizando el enfoque de distribuciones de series de potencias. En el capítulo 3, para obtener la distribución fiducial se propone el generador derivado del trabajo ya citado de Engen y Lillegård. Cuando el parámetro de la distribución es real, se considera también el generador fiducial obvio obtenido de la expresión de la función de distribución fiducial. Adicionalmente, derivado de las ideas desarrolladas en [22], se propone un método alternativo para generar distribuciones fiduciales, el cual se compara con el de Engen y Lillegård en dos ejemplos, en los que se ve que las distribuciones fiduciales simuladas correspondientes de hecho coinciden. Con el generador de Engen y Lillegård se obtiene la distribución fiducial conjunta de (α, β) de la distribución gamma, así como la distribución fiducial del parámetro θ de la distribución exponencial truncada. En estos dos casos de familias paramétricas, en las cuales el parámetro, (α, β) en el caso gamma o θ en el caso de la exponencial truncada, no se puede tomar como elemento (o función de un elemento) de un grupo, en el sentido clásico, se muestra numéricamente que ambas distribuciones, la fiducial y la distribución final Bayesiana, resultan muy parecidas cuando la distribución inicial es la de Jeffreys.

Capítulo 1

Antecedentes de la distribución fiducial

En este capítulo se presenta el argumento fiducial expuesto por Fisher en 1930 [9]. También se presentan los resultados que creemos más importantes para entender tanto sus ventajas como las desventajas que tradicionalmente se le han atribuido.

1.1. La distribución fiducial de Fisher

Como una alternativa a la distribución final Bayesiana, Fisher [9] introdujo las distribuciones de probabilidad fiduciales para los parámetros de la distribución de probabilidad de la población de interés.

En 1935 Fisher, [10], da la siguiente presentación del argumento fiducial. Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución $N(\mu, \sigma^2)$, con μ y σ^2 desconocidas, donde el parámetro de interés es μ . Si $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ y $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, entonces la variable aleatoria

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S}$$

tiene una distribución t de Student con $n - 1$ grados de libertad. Si x_1, x_2, \dots, x_n son los respectivos valores observados de X_1, X_2, \dots, X_n , y por ende \bar{X} y S^2 toman los valores \bar{x} y s^2 respectivamente, entonces Fisher dice que la cantidad

$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s}$$

está distribuida de acuerdo a la distribución t de Student con $n-1$ grados de libertad. Observa

que “se debe notar que t es una función continua de μ , junto con los valores observables de \bar{x} , s y n , solamente. Por lo tanto, la desigualdad

$$t > t_1$$

es equivalente a la desigualdad

$$\mu < \bar{x} - \frac{st_1}{\sqrt{n}}$$

de modo que esta última desigualdad debe ser satisfecha con la misma probabilidad que la primera. Esta probabilidad es conocida para todos los valores de t_1 , y decrece continuamente conforme t_1 crece. Como el lado derecho de la desigualdad toma todos los valores reales variando t_1 , podemos establecer la probabilidad de que μ sea menor que cualquier valor, o la probabilidad de que μ esté entre cualesquiera dos valores, o, brevemente, su distribución de probabilidad a la luz de la muestra observada.”

En el mismo artículo, considera el siguiente problema: Dados los valores x_1, x_2, \dots, x_n de una muestra aleatoria X_1, X_2, \dots, X_n de una distribución $N(\mu, \sigma^2)$, con μ y σ^2 desconocidas, encontrar la distribución de una observación x , aún no hecha, de esta distribución. Siguiendo un razonamiento similar al anterior, concluye que la cantidad

$$t = \sqrt{\frac{n}{n+1}} \frac{x - \bar{x}}{s},$$

donde \bar{x} y s están definidas como antes, está distribuida de acuerdo a la distribución t de Student con $n - 1$ grados de libertad. De aquí deduce que para $t_1 \in \mathbb{R}$ conocido, la probabilidad de que

$$t > t_1,$$

es igual a la probabilidad de que

$$x > \bar{x} + st_1 \sqrt{\frac{n+1}{n}}.$$

Por lo tanto la distribución t de Student, con el factor $\sqrt{1 + \frac{1}{n}}$, proporciona la distribución fiducial de

$$\frac{x - \bar{x}}{s},$$

en donde la única cantidad desconocida es la observación futura x .

Enseguida considera el problema análogo más general: dados los valores de una muestra aleatoria de tamaño n de una distribución $N(\mu, \sigma^2)$, con μ y σ^2 desconocidas, de la cual se tienen \bar{x} y s definidas como antes, encontrar la distribución fiducial de \bar{x}' y s' , obtenidas de las observaciones de una muestra aleatoria futura de tamaño n' . Obtiene que la función de densidad fiducial conjunta de \bar{x}' y s' es

$$f(\bar{x}', s') = \left(\frac{\frac{1}{2}(n+n'-3)!}{\frac{1}{2}(n-3)! \frac{1}{2}(n'-3)! (\sqrt{\pi})} \right) \left(\frac{2(n-1)^{\frac{1}{2}(n-1)} (n'-1)^{\frac{1}{2}(n'-1)} s^{n-1} s'^{n'-2}}{\sqrt{\frac{1}{n} + \frac{1}{n'}} \left((n-1)s^2 + (n'-1)s'^2 + \frac{(\bar{x}-\bar{x}')^2}{\sqrt{\frac{1}{n} + \frac{1}{n'}}} \right)^{\frac{1}{2}(n+n'-1)}} \right).$$

De aquí obtiene la función de densidad fiducial conjunta de μ y σ :

$$\begin{aligned} f(\mu, \sigma) &= \lim_{n' \rightarrow \infty} f(\bar{x}', s') \\ &= \frac{\sqrt{n}}{\sigma \sqrt{2\pi}} e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2} \frac{1}{\frac{1}{2}(n-3)!} \left(\frac{(n-1)s^2}{2\sigma^2} \right)^{\frac{1}{2}(n-1)} e^{-\frac{(n-1)s^2}{2\sigma^2}} \frac{2}{\sigma}. \end{aligned}$$

Integrando con respecto a σ de 0 a ∞ , encuentra la función de densidad fiducial marginal de μ , la cual, como era de esperarse, es

$$f(\mu) = \frac{\frac{1}{2}(n-2)!}{\frac{1}{2}(n-3)! \left(\sqrt{\pi(n-1)} \right)} \frac{1}{\left(1 + \frac{(\bar{x}-\mu)^2}{(n-1)s^2} \right)^{\frac{n}{2}}} \frac{\sqrt{n}}{s},$$

de acuerdo al hecho de que

$$\frac{\sqrt{n}(\bar{X} - \mu)}{S}$$

tiene distribución t de Student con $n-1$ grados de libertad. Similarmente obtiene la función de densidad fiducial marginal de σ , la cual es

$$f(\sigma) = \frac{1}{\frac{1}{2}(n-3)!} \left(\frac{(n-1)s^2}{2\sigma^2} \right)^{\frac{1}{2}(n-1)} e^{-\frac{(n-1)s^2}{2\sigma^2}} \frac{2}{\sigma},$$

de acuerdo al hecho de que

$$\frac{(n-1)S^2}{\sigma^2}$$

tiene distribución ji-cuadrada con $n-1$ grados de libertad.

Fisher concluye que “*parece que, para el caso especial de la media y la varianza de la distribución normal, no hay dificultad para extender sin ambigüedad la noción de probabilidad fiducial a la distribución conjunta de dos parámetros.*”

En general, parece que si las estadísticas T_1, T_2, T_3, \dots conjuntamente contienen toda la información disponible respecto a los parámetros $\theta_1, \theta_2, \theta_3, \dots$, y si se pueden encontrar funciones t_1, t_2, t_3, \dots de las T 's y los θ 's, para las cuales su distribución conjunta no depende de $\theta_1, \theta_2, \dots$, entonces la distribución fiducial conjunta de $\theta_1, \theta_2, \dots$ puede ser encontrada por sustitución. Para una estadística individual que tiene una distribución que depende sólo de un parámetro, del cual es un estimador suficiente, tal función siempre está dada por la probabilidad integral. De aquí la generalidad del método en los casos univariados”.

Para aclarar aun más el argumento fiducial, en 1945 [11] presenta el siguiente ejemplo:

“*Sea μ la mediana de una distribución, de la cual sólo se sabe que es continua. Sean x_1 y x_2 dos observaciones (independientes) de la distribución; entonces, para cualquier valor de μ será cierto que*

- *en un caso de cada cuatro, x_1 y x_2 serán mayores que μ ,*
- *en dos casos de cada cuatro, una observación será mayor que μ y la otra observación será menor que μ ,*
- *en un caso de cada cuatro, x_1 y x_2 serán menores que μ .*

Si Y es el número de observaciones menores que la mediana, entonces Y es una cantidad pivotal que involucra al parámetro y a las observaciones, y tiene una distribución de muestreo independiente del parámetro, es decir, Y toma los valores 0, 1 y 2, con probabilidades $1/4$, $1/2$, $1/4$, respectivamente.

Reconociendo esta propiedad, podemos decir que para dos observaciones dadas, ahora consideradas como parámetros fijos, la probabilidad de que μ sea menor que x_1 y x_2 es $1/4$, la probabilidad de que μ esté entre x_1 y x_2 es $1/2$, y la probabilidad de que μ sea mayor que x_1 y x_2 es $1/4$. De este modo, el argumento lleva a una distribución de probabilidad para μ , considerada ahora como una variable aleatoria”.

1.2. Distribuciones fiduciales y el Teorema de Bayes

En 1958 D. V. Lindley [19] estableció los resultados siguientes:

Sea X una variable aleatoria cuya distribución depende de un parámetro escalar θ . Entonces

1. Una condición necesaria y suficiente para que la distribución fiducial de θ corresponda a una distribución final bayesiana, es que existan transformaciones de x a u , y de θ a τ , tal que τ es un parámetro de localización para u . Estas condiciones serán referidas como (A). En este caso, Lindley llama consistente al argumento fiducial.
2. Si para una muestra aleatoria de cualquier tamaño de la distribución de X , existe una estadística suficiente simple para θ , entonces el argumento fiducial es inconsistente, a menos que se tengan las condiciones (A); cuando esto se cumple, el argumento fiducial es equivalente al argumento Bayesiano cuando se utiliza como distribución inicial la uniforme para τ .

Para demostrarlos se procede como sigue:

1. Sea $F(x | \theta)$ la distribución de la variable aleatoria X , donde el parámetro θ pertenece a un intervalo de \mathbb{R} , digamos (L, U) . Supondremos que F es derivable con respecto a θ y decreciente como función de θ , y además

$$\lim_{\theta \rightarrow U} F(x | \theta) = 0 \quad \text{y} \quad \lim_{\theta \rightarrow L} F(x | \theta) = 1,$$

donde L y U pueden ser $-\infty$ e ∞ , respectivamente. Llamemos (B) a estas condiciones. Entonces de acuerdo con Fisher la función de distribución fiducial de θ está dada por

$$H(\theta; x) = 1 - F(x | \theta),$$

y la función de densidad fiducial está dada como

$$h(\theta; x) = \phi_x(\theta) = -\frac{\partial}{\partial \theta} F(x | \theta) \tag{1.1}$$

Veamos bajo qué condiciones (1.1) se puede obtener bajo un argumento bayesiano, a partir de una distribución inicial $p(\theta)$. Supongamos que existe la densidad de probabilidad

$$f(x | \theta) = \frac{\partial}{\partial x} F(x | \theta).$$

Sea

$$\rho(x) = \int f(x | \theta) p(\theta) d\theta. \tag{1.2}$$

Entonces la densidad final de θ es

$$\frac{f(x | \theta)p(\theta)}{\rho(x)}. \quad (1.3)$$

Busquemos bajo qué condiciones sobre F se puede encontrar una distribución inicial de θ tal que (1.1) y (1.3) sean iguales. De aquí se tiene que se debe encontrar $p(\theta)$ tal que se cumpla la siguiente ecuación

$$-\frac{\partial}{\partial \theta} F(x | \theta) = \frac{\partial}{\partial x} F(x | \theta) \frac{p(\theta)}{\rho(x)}, \quad (1.4)$$

con $\rho(x)$ dado por (1.2). De (1.4) se tiene

$$-\frac{\frac{\partial}{\partial \theta} F(x | \theta)}{\frac{\partial}{\partial x} F(x | \theta)} = \frac{p(\theta)}{\rho(x)}. \quad (1.5)$$

De aquí, una condición necesaria para que (1.4) tenga una solución es que el lado izquierdo de (1.5) sea igual al producto de una función de θ y una función de x . Por lo tanto

$$\frac{1}{\rho(x)} \frac{\partial F}{\partial x} + \frac{1}{p(\theta)} \frac{\partial F}{\partial \theta} = 0.$$

Resolviendo esta ecuación diferencial parcial se concluye que o bien F es constante, lo cual es imposible, o bien

$$F(x | \theta) = G(R(x) - P(\theta)),$$

para alguna función G , donde

$$R(x) = \int_{-\infty}^x \rho(s) ds \quad \text{y} \quad P(\theta) = \int_{-\infty}^{\theta} p(\omega) d\omega$$

Si $U = R(X)$ y $\tau = P(\theta)$, entonces

$$F = G(u - \tau), \quad (1.6)$$

cuando las condiciones (B) se cumplen.

Recíprocamente, supongamos ahora que $F(x | \theta)$ satisface (1.6), donde u es sólo función de x y τ es sólo función de θ , y que además se cumplen las condiciones (B). Sea $w = u - \tau$. Entonces

$$\frac{\partial F}{\partial \theta} = \frac{dG}{dw} \frac{\partial w}{\partial \tau} \frac{d\tau}{d\theta} = \frac{dG}{dw} (-1) \frac{d\tau}{d\theta} = -\frac{dG}{dw} \frac{d\tau}{d\theta},$$

$$\frac{\partial F}{\partial x} = \frac{dG}{dw} \frac{\partial w}{\partial u} \frac{du}{dx} = \frac{dG}{dw} \frac{du}{dx},$$

de donde

$$-\frac{\partial F}{\partial \theta} = \frac{\partial F}{\partial x} \frac{\frac{d\tau}{d\theta}}{\frac{du}{dx}}. \quad (1.7)$$

Ahora,

$$\rho(x) = \int \frac{\partial F}{\partial x} p(\theta) d\theta = \int \frac{dG}{dw} \frac{du}{dx} p(\theta) d\theta = \frac{du}{dx} \int \frac{dG}{dw} p(\theta) d\theta.$$

Si

$$p(\theta) = \frac{d\tau}{d\theta}, \quad (1.8)$$

entonces

$$\begin{aligned} \rho(x) &= \frac{du}{dx} \int \frac{dG}{dw} \frac{d\tau}{d\theta} d\theta = \frac{du}{dx} \int (-1) \frac{dG}{dw} (-1) \frac{d\tau}{d\theta} d\theta = \frac{du}{dx} \int (-1) \frac{dG}{dw} \frac{\partial w}{\partial \tau} \frac{d\tau}{d\theta} d\theta \\ &= \frac{du}{dx} \int (-1) \frac{\partial F}{\partial \theta} d\theta = \frac{du}{dx} \int \frac{\partial}{\partial \theta} (1 - F(x | \theta)) d\theta = \frac{du}{dx}, \end{aligned}$$

es decir,

$$\rho(x) = \frac{du}{dx}. \quad (1.9)$$

De (1.7), (1.8) y (1.9) se tiene

$$-\frac{\partial F}{\partial \theta} = \frac{\partial F}{\partial x} \frac{p(\theta)}{\rho(x)},$$

de donde se sigue (1.4). Así, las condiciones (A) son necesarias y suficientes.

Finalmente, si θ es un parámetro de escala, entonces

$$F(x | \theta) = F_0\left(\frac{x}{\theta}\right) = F_0\left(e^{\log\left(\frac{x}{\theta}\right)}\right) = F_0\left(e^{\log(x) - \log(\theta)}\right) = G_0(u - \tau),$$

donde $F_0(\cdot) = F(\cdot | 1)$, $u = \log(x)$, $\tau = \log(\theta)$ y G_0 es la función de distribución de $\log(X)$ evaluada en u , cuyo parámetro τ es de localización; esto es, $G_0(u - \tau) = G(u | \tau)$.

De (1.8), la distribución inicial de θ es $p(\theta) = \theta^{-1}$.

Como aquí se demuestra, cuando el parámetro de la distribución es de localización, entonces una distribución inicial *natural* de él es la uniforme, la cual denota *ignorancia total* acerca del mismo. Dado que la distribución fiducial se obtiene sólo a partir de la muestra, por este motivo, cuando el parámetro es de localización o escala, o bien existen las transformaciones indicadas por Lindley, la distribución fiducial y la correspondiente distribución final bayesiana coinciden.

2. En el análisis anterior, sea X una estadística suficiente para θ . Sea Y una estadística suficiente para θ , obtenida de una segunda muestra, independiente de la primera. Sea $\psi_{x,y}(\theta)$ la distribución final de Bayes para θ , obtenida usando el teorema de Bayes para Y con distribución inicial $\phi_x(\theta)$. Sea $\phi_{x,y}(\theta)$ la distribución fiducial para θ , dados x, y , cuando inicialmente nada se conoce sobre θ . Veamos cuando se cumple que $\phi_{x,y}(\theta) = \psi_{x,y}(\theta)$. Sólo se considera el caso en el que X y Y tienen la misma distribución, la cual admite una estadística suficiente simple para cualquier tamaño de muestra. Suponemos que también se cumplen las condiciones (B). Entonces la distribución tiene una densidad de la forma

$$f(x)g(\theta)e^{u(x)h(\theta)},$$

la que sin pérdida de generalidad se puede escribir como

$$f(x)g(\theta)e^{x\theta} = f(x | \theta). \quad (1.10)$$

Puesto que

$$f(x | \theta)f(y | \theta) = f(x)f(y) (g(\theta))^2 e^{(x+y)\theta},$$

la distribución conjunta de X y Y admite a $X + Y$ como una estadística suficiente, de donde $\phi_{x,y}(\theta)$ sólo es función de $x + y$ y θ . Veamos bajo qué condiciones $\psi_{x,y}(\theta)$ sólo es función de $x + y$ y θ . Se tiene

$$\psi_{x,y}(\theta) = \frac{\phi_x(\theta)f(y | \theta)}{\int \phi_x(\theta)f(y | \theta)d\theta}.$$

Si el lado derecho de la igualdad anterior sólo es función de $x + y$ y θ , entonces $\psi_{x,y}(\theta')$ también es función sólo de $x + y$ y θ' . Puesto que

$$\frac{\psi_{x,y}(\theta)}{\psi_{x,y}(\theta')} = \frac{\phi_x(\theta)f(y | \theta)}{\phi_x(\theta')f(y | \theta')},$$

entonces

$$\frac{\phi_x(\theta)f(y | \theta)}{\phi_x(\theta')f(y | \theta')},$$

es función sólo de $x + y, \theta$ y θ' , donde $\theta \neq \theta'$; por lo tanto esta expresión no debe ser alterada al intercambiar x por y , es decir,

$$\frac{\phi_x(\theta)f(y | \theta)}{\phi_x(\theta')f(y | \theta')} = \frac{\phi_y(\theta)f(x | \theta)}{\phi_y(\theta')f(x | \theta')}.$$

Sustituyendo las expresiones para $f(y | \theta)$ y $f(y | \theta')$ dadas por (1.10) tenemos

$$\frac{\phi_x(\theta)g(\theta)e^{y\theta}}{\phi_x(\theta')g(\theta')e^{y\theta'}} = \frac{\phi_y(\theta)g(\theta)e^{x\theta}}{\phi_y(\theta')g(\theta')e^{x\theta'}},$$

o sea,

$$\frac{\phi_x(\theta)\phi_y(\theta')}{\phi_x(\theta')\phi_y(\theta)} = e^{(x-y)(\theta-\theta')}.$$

De esta última igualdad se tiene

$$\phi_x(\theta) = \frac{\phi_x(\theta')\phi_y(\theta)}{\phi_y(\theta')} e^{-x\theta'} e^{-y\theta} e^{y\theta'} e^{x\theta}.$$

Considerando a θ' y y como constantes, tal que las únicas variables son θ y x , de la igualdad anterior obtenemos

$$\phi_x(\theta) = F(x)G(\theta)e^{x\theta},$$

donde F y G son funciones indeterminadas. De aquí, por (1.10),

$$\frac{\phi_x(\theta)}{f(x|\theta)} = \frac{F(x)G(\theta)}{f(x)g(\theta)},$$

pero,

$$\frac{\phi_x(\theta)}{f(x|\theta)} = -\frac{\frac{\partial}{\partial\theta}F(x|\theta)}{\frac{\partial}{\partial x}F(x|\theta)},$$

de donde se sigue que

$$-\frac{\frac{\partial}{\partial\theta}F(x|\theta)}{\frac{\partial}{\partial x}F(x|\theta)} = \frac{F(x)G(\theta)}{f(x)g(\theta)}.$$

Así, se tiene la igualdad (1.5), la cual sólo es posible bajo las condiciones (A). Bajo estas condiciones vimos que la distribución fiducial $\phi_x(\theta)$ es una distribución final de Bayes con distribución inicial dada por (1.8), donde τ es el parámetro de localización, el cual es función sólo de θ . Entonces se tiene que $\phi_{x,y}(\theta) = \psi_{x,y}(\theta)$ porque el teorema de Bayes es consistente. De aquí las condiciones (A) son necesarias y suficientes. Así, el argumento fiducial es consistente si, y sólo si, es equivalente a un argumento Bayesiano bajo las condiciones (A), con distribución inicial dada por (1.8).

El siguiente es un ejemplo en donde no se cumple la igualdad $\phi_{x,y}(\theta) = \psi_{x,y}(\theta)$: Sea la función de densidad

$$f(x|\theta) = \frac{\theta^2}{\theta+1} (x+1) e^{-x\theta}, \quad x > 0, \quad \theta \geq 0, \quad (1.11)$$

la cual es de la forma (1.10). La distribución de $Z = X + Y$ es

$$f_z(z|\theta) = \left(\frac{\theta^2}{\theta+1}\right)^2 \left(z + z^2 + \frac{1}{6}z^3\right) e^{-z\theta},$$

de donde se deduce que

$$\phi_{x,y}(\theta) = \frac{e^{-z\theta}}{(\theta+1)^3} \left[\theta^3 \left(2z^2 + \frac{4}{3}z^3 + \frac{1}{6}z^4 \right) + \theta^4 \left(z^2 + z^3 + \frac{1}{6}z^4 \right) \right]. \quad (1.12)$$

De (1.11),

$$\phi_x(\theta) = \frac{\theta}{(\theta+1)^2} [1 + (1+\theta)(1+x)] x e^{-x\theta}, \quad (1.13)$$

luego, puesto que $\psi_{x,y}(\theta) \propto f(y|\theta)\phi_x(\theta)$, de (1.11) y (1.13) se tiene que

$$\psi_{x,y}(\theta) \propto \frac{\theta^3}{(\theta+1)^3} x(1+y) [1 + (1+\theta)(1+x)] e^{-(x+y)\theta}, \quad (1.14)$$

donde la constante de proporcionalidad involucra sólo a x y y . De (1.12) y (1.14) se tiene que $\phi_{x,y}(\theta) \neq \psi_{x,y}(\theta)$. También de (1.14) se tiene que $\psi_{x,y}(\theta) \neq \psi_{y,x}(\theta)$, es decir, el orden de las dos muestras afecta el análisis. También se comprueba que las distribuciones de X y Z satisfacen las condiciones (B).

En [12] Fisher afirmó que $\phi_{x,y}(\theta) = \psi_{x,y}(\theta)$, pero no lo demostró. Como acabamos de ver, tal afirmación no es cierta en general.

3. La familia exponencial y las condiciones (A)

Del análisis hecho en 2, de manera natural surge la pregunta: ¿bajo qué condiciones un miembro de la familia exponencial, dada por (1.10), satisface las condiciones (A)? La familia está dada como

$$f(x)g(\theta)e^{x\theta}, \quad (1.15)$$

donde

$$\int_{-\infty}^{\infty} f(x)e^{x\theta} dx < \infty$$

para todo θ en su rango de valores. De (1.15), para cada elemento $f(x)g(\theta)e^{x\theta}$ se tiene

$$\int_{-\infty}^{\infty} f(x)e^{x\theta} dx = g(\theta)^{-1}.$$

Se supone que las distribuciones están escritas en la forma (1.10). Si existen $u(x)$ y $\tau(\theta)$ tales que

$$F(x|\theta) = G(u + \tau)$$

(se evita el signo menos al escribir $u + \tau$ en lugar de $u - \tau$), entonces, haciendo $w = u + \tau$ y puesto que

$$F(x | \theta) = \int_{-\infty}^x f(s(u)) g(\theta(\tau)) e^{\theta(\tau)s(u)} ds,$$

se tiene

$$\frac{\partial F}{\partial u} = \frac{\partial F}{\partial x} \frac{dx}{du} = \frac{dx}{du} \frac{\partial F}{\partial x} = \frac{dx}{du} f(x(u)) g(\theta(\tau)) e^{\theta(\tau)x(u)} = \frac{dG}{dw},$$

ya que también

$$\frac{\partial F}{\partial u} = \frac{\partial G}{\partial u} = \frac{dG}{dw} \frac{dw}{du} = \frac{dG}{dw}.$$

Así, se requiere encontrar las condiciones bajo las cuales existen $u(x)$ y $\tau(\theta)$ tales que

$$\frac{dx}{du} f(x(u)) g(\theta(\tau)) e^{\theta(\tau)x(u)}$$

sea una función sólo de $u + \tau$. Así pues, supongamos que

$$\frac{dx}{du} f(x(u)) g(\theta(\tau)) e^{\theta(\tau)x(u)} = \frac{dG}{dw}. \quad (1.16)$$

Tomando logaritmos en (1.16) y derivando, primero con respecto a u y después con respecto a τ , se obtiene que

$$\theta'(\tau) x'(u)$$

es función sólo de $u + \tau$. Ya que $\log(\theta'(\tau))$ es sólo función de τ y $\log(x'(u))$ es sólo función de u , entonces se tiene que

$$\theta'(\tau) x'(u) = e^{a(u+\tau)+b},$$

donde a y b son constantes. No se pierde generalidad al suponer $b = 0$. Luego

$$\log(\theta'(\tau)) = a\tau \text{ y } \log(x'(u)) = au.$$

Hay dos casos a considerar, $a \neq 0$ y $a = 0$. En el primero

$$\theta(\tau) = c_1 e^{a\tau} + b_1, \quad x(u) = c_2 e^{au} + b_2, \quad (1.17)$$

y en el último

$$\theta(\tau) = c_1 \tau + b_1, \quad x(u) = c_2 u + b_2, \quad (1.18)$$

donde b_1, b_2, c_1 y c_2 son constantes. Consideremos primero (1.17). Dado que

$$\begin{aligned} e^{\theta(\tau)x(u)} &= \exp((c_1 e^{a\tau} + b_1)(c_2 e^{au} + b_2)) \\ &= \exp(c_1 c_2 e^{a(u+\tau)} + c_1 b_2 e^{a\tau} + b_1 c_2 e^{au} + b_1 b_2) \\ &= \exp(c_1 c_2 e^{a(u+\tau)}) \exp(c_1 b_2 e^{a\tau}) \exp(b_1 c_2 e^{au}) \exp(b_1 b_2), \end{aligned}$$

vemos que (1.16) se puede escribir como

$$f_1(u)g_1(\tau) \exp(ce^{a(u+\tau)}) = \frac{dG}{dw}, \quad (1.19)$$

de donde se sigue que $f_1(u)g_1(\tau)$ es función sólo de $u + \tau$; como antes, obtenemos

$$f_1(u) = \beta_1 e^{\alpha u + d_1}, \quad g_1(\tau) = \beta_2 e^{\alpha \tau + d_2}.$$

De aquí, el lado izquierdo de (1.19) se convierte en

$$\gamma e^{\alpha(u+\tau)} \exp(ce^{a(u+\tau)}), \quad (1.20)$$

o, de (1.17), en términos de x y θ ,

$$Cx^{-1} (\theta x)^{\frac{\alpha}{a}} e^{\theta x}, \quad (1.21)$$

donde C se elige de modo que la integral sea uno. Notemos que al derivar la familia exponencial consideramos el caso en el que el rango de x no depende del parámetro: de aquí, en (1.20) el rango de u y τ debe ser todo \mathbb{R} , luego en (1.21) los rangos de x y θ deben ser las semirectas reales tales que uno de sus extremos es el origen. Se puede suponer que $x \geq 0$, cuando por convergencia se debe tener $\theta < 0$ y $\alpha/a > 0$. Bajo estas condiciones se puede encontrar C ; escribiendo $\alpha/a = \lambda$ y cambiando el signo de θ se obtiene

$$\frac{\theta^\lambda x^{\lambda-1} e^{-x\theta}}{\Gamma(\lambda)}, \quad (1.22)$$

la distribución gamma, pero donde λ es conocido y θ es el único parámetro desconocido. Consideraciones análogas, pero ahora con (1.18) en lugar de (1.17), permiten obtener, en lugar de (1.22),

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \frac{(x-\theta)^2}{\sigma^2}\right), \quad (1.23)$$

la distribución normal, pero con σ conocido y θ el único parámetro desconocido (ver [19]).

Las distribuciones (1.22) y (1.23) son las únicas de la familia exponencial que satisfacen las condiciones (A), y por lo tanto el argumento fiducial es consistente sólo cuando el parámetro es de localización o una transformación de éste.

Estos resultados fueron muy importantes, ya que con ellos Lindley por primera vez estableció claramente, en el caso de una distribución uniparamétrica, las condiciones bajo las cuales la distribución fiducial coincide con una distribución final bayesiana.

1.3. El método fiducial e invarianza

Uno de los requisitos esenciales que Fisher [12] estableció para el uso del método fiducial es la ausencia de información previa respecto al parámetro que se está estimando. D. A. S. Fraser [13] interpretó este requisito en el sentido de que todos los valores del parámetro son equivalentes en la manera en la cual la distribución de probabilidad de la variable observable está relacionada al valor del parámetro que determina esa distribución. Como consecuencia de tal interpretación, en 1961 [13] presentó un modelo matemático que involucra un conjunto de transformaciones en el que desarrolló el argumento fiducial. Una consecuencia del modelo es que la información que sobre el parámetro proporciona un valor observado de la variable, está en la forma de una distribución de probabilidad, la distribución fiducial, la cual tiene una interpretación frecuentista en términos de una clase bien definida de muestreo repetido. Esto coincide con la afirmación de Fisher [12] de que *“el argumento fiducial usa las observaciones para cambiar el estatus lógico del parámetro, de uno en el cual nada se sabe de él, y ninguna afirmación probabilística se puede hacer, al estatus de una variable aleatoria que tiene una distribución bien definida”*.

Otra consecuencia de este marco especial es que no hay necesidad de requerir una distribución inicial para el parámetro; resulta, además, que la distribución fiducial coincide con la distribución final bayesiana, cuando en el argumento bayesiano se usa una distribución inicial adecuada.

Una consecuencia más concierne a la información inicial: los valores del parámetro quedan restringidos a algún rango específico. Esta restricción se puede usar para condicionar la distribución fiducial, lo cual resulta en una distribución fiducial condicionada. Una combinación probabilística de tales restricciones puede generar una distribución inicial, y una combinación apropiada de distribuciones fiduciales condicionales produce la distribución final bayesiana.

También considera el caso de dos conjuntos separados de variables, donde cada uno produce una distribución fiducial respecto a un parámetro de interés. Propone un método lógico para combinar tales distribuciones fiduciales, de modo que se obtenga una distribución fiducial resultante. Esta misma distribución se obtiene cuando, bajo el enfoque bayesiano, la distribución fiducial de uno de los conjuntos se toma como la distribución inicial para el otro conjunto. Se obtiene la misma distribución cuando el sistema combinado produce una estadística suficiente, y por lo tanto una distribución fiducial.

Por ser de interés para nosotros, de este trabajo de Fraser [13] presentamos las partes que consideramos más relevantes para nuestros propósitos, a las cuales por conveniencia se

les han hecho algunas modificaciones en la forma de presentar ciertos aspectos, aunque sin alterar su contenido.

1.3.1. El modelo de las transformaciones

Se supone que se cumplen las condiciones siguientes:

1. Existen un espacio muestral básico \mathcal{X} y una familia de funciones de densidad de probabilidad $f(x; \theta)$, donde $\theta \in \Omega$, con Ω el espacio parametral.
2. Existe $T : \mathcal{X} \rightarrow \mathcal{T}$, una estadística suficiente minimal de θ , donde $\mathcal{T} = T(\mathcal{X})$.
3. Existe un grupo (\mathcal{G}, \cdot) de transformaciones de \mathcal{X} en \mathcal{X} . Para cualesquiera $g_1, g_2 \in \mathcal{G}$, $g_1 \cdot g_2 = g_1 \circ g_2$, la composición de g_1 y g_2 , la cual denotaremos por $g_1 g_2$.
4. Cada transformación $g \in \mathcal{G}$ induce una transformación, la cual, por simplicidad y por no haber riesgo de confusión, también denotaremos por g , de modo que $g : \mathcal{T} \rightarrow \mathcal{T}$ y $T(g(x)) = g(T(x))$ para cada $x \in \mathcal{X}$.
5. Para cualesquiera $t_1, t_2 \in \mathcal{T}$, existe una única $g \in \mathcal{G}$ que satisface $g(t_1) = t_2$. Entonces, relacionando un punto de referencia arbitrario $t_0 \in \mathcal{T}$ con la identidad $e \in \mathcal{G}$, cualquier $t \in \mathcal{T}$ se puede identificar con la transformación única g_t tal que $g_t(t_0) = t$.
6. Cada transformación $g \in \mathcal{G}$ induce una transformación g^* de Ω sobre Ω , de modo que si T tiene distribución con parámetro θ , entonces $g(T)$ tiene distribución con parámetro $g^*(\theta)$.
7. Para cualesquiera $\theta, \theta' \in \Omega$ existe una única transformación $g \in \mathcal{G}$ de modo que $g^*(\theta) = \theta'$. Entonces, seleccionando un punto de referencia arbitrario $\theta_0 \in \Omega$, cada $\theta \in \Omega$ se puede identificar con la transformación única g_θ^* tal que $g_\theta^*(\theta_0) = \theta$.

Observación 1.1 De la suposición 5 vemos que de manera natural se induce una operación en \mathcal{T}^2 , de modo que \mathcal{T} y \mathcal{G} son grupos isomorfos. Similarmente, de la suposición 7 se induce una operación en Ω^2 de modo que \mathcal{G} y Ω son grupos isomorfos. Así, \mathcal{G} , \mathcal{T} y Ω son grupos isomorfos.

Ejemplo 1.2 Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de una distribución $N(\mu, \sigma^2)$. En este caso $\mathcal{X} = \mathbb{R}^n$. El parámetro es $\theta = (\mu, \sigma)$ y $\Omega = (-\infty, \infty) \times (0, \infty)$. La estadística suficiente minimal es $T = (\bar{X}, S)$, donde \bar{X} es la media muestral y S es la desviación estándar muestral de (X_1, X_2, \dots, X_n) . $\mathcal{T} = (-\infty, \infty) \times (0, \infty)$. Notemos que $\mathcal{T} = \Omega$.

Sea $[a, b]$ una transformación que mueve el origen en a unidades y cambia la escala por un factor positivo b , definida como

$$[a, b](x_1, x_2, \dots, x_n) = (a, b) \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} = (a + bx_1, a + bx_2, \dots, a + bx_n),$$

la cual también está dada por

$$\begin{aligned} [a, b](x_1, x_2, \dots, x_n) &= (0, 1) \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \\ &= (0, 1) \begin{pmatrix} 1 & 1 & \dots & 1 \\ a + bx_1 & a + bx_2 & \dots & a + bx_n \end{pmatrix} \\ &= (a + bx_1, a + bx_2, \dots, a + bx_n). \end{aligned}$$

Sean \mathcal{G} la clase de transformaciones

$$\mathcal{G} = \{[a, b] : a \in \mathbb{R}, b > 0\},$$

y \mathcal{M} el grupo de matrices

$$\mathcal{M} = \left\{ \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} : a \in \mathbb{R}, b > 0 \right\}.$$

Puesto que

$$\begin{aligned} ([a, b][c, d])(x_1, \dots, x_n) &= [a, b]([c, d](x_1, \dots, x_n)) \\ &= [a, b] \left((c, d) \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \right) \\ &= [a, b](c + dx_1, c + dx_2, \dots, c + dx_n) \\ &= (0, 1) \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ c + dx_1 & c + dx_2 & \dots & c + dx_n \end{pmatrix} \\ &= (0, 1) \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \\ &= (0, 1) \begin{pmatrix} 1 & 0 \\ a + bc & bd \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix}, \end{aligned}$$

se tiene entonces que la composición en \mathcal{G} es

$$[a, b][c, d] = [a + bc, bd],$$

y que \mathcal{G} es un grupo isomorfo a \mathcal{M} . Definiendo en \mathcal{T} el producto como

$$(a, b)(c, d) = (a + bc, bd),$$

\mathcal{T} es un grupo isomorfo a \mathcal{G} . Así \mathcal{T} , \mathcal{G} y \mathcal{M} son isomorfos.

Ahora,

$$\begin{aligned} T([a, b](x_1, x_2, \dots, x_n)) &= T(a + bx_1, a + bx_2, \dots, a + bx_n) \\ &= (a + b\bar{x}, bs) \\ &= (a, b) \begin{pmatrix} 1 & 0 \\ \bar{x} & s \end{pmatrix} \\ &= (0, 1) \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \bar{x} & s \end{pmatrix}, \end{aligned}$$

luego la transformación inducida en \mathcal{T} es

$$[a, b](\bar{x}, s) = (a, b) \begin{pmatrix} 1 & 0 \\ \bar{x} & s \end{pmatrix} = (0, 1) \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \bar{x} & s \end{pmatrix},$$

de modo que

$$T([a, b](x_1, x_2, \dots, x_n)) = [a, b](T(x_1, x_2, \dots, x_n)).$$

La clase de transformaciones inducida en \mathcal{T} también es un grupo isomorfo a \mathcal{M} .

Si (\bar{x}', s') y (\bar{x}'', s'') son dos valores de la estadística suficiente T , existe una única transformación $[a, b]$ de modo que $[a, b](\bar{x}', s') = (\bar{x}'', s'')$. Como $[\bar{x}, s](0, 1) = (\bar{x}, s)$, el isomorfismo

$$[\bar{x}, s] \leftrightarrow (\bar{x}, s),$$

es equivalente a elegir como punto de referencia a $(0, 1)$, de manera que se pueden usar las transformaciones para identificar los puntos en \mathcal{T} .

Una muestra de una distribución $N(\mu, \sigma^2)$ es transformada por $[a, b]$ en una muestra de la distribución $N(a + \mu b, b^2 \sigma^2)$. De aquí, usando la notación anterior, para cada transformación $[a, b] \in \mathcal{G}$ la transformación $[a, b]^*$ inducida está dada por

$$[a, b]^*(\mu, \sigma) = [a, b](\mu, \sigma) = (a + b\mu, b\sigma).$$

Si (μ', σ') y (μ'', σ'') son dos valores del parámetro, existe una única transformación $[a, b]$ de modo que $[a, b](\mu', \sigma') = (\mu'', \sigma'')$. Puesto que $[\mu, \sigma](0, 1) = (\mu, \sigma)$, el isomorfismo

$$[\mu, \sigma] \leftrightarrow (\mu, \sigma),$$

es equivalente a elegir como punto de referencia a $(0, 1)$, de modo que se pueden usar las transformaciones para identificar los puntos en Ω .

Debido a que \mathcal{G} y \mathcal{M} son isomorfos, podemos representar a los elementos $[a, b]$ como

$$\begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix}.$$

Análogamente, es conveniente hacer lo mismo con los elementos de \mathcal{T} y Ω .

1.3.2. Una cantidad pivotal

Comúnmente las distribuciones fiduciales se obtienen a partir de una cantidad pivotal, la cual es una función del parámetro y de la estadística suficiente considerada; tiene una distribución fija, independiente del valor del parámetro.

Aquí se desarrolla una cantidad pivotal, la cual se demuestra que es única en el sentido siguiente: cualquier otra cantidad pivotal invariante en forma bajo las transformaciones en el grupo, es una función de la cantidad pivotal. Una consecuencia de este resultado es que si se acepta el principio de que el procedimiento de estimación debería ser invariante con respecto a la transformación natural, entonces la distribución fiducial estará determinada de manera única.

El desarrollo que aquí se hace es con referencia al modelo en 1.3.1. Sean t_0 y θ_0 puntos de referencia fijos pero arbitrarios en \mathcal{T} y Ω , respectivamente. Para cada $t \in \mathcal{T}$, sea g_t el único elemento de \mathcal{G} tal que $g_t(t_0) = t$; similarmente, para cada $\theta \in \Omega$ sea h_θ el único elemento de \mathcal{G} tal que $h_\theta^*(\theta_0) = \theta$.

Como una transformación en \mathcal{T} , a una variable con distribución indizada por θ_0 , h_θ le asigna una variable con distribución etiquetada por θ . Por lo tanto, a una variable con distribución indizada por θ , h_θ^{-1} le asigna una variable con distribución indizada por θ_0 .

Proposición 1.3 *Se tiene que $T = g_T(t_0)$. Además, la variable siguiente es una cantidad pivotal*

$$h_\theta^{-1}(g_T(t_0)). \tag{1.24}$$

Demostración. Puesto que para cada $t \in \mathcal{T}$, g_t es el único elemento de \mathcal{G} tal que $g_t(t_0) = t$, entonces $T(x) = g_{T(x)}(t_0)$. Ahora, ya que $g_T(t_0)$ tiene distribución dependiente de θ , entonces $h_\theta^{-1}(g_T(t_0))$, que es función de θ y T , es una variable con distribución dependiente de θ_0 ; es decir, $h_\theta^{-1}(g_T(t_0))$ es función de θ y T , y tiene una distribución fija e independiente del valor del parámetro θ . Así, la variable $h_\theta^{-1}(g_T(t_0))$ es una cantidad pivotal. ■

Proposición 1.4 *Como una función de T y θ , la cantidad pivotal $h_\theta^{-1}(g_T(t_0))$ es invariante bajo transformaciones en \mathcal{G} .*

Demostración. Sea $f \in \mathcal{G}$; bajo f , la variable $T = g_T(t_0)$ es transformada en $f(T) = f(g_T(t_0))$, y el parámetro $\theta = h_\theta^*(\theta_0)$ de la distribución de T es transformado en el parámetro $f^*(\theta) = f^*(h_\theta^*(\theta_0))$ de la distribución de $f(T)$. De aquí

$$h_\theta^{-1}(g_T) = (f(h_\theta))^{-1}(f(g_T)).$$

■

Proposición 1.5 *Cualquier función de T y θ que es invariante con respecto a \mathcal{G} es función de la pivotal $h_\theta^{-1}(g_T)$.*

Demostración. Sea $P(t; \theta)$ invariante con respecto a \mathcal{G} : $P(f(t); f(\theta)) = P(t; \theta)$ para $f \in \mathcal{G}$. Usando esta invarianza,

$$\begin{aligned} P(t; \theta) &= P(g_t(t_0); h_\theta(\theta_0)) \\ &= P(h_\theta^{-1}(g_t(t_0)); h_\theta^{-1}(h_\theta(\theta_0))) \\ &= P(h_\theta^{-1}(g_t(t_0)); \theta_0). \end{aligned}$$

Así, $P(t; \theta)$ ha sido expresada como una función de $h_\theta^{-1}(g_t)$. De aquí se sigue que cualquier cantidad pivotal invariante es una función de la cantidad pivotal particular $h_\theta^{-1}(g_T)$. Así, esencialmente la pivotal $h_\theta^{-1}(g_T)$ es única. ■

Ejemplo 1.6 *Consideremos el mismo caso que en el ejemplo 1.2. El punto de referencia que se usa en \mathcal{T} y Ω es $(0, 1)$, que convenimos en representar como*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

De aquí, puesto que

$$g_T = T = \begin{pmatrix} 1 & 0 \\ \bar{X} & S \end{pmatrix}, \quad h_\theta = \theta = \begin{pmatrix} 1 & 0 \\ \mu & \sigma \end{pmatrix}$$

y

$$h_\theta^{-1} = \theta^{-1} = \begin{pmatrix} 1 & 0 \\ \mu & \sigma \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -\frac{\mu}{\sigma} & \frac{1}{\sigma} \end{pmatrix},$$

entonces de (1.24) la cantidad pivotal es

$$\begin{aligned} g &= h_\theta^{-1}(g_T(t_0)) \\ &= \theta^{-1}T, \end{aligned} \tag{1.25}$$

es decir,

$$\begin{aligned} g &= \begin{pmatrix} 1 & 0 \\ -\frac{\mu}{\sigma} & \frac{1}{\sigma} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \bar{X} & S \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \frac{\bar{X}-\mu}{\sigma} & \frac{S}{\sigma} \end{pmatrix}, \end{aligned}$$

la cual esencialmente es la única cantidad pivotal invariante. Su distribución se puede describir por medio de la variable pivotal

$$g = \begin{pmatrix} 1 & 0 \\ \frac{Z}{\sqrt{n}} & \frac{\xi}{\sqrt{n-1}} \end{pmatrix}, \quad (1.26)$$

donde Z y ξ son variables aleatorias independientes, Z con distribución $N(0,1)$ y ξ con distribución χ con $n-1$ grados de libertad. De la pivotal g tenemos que

$$\frac{\sqrt{n} \left(\frac{\bar{X}-\mu}{\sqrt{\sigma^2}} \right)}{\sqrt{\frac{S^2}{\sigma^2}}} \sim t_{n-1},$$

la cual es esencialmente la única pivotal que se usa para hacer inferencia sobre μ , cuando σ es desconocida.

La discusión anterior muestra que la distribución de T produjo una distribución única para la pivotal g . Lo que quizá es de mayor interés es que cuando esta distribución única se usa junto con la ecuación pivotal $g = h_\theta^{-1}(g_T(t_0))$ describe al modelo completamente. En efecto, sea g una variable cuya distribución es la distribución pivotal única. La ecuación pivotal

$$g = h_\theta^{-1}(g_T(t_0))$$

puede ser resuelta, dando

$$g_T(t_0) = h_\theta(g),$$

$$T = g_T(t_0) = h_\theta(g),$$

lo que demuestra que la distribución de T se obtiene a través de una transformación h_θ del parámetro aplicada a g , la cual es una variable aleatoria obtenida al aplicar la transformación aleatoria g_T al punto de referencia t_0 .

1.3.3. La distribución fiducial

La técnica para obtener una distribución fiducial es la siguiente: el valor observado de la estadística suficiente se sustituye en la ecuación pivotal, la que se trata como una variable aleatoria con la distribución normalmente asociada con ella. En dicha ecuación el parámetro se trata como una variable aleatoria libre, a la cual la cantidad pivotal le transfiere su distribución a través de la expresión que la constituye.

Sea g una transformación aleatoria tal que la distribución de $g(t_0)$ es la fija de la pivotal; sea t el valor observado de la estadística suficiente T ; sea $\hat{\theta}$ la variable aleatoria cuyos valores son los posibles valores del parámetro a la luz de la información muestral y de la distribución de la cantidad pivotal. La ecuación pivotal puede ser resuelta para θ en términos de g y t , como sigue,

$$\begin{aligned} g &= h_{\theta}^{-1}(g_t), \\ h_{\theta} &= g_t(g^{-1}), \\ \hat{\theta} &= g_t(g^{-1}(\theta_0)). \end{aligned} \tag{1.27}$$

De (1.27) obtenemos la distribución fiducial de $\hat{\theta}$ a partir de la distribución fija de la pivotal.

En una aplicación de esta teoría el parámetro θ tiene un valor fijo, pero desconocido. Este valor determina la distribución de la estadística suficiente, la cual a su vez produce el valor observado t . En tal situación θ es fijo, pero esto no significa que no se puedan hacer afirmaciones probabilísticas respecto a él.

En 1.3.2 se mostró que la información frecuentista disponible concerniente a la estadística suficiente y relativa al parámetro, se puede resumir en la distribución fija de la variable pivotal. En términos de “muestreo repetido” de esta distribución fija para la variable $g(t_0)$, existe una distribución de frecuencias de valores $\hat{\theta}$'s, de posibles valores de θ correspondientes al valor observado t . Esta es la distribución fiducial y tiene una interpretación frecuentista.

Ejemplo 1.7 *Se considera otra vez el ejemplo de la distribución normal. En este caso, ya que $g_t = t$ y $g^{-1}(\theta_0) = g^{-1}$, de (1.27) obtenemos*

$$\hat{\theta} = tg^{-1}. \tag{1.28}$$

Nótese la similitud entre (1.25) y (1.28). Es como si al observar un valor t de T , de (1.25) simplemente despejamos θ . De (1.26) la distribución de la cantidad pivotal se puede describir por medio de la variable pivotal

$$g = \begin{pmatrix} 1 & 0 \\ \frac{Z}{\sqrt{n}} & \frac{\xi}{\sqrt{n-1}} \end{pmatrix},$$

donde Z y ξ son variables aleatorias independientes, Z con distribución $N(0, 1)$ y ξ con distribución χ con $n - 1$ grados de libertad. Luego,

$$\begin{aligned}\hat{\theta} &= tg^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ \bar{x} & s \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{Z}{\sqrt{n}} & \frac{\xi}{\sqrt{n-1}} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ \bar{x} & s \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{Z}{\frac{\xi}{\sqrt{n-1}}\sqrt{n}} & \frac{\sqrt{n-1}}{\xi} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \bar{x} - \frac{Z}{\frac{\xi}{\sqrt{n-1}}\sqrt{n}} & \frac{\sqrt{n-1}}{\xi}s \end{pmatrix},\end{aligned}$$

o sea,

$$\begin{pmatrix} 1 & 0 \\ \hat{\mu} & \hat{\sigma} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \bar{x} - \frac{Z}{\frac{\xi}{\sqrt{n-1}}\sqrt{n}} & \frac{\sqrt{n-1}}{\xi}s \end{pmatrix}.$$

Así,

$$\hat{\sigma} = \frac{\sqrt{n-1}}{\xi}s, \quad (1.29)$$

$$\hat{\mu} = \bar{x} - \frac{Z}{\frac{\xi}{\sqrt{n-1}}\sqrt{n}}s, \quad (1.30)$$

es decir,

$$\hat{\mu} = \bar{x} - Z\frac{\hat{\sigma}}{\sqrt{n}}, \quad (1.31)$$

que también puede ser escrita como

$$\hat{\mu} = \bar{x} - \tau\frac{s}{\sqrt{n}},$$

donde τ es una variable aleatoria "t de Student" con $n - 1$ grados de libertad. Las variables fiduciales $\hat{\mu}$ y $\hat{\sigma}$ se obtienen a partir de los valores observados \bar{x} y s por medio de las pivotaes Z y ξ . De (1.29) obtenemos la densidad fiducial marginal de σ , la cual es

$$f_{\sigma}(\sigma) = \frac{1}{\frac{1}{2}(n-3)!} \left(\frac{(n-1)s^2}{2\sigma^2} \right)^{\frac{1}{2}(n-1)} e^{-\frac{(n-1)s^2}{2\sigma^2}} \frac{2}{\sigma}, \quad (1.32)$$

y de (1.31) obtenemos la densidad fiducial condicional de μ dado σ , la cual es

$$f_{\mu|\sigma}(\mu) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2}. \quad (1.33)$$

De (1.32) y de (1.33) obtenemos la densidad fiducial conjunta de μ y σ :

$$\begin{aligned} f_{\mu,\sigma}(\mu, \sigma) &= f_{\mu\sigma}(\mu) f_{\sigma}(\sigma) \\ &= \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2} \frac{1}{\frac{1}{2}(n-3)!} \left(\frac{(n-1)s^2}{2\sigma^2}\right)^{\frac{1}{2}(n-1)} e^{-\frac{(n-1)s^2}{2\sigma^2}} \frac{2}{\sigma}. \end{aligned}$$

Observación 1.8 En este ejemplo vimos cómo de manera directa y sencilla hemos obtenido la densidad fiducial conjunta de μ y σ de la distribución $N(\mu, \sigma^2)$.

Ahora, sea t un valor observado de la estadística suficiente. Sea T la estadística suficiente observable, con distribución que depende de θ , dada por

$$T = h_{\theta}(g(t_0)),$$

donde la aleatoriedad es producida por la transformación aleatoria g , h_{θ} es la transformación del parámetro, y t_0 es el punto de referencia; θ no es fijo, pero un valor de él y un valor observado de g producen un valor observado de T , $T = t$.

La transformación que a T le asigna un valor t es $g_t(g^{-1}(h_{\theta}^{-1}(T)))$, como se puede verificar,

$$\begin{aligned} g_t(g^{-1}(h_{\theta}^{-1}(T))) &= g_t(g^{-1}(h_{\theta}^{-1}(h_{\theta}(g(t_0)))))) \\ &= g_t(t_0) \\ &= t. \end{aligned}$$

Aplicando esta transformación a θ se tiene

$$g_t(g^{-1}(h_{\theta}^{-1}(\theta))) = g_t(g^{-1}(\theta)),$$

la cual es una variable aleatoria generada por la transformación aleatoria g . Esta distribución de θ es justamente la distribución fiducial desarrollada antes en esta sección.

Observación 1.9 Sea X_1, X_2, \dots, X_n una muestra aleatoria de la función de densidad $\frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$ con parámetro de localización μ y parámetro de escala σ , y sean \bar{X} la media muestral y S la desviación estándar muestral. Entonces el conjunto de transformaciones

$$\mathcal{G}_1 = \{g_{a,b} : \mathbb{R}^n \rightarrow \mathbb{R}^n \mid a \in \mathbb{R}, b > 0, g_{a,b}(x_1, x_2, \dots, x_n) = (a + bx_1, a + bx_2, \dots, a + bx_n)\}$$

es un grupo. Además se satisfacen las suposiciones 4-7 del modelo de las transformaciones, y la variable

$$\frac{\bar{X} - \mu}{S}$$

es una cantidad pivotal. Si $\sigma = 1$, entonces $b = 1$ y $\bar{X} - \mu$ es una cantidad pivotal. Si $\mu = 0$, entonces $a = 0$ y $\frac{\bar{X}}{\sigma}$ es una cantidad pivotal. La demostración de estos hechos se hace de forma similar al ejemplo 1.2.

1.4. Ejemplos relevantes sobre la definición de probabilidad fiducial

En 1961, David R. Brillinger [3] señala que a 30 años de haber sido dada a conocer por Fisher, entre los estadísticos no se ha alcanzado un acuerdo respecto a la obtención, manipulación o interpretación de la probabilidad fiducial. La razón para esta falta de acuerdo, y la controversia resultante, posiblemente se deba al hecho de que el método fiducial ha sido presentado como un principio lógico general, sin embargo ha sido ilustrado principalmente por medio de ejemplos particulares, en lugar de hacerlo de manera general.

Un paso esencial en el proceso de inferencia fiducial es la obtención de la distribución fiducial. Entre las propiedades que han sido enfatizadas por Fisher para poder obtener una distribución fiducial genuina, están la suficiencia, la ausencia de cualquier información inicial, y la existencia de una cantidad pivotal. En 1947 Tukey [28] descubrió varios ejemplos que demostraron que la distribución fiducial inducida por un conjunto de cantidades pivotaes, suficientes y suavemente invertibles, no es necesariamente única; es decir, la distribución inducida depende del conjunto particular elegido de cantidades pivotaes. Fisher [12] argumentó que las cantidades pivotaes usadas en estos ejemplos habían sido construidas artificialmente, y se propusieron requerimientos adicionales. En el caso bivariado los requerimientos se reducen a

$$f(x, y \mid \alpha, \beta) = f(x \mid \alpha)f(y \mid x, \alpha, \beta),$$

o bien a

$$f(x, y \mid \alpha, \beta) = f(x \mid \alpha, \beta)f(y \mid x, \beta),$$

con $f(x, y \mid \alpha, \beta)$ la densidad dada. Sin embargo, como se muestra más adelante, ninguno de estos requerimientos es suficiente para asegurar la unicidad la distribución resultante.

En el caso unidimensional, Lindley [19] demostró un teorema en el que establece que una distribución fiducial es una distribución final de Bayes si y sólo si la función de distribución es invariante bajo un grupo de transformaciones continuas del parámetro escalar. Es natural buscar una extensión de este teorema para el caso multi-paramétrico, pues si se puede demostrar que existe un grupo, entonces existe una cantidad pivotal esencialmente única, una interpretación frecuentista y ciertas propiedades de consistencia.

En el caso multiparamétrico se sabe que si la función de distribución es invariante bajo un tipo particular de grupo, que representa al espacio paramétrico, entonces la distribución fiducial es una distribución final de Bayes (ver [13]). Sin embargo, el ejemplo final que aquí se presenta demuestra que una distribución fiducial bivariada puede ser una distribución final de Bayes sin tener fuertes propiedades de invarianza de grupo.

1.4.1. El caso unidimensional

La situación más simple a la cual se le puede aplicar el argumento fiducial es cuando se tiene un parámetro escalar θ y una estadística suficiente escalar T de θ con distribución continua. Hay muy poca controversia acerca de la distribución fiducial que se obtiene en este caso, probablemente por las razones siguientes: (i) se puede mostrar que cualquier cantidad pivotal debe ser función de la función de distribución $F(T, \theta)$, y por lo tanto la distribución fiducial es única, y (ii) cualquier intervalo fiducial de probabilidad α obtenido de la distribución es un intervalo del $100\alpha\%$ de confianza, en el sentido de que si

$$P_f(\theta_{\alpha_1}(T) \leq \theta \leq \theta_{\alpha_2}(T)) = \alpha_2 - \alpha_1 = \alpha,$$

entonces

$$P_\theta(\theta_{\alpha_1}(T) \leq \theta \leq \theta_{\alpha_2}(T)) = \alpha_2 - \alpha_1 = \alpha,$$

donde P_f denota la probabilidad obtenida de la distribución fiducial y P_θ la probabilidad obtenida de la variable aleatoria T .

En el caso unidimensional, la fórmula aceptada para la densidad fiducial es la dada por Fisher,

$$-\frac{\partial F}{\partial \theta} d\theta,$$

donde (i) $F(T, \theta)$ denota la función de distribución de una estadística suficiente T , (ii) θ y T varían continuamente sobre el mismo rango, y (iii) F es una función monótona de θ , y T es una función decreciente de θ . (Estas son las condiciones de Fisher.) La fórmula anterior se toma aquí como la definición de la distribución fiducial unidimensional.

1.4.2. El caso multidimensional

Se han propuesto una gran cantidad de métodos y condiciones para obtener distribuciones fiduciales multivariadas. Parece que el primer método fue obtener la distribución fiducial a través de cantidades pivotaes. Sin embargo, varias personas notaron que si una distribución fiducial se define simplemente como la que induce un conjunto de cantidades pivotaes, entonces ello puede resultar en no-unicidad. La existencia de tales ejemplos motivó a varios autores a proponer condiciones adicionales que se deben satisfacer a fin de obtener una distribución fiducial única.

Fisher dio ejemplos acerca de la obtención de distribuciones fiduciales multivariadas. Si x, y son las variables aleatorias con densidad bivariada $f(x, y | \alpha, \beta)$, los ejemplos parecen indicar que Fisher requería que la densidad tenga la propiedad

$$f(x, y | \alpha, \beta) = f(x | \alpha)f(y | x, \alpha, \beta). \tag{1.34}$$

Entonces la distribución fiducial se obtiene como sigue: de $f(x, \alpha)$ encuentra la densidad fiducial de α , un problema univariado. Después, considerando α fijo, a partir de $f(y | x, \alpha, \beta)$ encuentra una densidad fiducial de β , otra vez un problema univariado. La densidad fiducial conjunta requerida es el producto de las dos últimas densidades fiduciales.

En contraste con Fisher, Quenouille [25] requiere que la densidad pueda factorizarse como

$$f(x, y | \alpha, \beta) = f(x | \alpha, \beta)f(y | x, \beta). \quad (1.35)$$

Justifica esta factorización por medio del siguiente argumento de suficiencia: x es suficiente para α , por lo tanto, para β fijo se puede usar $F(x, \alpha, \beta)$ para obtener la distribución fiducial de α dado β . Ahora, con x fija, y es casi suficiente para β , de donde se sigue que se puede emplear $F(y | x, \beta)$ para obtener la distribución fiducial de β . Otra vez, la distribución fiducial conjunta que se requiere es el producto de las dos distribuciones fiduciales individuales.

En el caso de independencia, las factorizaciones (1.34) y (1.35) llevan a la misma expresión. Esto es lo que pasa cuando en el caso de la distribución $N(\mu, \sigma^2)$ se determina la distribución fiducial conjunta de μ y σ^2 .

Desafortunadamente, en el caso de no-independencia ninguna de las factorizaciones (1.34) y (1.35) es suficiente para obtener una distribución fiducial única. Aquí se van a presentar ejemplos de no-unicidad, los que serán tales que (i) las estadísticas involucradas son suficientes, (ii) las estadísticas y los parámetros tienen el mismo rango, (iii) las cantidades pivotaes son monótonas como funciones de las estadísticas y los parámetros, y (iv) todas las distribuciones fiduciales involucradas son distribuciones propias.

Sea (x, y) con la función de distribución

$$F(x, y | \alpha, \beta) = (1 - e^{-\alpha x}) (1 - e^{-\beta y}) (1 - \gamma e^{-\phi x - \mu y}),$$

con $x, y, \alpha, \beta > 0$; ϕ es función de α , μ es función de β , y γ es una constante que será especificada después.

Ahora procede a deducir la distribución fiducial conjunta de α y β . Primero, las distribuciones marginales de x y y son

$$F(x | \alpha, \beta) = F(x | \alpha) = 1 - e^{-\alpha x} \quad \text{y} \quad F(y | \alpha, \beta) = F(y | \beta) = 1 - e^{-\beta y}, \quad (1.36)$$

respectivamente. Por lo tanto se obtiene la factorización (1.34), y, lo que es más importante, se obtiene para ambas combinaciones de variables aleatorias y parámetros. Se puede mostrar que

$$F(y | x, \alpha, \beta) = (1 - e^{-\beta y}) (1 - \gamma e^{-\phi x - \mu y}) \left[1 - \left(\frac{\phi}{\alpha} \right) (e^{\alpha x} - 1) \right]. \quad (1.37)$$

De (1.36) se obtiene la densidad fiducial marginal de α , la cual es $xe^{-\alpha x}$. La densidad fiducial de β dados x, y, α está dada por $\frac{\partial}{\partial \beta} (F(y | x, \alpha, \beta))$.

La densidad fiducial conjunta de α y β es el producto de las dos últimas densidades. Es

$$xye^{-\alpha x - \beta y} \left(1 - \gamma e^{-(\phi - \alpha)x - (\mu - \beta)y} \left[e^{-\alpha x} - \left(\frac{\phi}{\alpha} \right) (1 - e^{-\alpha x}) \right] \left[e^{-\beta y} - \frac{\partial \mu}{\partial \beta} (1 - e^{-\beta y}) \right] \right), \quad (1.38)$$

donde ahora ϕ, μ y γ pueden ser elegidos de forma tal que se satisfagan las condiciones (i) a (iv). Ejemplos de tales elecciones son

$$\phi = \alpha (1 + e^{-\alpha}), \quad \mu = \beta (1 + e^{-\beta}), \quad |\gamma| \leq \frac{1}{4}.$$

Sin embargo, con esta elección (1.38) no es una función simétrica de (x, α) y (y, β) , mientras que sí lo es la función de distribución. Por lo tanto la densidad fiducial que se obtiene depende de la marginal con la que se empieza. Así, este es un contraejemplo sobre la suposición de la unicidad cuando se tiene la factorización (1.34).

En este caso se puede ver que una condición suficiente para que la distribución fiducial que se obtenga sea única, es que la función de distribución sea una función sólo de las marginales; no obstante, esta condición no es necesaria.

Un ejemplo que demuestra que la factorización (1.35) no es suficiente para asegurar la unicidad se debe a Savage (ver [28]), y es el siguiente: sean $0 \leq x, y, \alpha, \beta < \infty$, donde (x, y) tiene densidad

$$f(x, y | \alpha, \beta) = \left(\frac{\alpha^2 \beta^2}{\alpha + \beta} \right) (x + y) e^{-(\alpha x + \beta y)}. \quad (1.39)$$

Se tiene

$$f(x | \alpha, \beta) = \frac{\alpha^2 \beta^2}{\alpha + \beta} \left(x e^{-\alpha x} \int_0^\infty e^{-\beta y} dy + e^{-\alpha x} \int_0^\infty y e^{-\beta y} dy \right) = \frac{\alpha^2 \beta^2}{\alpha + \beta} \left(\frac{1 + \beta x}{\beta^2} \right) e^{-\alpha x}. \quad (1.40)$$

También,

$$f(y | x, \alpha, \beta) = \frac{f(x, y, \alpha, \beta)}{f(x, \alpha, \beta)} = \frac{f(x, y | \alpha, \beta)}{f(x | \alpha, \beta)} = \frac{\beta^2 (x + y) e^{-\beta y}}{1 + \beta x},$$

es decir,

$$f(y | x, \beta) = \frac{\beta^2 (x + y) e^{-\beta y}}{1 + \beta x}. \quad (1.41)$$

De (1.40) y (1.41) vemos que se satisface la factorización (1.35).

De (1.40) se tiene

$$F(x | \alpha, \beta) = \frac{\alpha^2 \beta^2}{\alpha + \beta} \left(\int_0^x \frac{ue^{-\alpha u}}{\beta} du + \int_0^x \frac{e^{-\alpha u}}{\beta^2} du \right) = 1 - \left(1 + \frac{\alpha \beta x}{\alpha + \beta} \right) e^{-\alpha x},$$

y de (1.41)

$$F(y | x, \beta) = \frac{\beta^2}{1 + \beta x} \left(x \int_0^y e^{-\beta u} du + \int_0^y ue^{-\beta u} du \right) = 1 - \left(1 + \frac{\beta y}{1 + \beta x} \right) e^{-\beta y},$$

de donde tenemos que

$$S = \left(1 + \frac{\alpha \beta x}{\alpha + \beta} \right) e^{-\alpha x}$$

y

$$T = \left(1 + \frac{\beta y}{1 + \beta x} \right) e^{-\beta y}$$

son cantidades pivotaes cuya densidad conjunta es la uniforme en el cuadrado unitario (ver [26]). Sean x y y fijos; puesto que $\lim_{\beta \rightarrow \infty} T = 0$, $\lim_{\beta \rightarrow 0} T = 1$ y

$$\frac{dT}{d\beta} = -\frac{y}{(1 + \beta x)^2} ((1 + \beta x)(1 + \beta x + \beta y) - 1) < 0,$$

entonces considerando a T como función de β , $T : (0, \infty) \rightarrow (0, 1)$ es biyectiva. Del mismo modo, para x, y y β fijos, considerando a S como función de α , $S : (0, \infty) \rightarrow (0, 1)$ es biyectiva, ya que

$$\frac{dS}{d\alpha} = -\frac{\alpha x}{(\alpha + \beta)^2} ((\alpha + 2\beta) + \beta(\alpha + \beta)x) < 0.$$

Puesto que $\frac{\partial T}{\partial \alpha} = 0$, el Jacobiano de la transformación $(S, T) \mapsto (\alpha, \beta)$ es $\frac{\partial T}{\partial \beta} \frac{\partial S}{\partial \alpha} > 0$, y puesto que la densidad conjunta de (S, T) es la uniforme en el cuadrado unitario, la densidad conjunta inducida de (α, β) es

$$h_1(\alpha, \beta) = \frac{\partial T}{\partial \beta} \frac{\partial S}{\partial \alpha} = \frac{\alpha x y}{(\alpha + \beta)^2} e^{-\alpha x - \beta y} \frac{(\alpha + 2\beta + \beta(\alpha + \beta)x) ((1 + \beta x)(1 + \beta x + \beta y) - 1)}{(1 + \beta x)^2}. \quad (1.42)$$

Como la densidad en (1.39) es simétrica con respecto a (α, x) y (β, y) , si ahora $1 - S$ es la función de distribución de y (dados α y β), y $1 - T$ es la función de distribución condicional de x dado y , siguiendo el mismo procedimiento vemos que ahora la densidad conjunta inducida

de (α, β) , $h_2(\alpha, \beta)$, la obtenemos por simetría de (1.42) intercambiando los papeles de α y β y los papeles de x y y . Observamos que

$$h_1(\alpha, \beta) \neq h_2(\alpha, \beta).$$

Finalmente,

$$\frac{\partial S}{\partial x} = -\frac{1 + \beta x}{\alpha + \beta} \alpha^2 e^{-\alpha x} < 0,$$

$$\frac{\partial S}{\partial \beta} = \frac{\alpha^2 x}{(\alpha + \beta)^2} e^{-\alpha x} > 0$$

y

$$\frac{\partial T}{\partial y} = -\frac{x + y}{1 + \beta x} \beta^2 e^{-\beta y} < 0,$$

y así este ejemplo también cumple con las demás propiedades establecidas (i) a (iv), pero no lleva a una distribución fiducial única.

1.4.3. Distribuciones fiduciales multivariadas y el Teorema de Bayes

Brillinger [3] se refiere al teorema de Lindley [19] que aparece en la sección 1.2. Lo enuncia como sigue:

TEOREMA. *Para que la distribución fiducial de θ dada x sea una distribución final de Bayes es necesario y suficiente que exista una transformación de x a μ y de θ a τ tal que τ es un parámetro de localización para μ .*

Una generalización natural de la transformación descrita en el Teorema de Lindley al caso bi-dimensional es la siguiente: existen transformaciones x a $a(x, y, \mu, \nu)$, y a $b(x, y, \mu, \nu)$, α a $c(\alpha, \beta, \mu, \nu)$ y β a $d(\alpha, \beta, \mu, \nu)$, las cuales forman un grupo para los parámetros arbitrarios μ, ν , y tales que la función de distribución satisface $F(x, y, \alpha, \beta) = F(a, b, c, d)$. Si las transformaciones son continuas, entonces se dice que forman un grupo de transformaciones continuas bi-paramétricas, o un grupo de transformaciones de Lie bi-paramétricas.

Enseguida se presenta un ejemplo de una distribución fiducial bivariada, obtenida de una densidad que satisface las factorizaciones (1.34) y (1.35), que también es una densidad final Bayesiana, pero que no es invariante bajo cualquier grupo de transformaciones continuas bi-paramétricas.

Considera la función de distribución bivariada

$$H(x, y) = \frac{1}{(1 + e^{-x-\alpha})(1 + e^{-\phi(\alpha)(y+\beta)})},$$

donde $-\infty < x, y, \alpha, \beta < \infty$ y $\phi > 0$ será especificado después. Este ejemplo satisface las factorizaciones (1.34) y (1.35); además se encuentra que la densidad fiducial conjunta de α y β es

$$\phi(\alpha)e^{-(x+\alpha)-\phi(\alpha)(y+\beta)}H^2. \quad (1.43)$$

Siempre que $\phi(\alpha)$ sea seleccionada de manera tal que la densidad anterior sea integrable en α y β , se ve que (1.43) es igual a

$$\frac{h(x, y, \alpha, \beta)}{\int h(x, y, \alpha, \beta) d\alpha d\beta},$$

o sea, la densidad fiducial es una densidad final de Bayes con densidad inicial uniforme. Un ejemplo de un tipo de función ϕ requerida es $\phi(\alpha) = 2 + \frac{\alpha}{(1+\alpha^2)^{\frac{1}{2}}}$.

Para probar que $H(x, y)$ no es invariante bajo cualquier grupo de transformaciones continuas bi-paramétricas, se requiere el siguiente teorema, el cual está demostrado por Brillinger.

TEOREMA. *La función de distribución $H(x, y, \alpha, \beta)$ es invariante bajo un grupo de transformaciones continuas bi-paramétricas si, y sólo si, existen funciones linealmente independientes*

$$(\epsilon(x, y), \eta(x, y), a(\alpha, \beta), b(\alpha, \beta)), (\bar{\epsilon}(x, y), \bar{\eta}(x, y), \bar{a}(\alpha, \beta), \bar{b}(\alpha, \beta)),$$

donde cada función componente es analítica, tal que

$$\epsilon H_x + \eta H_y + a H_\alpha + b H_\beta = 0, \quad \bar{\epsilon} H_x + \bar{\eta} H_y + \bar{a} H_\alpha + \bar{b} H_\beta = 0.$$

Aplicando este teorema al caso particular bajo consideración. $H(x, y)$ es invariante bajo algún grupo de transformaciones bi-paramétricas sólo si existen $\epsilon(x, y), \eta(x, y), a(\alpha, \beta)$ y $b(\alpha, \beta)$ tales que

$$\frac{\epsilon + a}{e^{x+\alpha} + 1} = - \left[\eta + b + a \frac{\phi'}{\phi} \left(\frac{(y + \beta) \phi}{e^{\phi(y+\beta)}} \right) \right].$$

Se consideran las implicaciones de esta relación al diferenciarla con respecto a x y después dos veces con respecto a β . Combinando las relaciones obtenidas de este modo se ve primero que $\alpha_\beta = \eta_x = 0$. Usando estos resultados se ve que $\epsilon = Ke^x + R$, $a = Ke^{-\alpha} - R$, con K y R constantes. Sustituyendo en las relaciones originales se encuentra $K = R = 0$, lo que implica $\eta + b = 0$, donde cada uno es constante. Por lo tanto se ha encontrado que $\epsilon = 0$, $\eta = C$, $a = 0$, $b = -C$ son las únicas funciones compatibles con la relación dada. Por lo tanto $H(x, y)$ es invariante sólo bajo el grupo de transformaciones uni-paramétricas $y \rightarrow y + A$, $\beta \rightarrow \beta - A$.

Así, el anterior es un contraejemplo a la extensión propuesta del Teorema de Lindley.

Capítulo 2

La distribución fiducial

Haciendo uso de la significancia, o *p-value*, O'Reilly [22] hace una presentación de la distribución fiducial en el caso en el que el parámetro es real y continuo. Por considerarlo relevante, en este capítulo se hace una revisión de dicha presentación, en la cual se analizan las familias exponencial y exponencial natural. Para las distribuciones asociadas a la Bernoulli (binomial, binomial negativa y Poisson), presenta una distribución fiducial única invariante en cuanto al intercambio entre “éxito” o “fracaso”; además se obtiene que la densidad fiducial es la misma en los casos binomial y binomial negativa. Así, la distribución fiducial propuesta satisface el principio fuerte de verosimilitud. Además, en este capítulo también se presentan dos resultados nuevos sobre la distribución fiducial cuando la distribución de la estadística pertenece a la familia exponencial.

2.1. Contrastabilidad monótona y extrema

Sean $f(x; \theta)$ una función de densidad, donde $\theta \in \Theta = (\underline{\theta}, \bar{\theta})$ es un parámetro real, T una estadística suficiente minimal de θ , $g(t; \theta)$ la función de densidad de T , y $G(t; \theta)$ su función de distribución. $G(t; \theta)$ también contiene toda la información relevante para hacer inferencia sobre θ .

Suponemos que valores grandes de θ implican, estadísticamente hablando, valores grandes de T , y también que valores pequeños de θ implican valores pequeños de T . Esto significa que un valor grande de T , digamos $T = t$, da evidencia contra $H_0 : \theta \leq \theta_0$ en el contexto de una prueba de hipótesis, donde la alternativa es $H_1 : \theta > \theta_0$.

En la teoría clásica de prueba de hipótesis, si $T = t$, entonces el valor p asociado es $1 - G(t; \theta_0)$, el cual, al estar dado en términos de la función de distribución, es una función de t monótonamente decreciente, ya que cuando t crece es más fuerte la evidencia contra H_0 .

Si ahora nos fijamos en el valor p , pero dejamos la evidencia t fija y pensamos en valores de θ_0 más discrepantes respecto a la evidencia, digamos $\theta'_0 < \theta_0$, sería lógico esperar que la

cantidad $1 - G(t; \theta'_0)$, el valor p para $H_0 : \theta \leq \theta'_0$, sea menor que el valor p original asociado a $H_0 : \theta \leq \theta_0$, pues ahora la hipótesis nula es más discrepante con respecto a la evidencia. Esto significa que $G(t; \theta)$ debe ser monótonamente decreciente como función de θ para t fijo. Esta propiedad será llamada **contrastabilidad monótona**.

Esta idea puede ser llevada más lejos si pedimos que para los casos donde tiende a ser extrema la discrepancia entre la hipótesis nula y la evidencia, el valor p converja a cero. Si θ_0 es mantenida fija pero t se aproxima a su límite superior, el valor p ciertamente se aproxima a cero ya que $G(t; \theta_0)$ es una función de distribución. En el caso donde la evidencia t se mantiene fija pero movemos a θ_0 más y más lejos respecto a la evidencia, no es obvio, ni es generalmente cierto, que el valor p converge a cero. Esta última propiedad ocurre si $G(t; \theta) \rightarrow 1$ cuando $\theta \rightarrow \underline{\theta}$ y será llamada **contrastabilidad extrema**.

Ya que se puede aplicar el mismo razonamiento a $H_0 : \theta \geq \theta_0$, donde valores pequeños de t dan evidencia contra H_0 , damos las siguientes definiciones:

1. En un problema de inferencia, se dice que se tiene **contrastabilidad monótona** si $G(t; \theta)$ es monótonamente decreciente como función de θ .
2. En un problema de inferencia, se dice que se tiene **contrastabilidad extrema** si, para t fijo, $G(t; \theta) \rightarrow 1$ cuando $\theta \rightarrow \underline{\theta}$ y $G(t; \theta) \rightarrow 0$ cuando $\theta \rightarrow \bar{\theta}$.

2.2. Ejemplos

1. **Parámetro de localización.** Sean θ un parámetro de localización, $\Theta = \mathbb{R}$ y $G(t; \theta) = G_0(t - \theta)$, con G_0 alguna función de distribución conocida. Como función de θ (con t fijo), $G(t; \theta)$ es monótonamente decreciente, es decir, cumple la propiedad de **contrastabilidad monótona**. También,

$$\lim_{\theta \rightarrow -\infty} G(t; \theta) = \lim_{\theta \rightarrow -\infty} G_0(t - \theta) = \lim_{y \rightarrow \infty} G_0(y) = 1,$$

y

$$\lim_{\theta \rightarrow \infty} G(t; \theta) = \lim_{\theta \rightarrow \infty} G_0(t - \theta) = \lim_{y \rightarrow -\infty} G_0(y) = 0,$$

de donde se ve que $G(t; \theta)$ cumple la propiedad de **contrastabilidad extrema**.

2. **Parámetro de escala.** Ahora $\theta > 0$ y $\Theta = \mathbb{R}^+$, con $G(t; \theta) = G_0\left(\frac{t}{\theta}\right)$, donde G_0 es alguna función de distribución conocida; entonces de manera similar a como se hizo en el ejemplo anterior, se ve que se cumplen las propiedades de **contrastabilidad monótona** y **extrema**.

Como ya se vió de manera más general, en los dos ejemplos previos se tiene modelo de grupo, por lo que la distribución fiducial se deduce de la pivotal.

Es interesante observar que cuando se satisface la propiedad de contrastabilidad extrema, entonces $G(t; \theta)$ converge a una distribución degenerada cuando θ converge a los extremos. Así, es interesante ver qué sucede cuando la distribución límite no es degenerada.

3. En Pedersen [24] con $X \sim N(\mu, 1)$, el problema de inferencia es inferir sobre μ^2 . Un problema de inferencia equivalente es inferir sobre $\theta = |\mu|$, donde la correspondiente variable aleatoria $T = |X|$ tiene función de distribución

$$G(t; \theta) = \Phi(t - \theta) - \Phi(-t - \theta), \quad t > 0,$$

en donde Φ es la función de distribución $N(0, 1)$. Se puede verificar que

- i. $G(t; \theta)$ es monótonamente decreciente como función de θ para cada t fijo, pero
- ii. aún cuando $\lim_{\theta \rightarrow \infty} G(t; \theta) = 0$, $\lim_{\theta \rightarrow 0} G(t; \theta) = 1 - 2\Phi(-t) < 1$, para toda $t > 0$.
La interpretación para este límite, el cual es menor que la unidad para $t > 0$ cuando $\theta \rightarrow 0$, es que su complemento con respecto a la unidad es el valor p , el cual es $2\Phi(-t)$, asociado a las hipótesis $H_0 : \theta = 0$ vs $H_1 : \theta > 0$, una vez que t ha sido observado.

2.3. Distribución fiducial

La distribución fiducial, a la que también se le refiere como distribución de confianza o aún como valor p o distribución de significancia, fue introducida por Fisher en 1930 [9], como ya se dijo en el capítulo 1.

Si $G(t; \theta)$ satisface la propiedad de contrastabilidad monótona, entonces la distribución fiducial de θ se define simplemente como

$$H(\theta; t) = 1 - G(t; \theta),$$

si $G(t; \theta)$ es decreciente como función de θ para t fijo, o bien como

$$H(\theta; t) = G(t; \theta),$$

si $G(t; \theta)$ es creciente como función de θ para t fijo.

Si también se satisface la condición de contrastabilidad extrema, entonces la distribución fiducial no tiene masa fiducial en los extremos del intervalo Θ .

El uso de la distribución fiducial H es directo y da intervalos de confianza (fiduciales).

Es digno de mencionar que en la definición de la distribución fiducial, la única condición que debe satisfacer $G(t; \theta)$, además de la continuidad en t , es la de contrastabilidad monótona y extrema. Esto se debe a que $G(t; \theta)$ es en sí misma una cantidad pivotal cuya distribución es la uniforme en $(0, 1)$. Como una consecuencia de esto último, los intervalos de confianza unilaterales o bilaterales obtenidos a partir de una distribución fiducial tendrán cobertura exacta.

En cuanto a la posible existencia de distribuciones fiduciales diferentes, se puede mostrar que una vez que ha sido seleccionada la estadística suficiente minimal T de θ , la distribución fiducial inducida para θ es la misma, ya sea que se use T o cualquier otra estadística $V = s(T)$ que sea función inyectiva de T . Este resultado es relevante porque asegura que bajo continuidad de la distribución de T , y bajo contrastabilidad monótona, la distribución fiducial es única.

2.4. Familia exponencial

Definición 2.1 *Se dice que la familia $\{g^*(t^*; \eta) : \eta \in \mathbb{R}\}$, de funciones de densidad de probabilidad o de funciones de probabilidad, es una **familia exponencial** si $g(t^*; \eta)$ se puede expresar en la forma*

$$g^*(t^*; \eta) = k^*(t^*)c(\eta) \exp(w(\eta)q(t^*)),$$

donde $k^*(\cdot) \geq 0$ y $q(\cdot)$ son funciones de valor real que no dependen de η , y $c(\cdot) \geq 0$ y $w(\cdot)$ son funciones de valor real que no dependen de t^* .

Definición 2.2 *Se dice que la función de densidad de probabilidad $g(t; \theta)$ pertenece a la **familia exponencial natural**, si se puede expresar en la forma*

$$g(t; \theta) = k(t) \exp(\theta t - M(\theta)),$$

con $t \in R_t$, donde R_t es un intervalo abierto que no depende de θ , y donde $\theta \in \Theta$, con

$$\Theta = \{\theta \in \mathbb{R} : M(\theta) \text{ es finito}\}.$$

Si Θ es abierto, entonces se dice que la familia exponencial natural es **regular**.

Conjetura 2.3 *Sea $g(t; \theta) = k(t) \exp(\theta t - M(\theta))$ una familia exponencial, donde $\theta \in (\underline{\theta}, \bar{\theta})$. Entonces $g(t; \theta)$ es regular si y sólo si satisface la propiedad de contrastabilidad extrema.*

Observación 2.4 *Si $g(t; \theta)$ pertenece a la familia exponencial natural, entonces*

$$M(\theta) = \log \left\{ \int_{\mathbb{R}} k(t) e^{\theta t} dt \right\}$$

se conoce como el *cumulante*.

Proposición 2.5 *Si $g(t; \theta)$ pertenece a la familia exponencial natural, entonces $G(t; \theta)$ es monótonamente decreciente como función de θ , para todo $t \in R_t$.*

Demostración. Si

$$M(\theta; t) = \log \left\{ \int_{\{u \leq t\}} k(u) e^{\theta u} du \right\},$$

entonces

$$\begin{aligned} G(t; \theta) &= \int_{\{u \leq t\}} k(u) e^{\theta u - M(\theta)} du = e^{-M(\theta)} \int_{\{u \leq t\}} k(u) e^{\theta u} du \\ &= e^{-M(\theta)} e^{\log \left\{ \int_{\{u \leq t\}} k(u) e^{\theta u} du \right\}} = e^{-M(\theta)} e^{M(\theta; t)} \\ &= e^{M(\theta; t)} e^{-M(\theta)}, \end{aligned}$$

luego

$$\begin{aligned} \frac{d}{d\theta} \log G(t; \theta) &= \frac{dM(\theta; t)}{d\theta} - \frac{dM(\theta)}{d\theta} = \frac{\int_{\{u \leq t\}} u h(u) e^{\theta u} du}{\int_{\{u \leq t\}} h(u) e^{\theta u} du} - \frac{\int_{\mathbb{R}} t h(t) e^{\theta t} dt}{\int_{\mathbb{R}} h(t) e^{\theta t} dt} \\ &= \frac{\int_{\{u \leq t\}} u h(u) e^{\theta u - M(\theta)} du}{\int_{\{u \leq t\}} h(u) e^{\theta u - M(\theta)} du} - \frac{\int_{\mathbb{R}} t h(t) e^{\theta t - M(\theta)} dt}{\int_{\mathbb{R}} h(t) e^{\theta t - M(\theta)} dt} \\ &= E_{\theta}(T | T \leq t) - E_{\theta}(T) \leq 0 \quad \text{para toda } t \in R_t, \end{aligned} \tag{2.1}$$

excepto quizá para t en los extremos de R_t , en cuyo caso es tiene un conjunto de medida cero. Así, para todo valor observado t podemos suponer que $G(t; \theta)$ es monótonamente decreciente como función de θ . ■

Corolario 2.6 *Si $g(t; \theta)$ pertenece a la familia exponencial natural y además $G(t; \theta)$ satisface la propiedad de contrastabilidad extrema, entonces existe la distribución fiducial de θ , la cual está dada por*

$$H(\theta; t) = 1 - G(t; \theta). \tag{2.2}$$

De la proposición 2.5 es sencillo ver que:

Proposición 2.7 Si $g(t; \theta)$ pertenece a la familia exponencial natural y además $G(t; \theta)$ satisface la propiedad de contrastabilidad extrema, entonces la densidad fiducial de θ puede expresarse como

$$h(t; \theta) = \frac{d}{d\theta} H(t; \theta) = G(t; \theta) [E_\theta(T) - E_\theta(T | T \leq t)],$$

aunque esta relación no facilite el cómputo de dicha densidad fiducial.

Demostración. En efecto, por (2.2),

$$\begin{aligned} h(t; \theta) &= -\frac{d}{d\theta} G(t; \theta) \\ &= -G(t; \theta) \frac{d}{d\theta} \log G(t; \theta) \\ &= -G(t; \theta) [E_\theta(T | T \leq t) - E_\theta(T)] \\ &= G(t; \theta) [E_\theta(T) - E_\theta(T | T \leq t)], \end{aligned}$$

por (2.1). ■

La proposición 2.5 se puede generalizar para el caso de la familia exponencial en general.

Corolario 2.8 Sea $g(t; \theta) = h(t)c(\theta) \exp(w(\theta)q(t))$ una familia exponencial, con $t \in R_t$ y $\theta \in \Theta$, donde R_t y Θ son intervalos abiertos y R_t no depende de θ . Si w y q son monótonas y derivables tales que

$$\frac{d}{d\theta} w(\theta) \neq 0 \quad \text{para toda } \theta \in \Theta \quad \text{y} \quad \frac{d}{dt} q(t) \neq 0 \quad \text{para toda } t \in R_t,$$

entonces $G(t; \theta)$ es monótona como función de θ , para todo $t \in R_t$.

Ejemplo 2.9 Distribución gamma. Sea

$$F(x; \alpha) = \frac{1}{\Gamma(\alpha)} \int_0^x u^{\alpha-1} e^{-u} du$$

la función de distribución gamma de la variable aleatoria X , cuya densidad denotamos por f . De la proposición 2.5, se ve que $F(x; \alpha)$ satisface la propiedad de contrastabilidad monótona. Además se comprueba que $F(x; \alpha)$ también satisface la propiedad de contrastabilidad extrema. Así, la densidad fiducial de α , una vez que se observó x , es

$$h(\alpha; x) = -\frac{d}{d\alpha} (F(x; \alpha)) = \int_0^x (\psi(\alpha) - \log(u)) f(u; \alpha) du,$$

donde $\psi(\alpha)$ es la función digamma,

$$\psi(\alpha) = \frac{d}{d\alpha} (\log(\Gamma(\alpha))).$$

Usando la distribución inicial (o a priori) de Jeffreys $\pi(\alpha) \propto [\psi(\alpha)]^{\frac{1}{2}}$, obtenemos una distribución a posteriori que es muy parecida a la distribución fiducial, como se muestra en la figura 2.1, en la que se reproducen las gráficas correspondientes que aparecen en [22]. Las densidades fiducial y a posteriori fueron obtenidas numéricamente para varios valores de x , que es el caso de una muestra de tamaño $n = 1$.

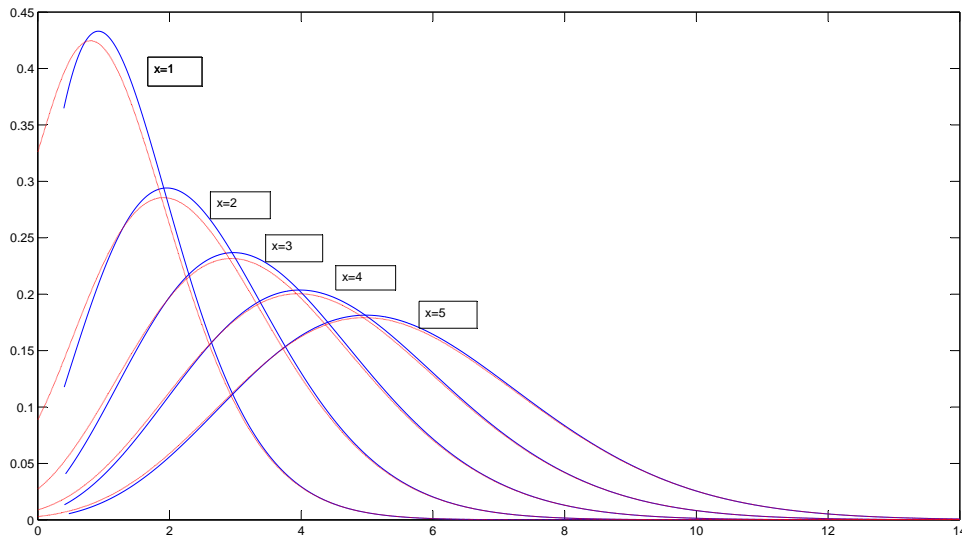


Figura 2.1: Densidad fiducial (en color azul) y densidad final (en color rojo) de α , para el caso $n = 1$.

En general, se tiene el siguiente resultado.

Proposición 2.10 Si x_1, x_2, \dots, x_n son observaciones de la muestra aleatoria X_1, X_2, \dots, X_n de la distribución $\text{gamma}(\alpha, 1)$, la densidad fiducial de α es

$$f_f(\alpha; t) = \int_{-\infty}^t (n\psi(\alpha) - u) g(u; \alpha) du,$$

donde $t = \sum_{i=1}^n \log(x_i)$ es el valor observado de $T = \sum_{i=1}^n \log(X_i)$, la estadística suficiente minimal de α ,

$$g(t; \alpha) = \frac{1}{(\Gamma(\alpha))^n} e^{(\alpha-1)t} h(t)$$

es la densidad de T y $h(t)$ es la densidad de la suma de los logaritmos de una muestra aleatoria de tamaño n de la distribución $\exp(1)$.

Demostración. Antes de encontrar la densidad de T , primero observemos que si W tiene distribución $\exp(1)$, entonces la densidad de $Y = \log(W)$ es

$$h_1(y) = e^y e^{-e^y}.$$

Ahora probemos por inducción sobre el tamaño de muestra n que la densidad de T es

$$g(t; \alpha) = \frac{1}{(\Gamma(\alpha))^n} e^{(\alpha-1)t} h(t), \quad (2.3)$$

donde $h(t)$ es la densidad de la suma de los logaritmos de una muestra aleatoria de tamaño n de la distribución $\exp(1)$.

(a) Si $n = 1$, entonces $T = \log(X_1)$, de donde se tiene que la densidad de T es

$$g(t; \alpha) = f(e^t; \alpha) \left| \frac{dx_1}{dt} \right| = \frac{1}{\Gamma(\alpha)} e^{t(\alpha-1)} e^{-e^t} e^t = \frac{1}{\Gamma(\alpha)} e^{t(\alpha-1)} h_1(t),$$

donde h_1 es la densidad de $\log(W)$, con W una variable aleatoria con distribución $\exp(1)$.

(b) Supongamos que la afirmación se cumple para tamaños de muestra $n - 1$, $n \geq 2$, y probemos que entonces también se cumple para tamaños de muestra n . Sean $T_1 = \log(X_1)$, $T_{n-1} = \sum_{i=2}^n \log(X_i)$, y $g_1(t_1; \alpha)$ y $g_{n-1}(t_{n-1}; \alpha)$ las densidades de T_1 y T_{n-1} respectivamente. Entonces la función de densidad de $T = T_1 + T_{n-1}$ es

$$\begin{aligned} g(t; \alpha) &= \int_{-\infty}^{\infty} g_1(u; \alpha) g_{n-1}(t - u; \alpha) du \\ &= \int_{-\infty}^{\infty} \frac{1}{\Gamma(\alpha)} e^{(\alpha-1)u} h_1(u) \frac{1}{(\Gamma(\alpha))^{n-1}} e^{(\alpha-1)(t-u)} h_{n-1}(t - u) du \\ &= \frac{1}{(\Gamma(\alpha))^n} e^{(\alpha-1)t} \int_{-\infty}^{\infty} h_1(u) h_{n-1}(t - u) du. \end{aligned}$$

Puesto que h_1 es la densidad de $\log(W_1)$ y h_{n-1} es la densidad de $\sum_{i=2}^n \log(W_i)$, donde las W_i son variables aleatorias independientes con distribución $\exp(1)$, $i = 1, \dots, n$, entonces la densidad de $Y_n = \sum_{i=1}^n \log(W_i) = \log(W_1) + \sum_{i=2}^n \log(W_i)$ es

$$h_n(y_n) = \int_{-\infty}^{\infty} h_1(u) h_{n-1}(y_n - u) du.$$

Así, la función de densidad de T es

$$g(t; \alpha) = \frac{1}{(\Gamma(\alpha))^n} e^{(\alpha-1)t} h_n(t),$$

donde $h_n(t)$ es la densidad de la suma de los logaritmos de una muestra aleatoria de tamaño n de la distribución $\exp(1)$. Haciendo $h(t) = h_n(t)$, la densidad de T es la función (2.3)

De la proposición 2.5, se ve que $G(t; \alpha)$ satisface la propiedad de contrastabilidad monótona. También se comprueba que $G(t; \alpha)$ satisface la propiedad de contrastabilidad extrema. Así, la densidad fiducial de α , una vez que se observó t , es

$$f_f(\alpha; t) = -\frac{d}{d\alpha}(G(t; \alpha)) = \int_{-\infty}^t (n\psi(\alpha) - u)g(u; \alpha)du.$$

■

Ejemplo 2.11 *Distribución exponencial truncada.* Sea la función de distribución

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

Como es señalado por Beltrán en su Tesis de Maestría [1], esta familia aparece en Deemer y Votaw [6], la cual es la referencia más antigua que conocemos. Fue utilizada por Lindqvist y Taraldsen [20], como un ejemplo de una distribución donde no hay una cantidad pivotal derivada de un modelo de grupo; esto es, una cantidad pivotal en el sentido clásico. La estadística suficiente minimal es $T = \sum_{i=1}^n X_i$, donde X_1, \dots, X_n es una muestra aleatoria de $F(x; \theta)$. O'Reilly y Rueda [23] obtuvieron la densidad fiducial de θ , la cual comparan con la densidad final bayesiana cuando la distribución inicial es la no-informativa de Jeffreys (ver [2] y [21]). También obtienen la función de densidad de T ,

$$g(t; \theta) = \left(\frac{\theta}{e^{\theta} - 1}\right)^n I_n(t) e^{\theta t}, \quad t \in (0, n),$$

donde

$$I_n(t) = \frac{1}{(n-1)!} \sum_{j=0}^n \binom{n}{j} (-1)^j (t-j)_+^{n-1}$$

es la función de densidad de la suma de n variables aleatorias independientes con distribución uniforme en $(0, 1)$; $(t-j)_+$ es la parte positiva de $(t-j)$. Se tiene que $g(t; \theta)$ pertenece a la familia exponencial natural. Además se puede demostrar que $\lim_{\theta \rightarrow -\infty} G(t; \theta) = 1$ y que $\lim_{\theta \rightarrow \infty} G(t; \theta) = 0$. Así $G(t; \theta)$ satisface las propiedades de contrastabilidad monótona y extrema, por lo que se puede definir la distribución fiducial de θ , la cual es no defectuosa, es decir, no tiene masa fiducial en ninguno de los extremos. La función de densidad fiducial de θ es

$$h(\theta; t) = A(\theta) \int_0^t (B(\theta) - u) e^{\theta u} I_n(u) du, \quad (2.4)$$

donde

$$A(\theta) = \left(\frac{\theta}{e^\theta - 1} \right)^n$$

y

$$B(\theta) = \left(\frac{n}{\theta(e^\theta - 1)} \right) ((\theta - 1)e^\theta + 1).$$

Aunque la integral en (2.4) no se puede evaluar de manera exacta, se puede aproximar numéricamente. En las figuras 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8 y 2.9 se muestran las gráficas de las densidades fiduciales (en color azul) y las densidades finales Bayesianas correspondientes (en color rojo), las cuales se obtuvieron para cada combinación de (θ, n, t) que aparece en cada figura. La densidad inicial no informativa de Jeffreys,

$$\pi(\theta) \propto \left(\frac{1}{\theta^2} + \frac{1}{2 - (e^\theta + e^{-\theta})} \right)^{\frac{1}{2}},$$

se empleó para obtener las densidades finales,

$$\pi(\theta | t) \propto \pi(\theta) \left(\frac{\theta}{e^\theta - 1} \right)^n e^{\theta t},$$

las cuales fueron obtenidas por medio del algoritmo de Metropolis-Hastings (ver [14] y [16]), como se muestra en el capítulo 3, cuando se aborda la distribución gamma (α, β) .

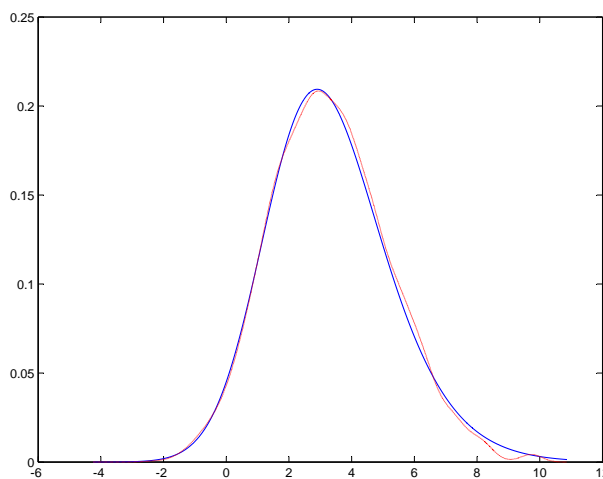


Figura 2.2: $(\theta = 1, n = 5, t = 3.6906)$

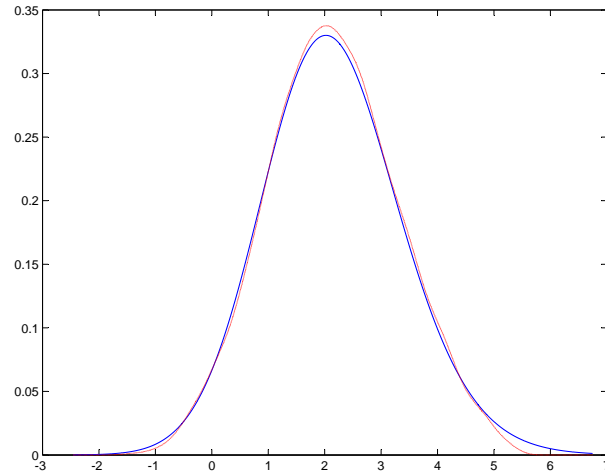


Figura 2.3: ($\theta = 2$, $n = 10$, $t = 6.6743$)

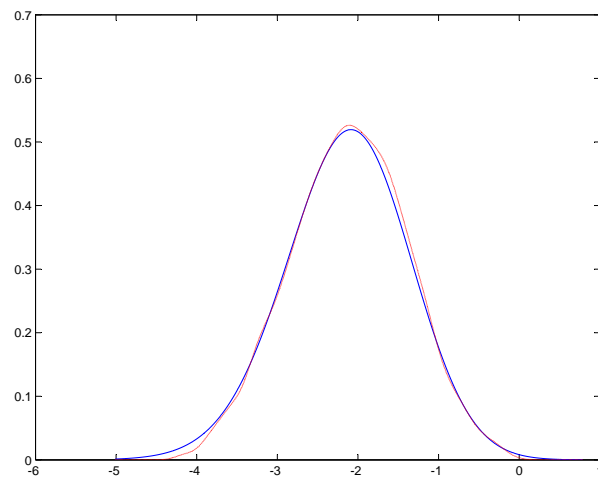


Figura 2.4: ($\theta = -2$, $n = 25$, $t = 8.3469$)

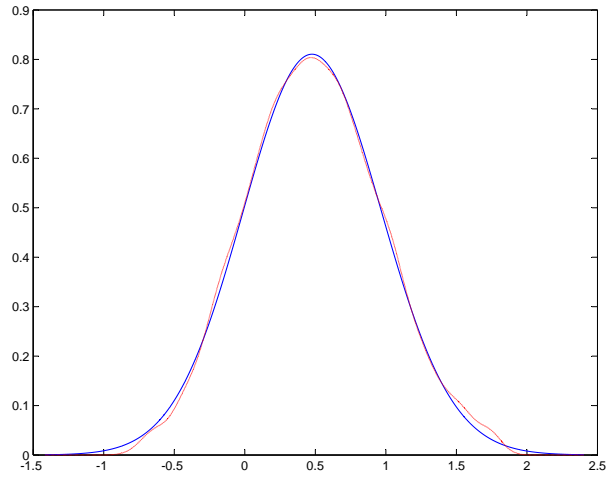


Figura 2.5: ($\theta = 0, n = 50, t = 27.0027$)

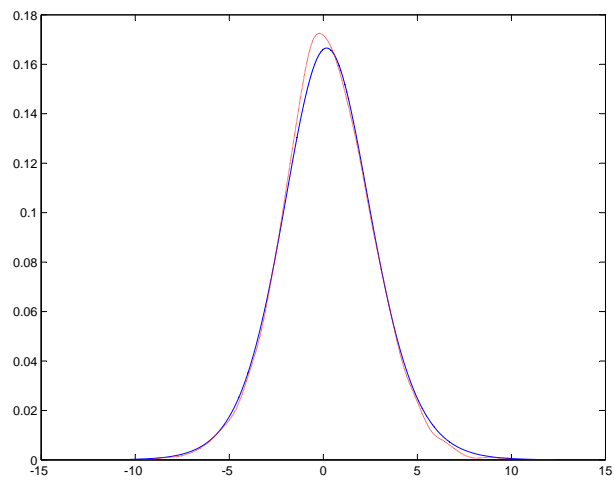


Figura 2.6: ($\theta = 4, n = 2, t = 1.0360$)

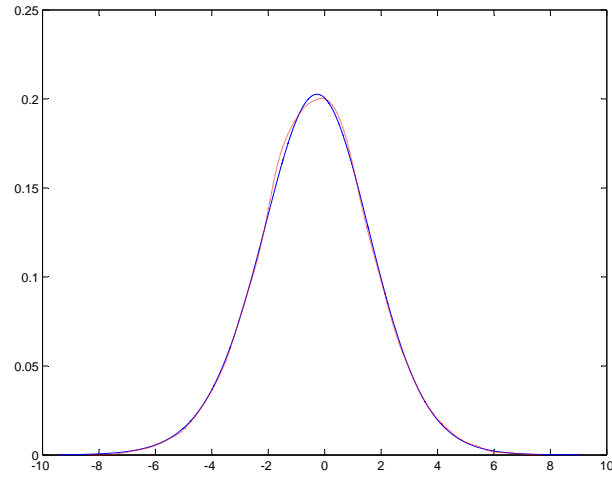


Figura 2.7: ($\theta = 0$, $n = 3$, $t = 1.4168$)

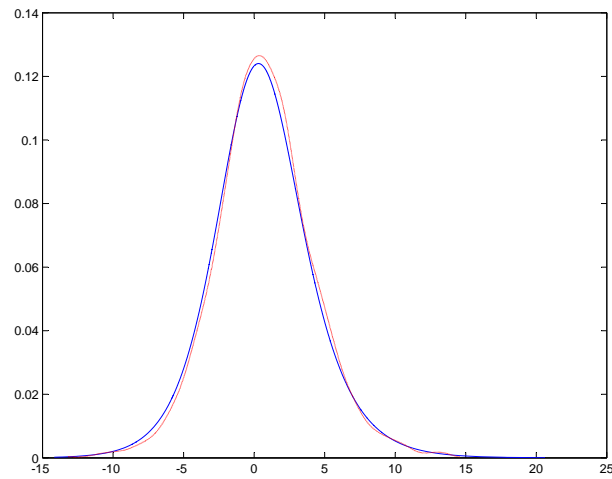


Figura 2.8: ($\theta = 2$, $n = 1$, $t = 0.5585$)

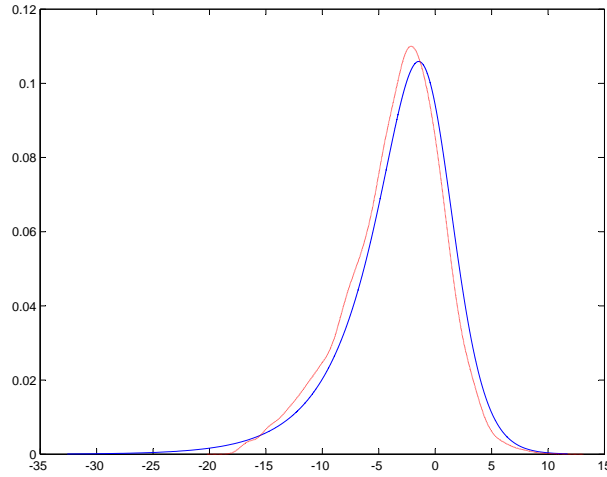


Figura 2.9: ($\theta = -3$, $n = 1$, $t = 0.2519$)

2.5. T discreta y θ continua

Siguiendo la idea de la significación o *p-value*, ahora tenemos dos elecciones posibles para una distribución fiducial, una utilizando la función de distribución $G(t; \theta)$, y otra utilizando la función continua por la izquierda $G^-(t; \theta) = P_\theta(T < t)$. Si para G y G^- se satisface la condición de contrastabilidad monótona, entonces por medio de la combinación convexa

$$H_\gamma(\theta; t) = \gamma(1 - G(t; \theta)) + (1 - \gamma)(1 - G^-(t; \theta)), \quad \gamma \in [0, 1],$$

se puede construir una “familia” de distribuciones fiduciales, las cuales pueden o no tener masa fiducial en los puntos extremos de $(\underline{\theta}, \bar{\theta})$, dependiendo de si G o G^- no satisfacen o sí, la condición de contrastabilidad extrema.

En [22] se discuten las siguientes distribuciones.

Ejemplo 2.12 *Distribución binomial.* Sea $T \sim B(n, \theta)$, donde n es conocido y $\theta \in (0, 1)$. La función de distribución de T es

$$G(t; \theta) = \sum_{j=0}^t \binom{n}{j} \theta^j (1 - \theta)^{n-j}.$$

Si $t = 0, 1, \dots, n - 1$, veamos que

$$\frac{d}{d\theta}(G(t; \theta)) = - \left(\frac{n!}{(n-t-1)!t!} \right) \theta^t (1 - \theta)^{n-t-1}. \quad (2.5)$$

Se comprueba directamente que (2.5) se cumple para $t = 0$. Supongamos entonces que $t \geq 1$.
Puesto que

$$\begin{aligned} G(t; \theta) &= \sum_{j=0}^t \binom{n}{j} \theta^j (1-\theta)^{n-j} = \binom{n}{0} (1-\theta)^n + \sum_{j=1}^t \binom{n}{j} \theta^j (1-\theta)^{n-j} \\ &= (1-\theta)^n + \sum_{j=1}^t \binom{n}{j} \theta^j (1-\theta)^{n-j}, \end{aligned}$$

entonces,

$$\begin{aligned} \frac{dG}{d\theta} &= -n(1-\theta)^{n-1} + \sum_{j=1}^t \left[\binom{n}{j} j \theta^{j-1} (1-\theta)^{n-j} - \binom{n}{j} (n-j) \theta^j (1-\theta)^{n-j-1} \right] \\ &= -n(1-\theta)^{n-1} + \sum_{j=1}^t \binom{n}{j} j \theta^{j-1} (1-\theta)^{n-j} - \sum_{j=1}^t \binom{n}{j} (n-j) \theta^j (1-\theta)^{n-j-1} \\ &= -n(1-\theta)^{n-1} + (1-\theta) \sum_{j=1}^t \binom{n}{j} j \theta^{j-1} (1-\theta)^{n-j-1} \\ &\quad - n\theta \sum_{j=1}^t \binom{n}{j} \theta^{j-1} (1-\theta)^{n-j-1} + \theta \sum_{j=1}^t \binom{n}{j} j \theta^{j-1} (1-\theta)^{n-j-1} \\ &= -n(1-\theta)^{n-1} + \sum_{j=1}^t \binom{n}{j} j \theta^{j-1} (1-\theta)^{n-j-1} - n \sum_{j=1}^t \binom{n}{j} \theta^j (1-\theta)^{n-j} \\ &= \sum_{j=1}^t \binom{n}{j} j \theta^{j-1} (1-\theta)^{n-j-1} - n \sum_{j=0}^t \binom{n}{j} \theta^j (1-\theta)^{n-j}. \end{aligned}$$

Si $t = 1$,

$$\begin{aligned}
\frac{dG}{d\theta} &= \binom{n}{1} (1-\theta)^{n-2} - n \left[\binom{n}{0} (1-\theta)^{n-1} + \binom{n}{1} \theta (1-\theta)^{n-2} \right] \\
&= n(1-\theta)^{n-2} - n(1-\theta)^{n-1} - n^2\theta(1-\theta)^{n-2} \\
&= n(1-\theta)^{n-2} (1 - (1-\theta) - n\theta) \\
&= n(1-\theta)^{n-2} (\theta - n\theta) \\
&= -n(n-1)\theta(1-\theta)^{n-2} \\
&= -\frac{n!}{(n-2)!1!} \theta (1-\theta)^{n-2} \\
&= -\frac{n!}{(n-1-1)!1!} \theta (1-\theta)^{n-1-1}.
\end{aligned}$$

Supongamos ahora que (2.5) se cumple para $t-1$ y comprobemos que entonces también se cumple para t , $2 \leq t \leq n-1$.

$$\begin{aligned}
\frac{dG}{d\theta} &= \sum_{j=1}^{t-1} \binom{n}{j} j\theta^{j-1}(1-\theta)^{n-j-1} - n \sum_{j=0}^{t-1} \binom{n}{j} \theta^j (1-\theta)^{n-j-1} \\
&\quad + \binom{n}{t} t\theta^{t-1}(1-\theta)^{n-t-1} - n \binom{n}{t} \theta^t (1-\theta)^{n-t-1} \\
&= -\frac{n!}{(n-(t-1)-1)!(t-1)!} \theta^{t-1} (1-\theta)^{n-(t-1)-1} \\
&\quad + \frac{n!t}{(n-t)!t!} \theta^{t-1} (1-\theta)^{n-t-1} - \frac{n(n!)}{(n-t)!t!} \theta^t (1-\theta)^{n-t-1} \\
&= -\frac{n!}{(n-t)!(t-1)!} \theta^{t-1} (1-\theta)^{n-t} + \frac{n!}{(n-t)!(t-1)!} \theta^{t-1} (1-\theta)^{n-t-1} \\
&\quad - \frac{n(n!)}{(n-t)!t!} \theta^t (1-\theta)^{n-t-1} \\
&= -\frac{n!}{(n-t)!(t-1)!} \theta^{t-1} (1-\theta)^{n-t-1} \left(1 - \theta - 1 + \frac{n\theta}{t} \right) \\
&= -\frac{n!}{(n-t)!(t-1)!} \theta^{t-1} (1-\theta)^{n-t-1} \left(\frac{n\theta - t\theta}{t} \right) \\
&= -\frac{n!}{(n-t-1)!t!} \theta^t (1-\theta)^{n-t-1}.
\end{aligned}$$

Así, hemos comprobado que (2.5) se cumple. Puesto que

$$\frac{d}{d\theta} (G(t; \theta)) \leq 0, \quad \theta \in (0, 1), \quad t \in \{0, 1, \dots, n-1\},$$

se tiene que G es monótonamente decreciente como función de θ . Si $t = n$, $G(n, \theta) = 1$, para todo $\theta \in (0, 1)$. Análogamente, si $t = 1, 2, \dots, n$, $G^-(t; \theta)$ es monótonamente decreciente como función de θ . Si $t = 0$, $G^-(0, \theta) = 0$ para todo $\theta \in (0, 1)$. Si se quiere hacer inferencias que sean invariantes al intercambiar θ por $\eta = 1 - \theta$, y T por $V = n - T$, una elección que satisface este requisito y que además siempre permite definir una distribución fiducial única, sin masa en los extremos, es

$$H(\theta; t) = \begin{cases} 1 - G(t; \theta), & \text{si } t = 0, \\ \frac{1}{2}(1 - G(t; \theta)) + \frac{1}{2}(1 - G^-(t; \theta)), & \text{si } t = 1, 2, \dots, n-1, \\ 1 - G^-(t; \theta), & \text{si } t = n, \end{cases}$$

cuya función de densidad es continua, la cual es

$$h(\theta; t) = \begin{cases} \beta(t+1, n-t), & \text{si } t = 0, \\ \frac{1}{2}\beta(t+1, n-t) + \frac{1}{2}\beta(t, n-t+1), & \text{si } t = 1, 2, \dots, n-1, \\ \beta(t, n-t+1), & \text{si } t = n, \end{cases}$$

donde $\beta(\alpha_1, \alpha_2)$ es la función de densidad beta con parámetros α_1 y α_2 positivos. Esta distribución es muy parecida a la distribución final $\beta(t + \frac{1}{2}, n - t + \frac{1}{2})$, que se obtiene usando la distribución inicial de Jeffreys $\pi(\theta) = \theta^{-\frac{1}{2}}(1 - \theta)^{-\frac{1}{2}}$; también es muy parecida a las distribuciones $\beta(t+1, n-t)$ y $\beta(t, n-t+1)$, las cuales han sido propuestas para encontrar los puntos extremos superior e inferior, respectivamente, de los intervalos de Clopper-Pearson (ver [5]), referidos también por algunos autores como el “procedimiento exacto” (ver [4]), en el cual en las colas izquierda y derecha se deja un área igual a $\alpha/2$ cuando se encuentra el intervalo de confianza $1 - \alpha$. En Wang [29] se da un “intervalo fiducial más corto” para θ ; esencialmente se usa el método de Clopper-Pearson, pero con una variación tal que al encontrar los extremos del intervalo, α_1 (el área de la cola izquierda) y α_2 (el área de la cola derecha) sumen α , y que dicha combinación produzca el intervalo de longitud mínima. Para $n = 10$ y $t = 7$ reportan un intervalo de confianza .9, el cual es (0.426, 0.935); también exhiben el intervalo de Clopper-Pearson (0.393, 0.913).

Ejemplo 2.13 *Distribución binomial negativa.* Sea $U \sim BN(r, p)$, donde $r \geq 1$ es conocido y $p \in (0, 1)$. La función de distribución es

$$G(u; p) = p^r \sum_{j=0}^u \binom{j+r-1}{r-1} (1-p)^j, \quad u \in \{0, 1, 2, \dots\}.$$

Procediendo de modo similar que en el ejemplo 2.12, se comprueba que en este caso la derivada es

$$\frac{d}{dp} (G(u; p)) = p^{r-1} \left(\frac{r(r+1)\dots(r+u)}{u!} \right) (1-p)^u$$

la cual es no negativa, de donde se tiene que $G(u; p)$ es monótonamente creciente como función de p . Por lo tanto la distribución fiducial de p es $H(p; u) = G(u; p)$, que corresponde a una distribución $\beta(r, u+1)$. Si hacemos $r = t$ y $u = n - t$, entonces $\beta(r, u+1) = \beta(t, n - t + 1)$.

Si intercambiamos p por $\eta = (1 - p)$ y definimos la variable aleatoria $W \sim BN(r^*, \eta)$, donde $r^* \geq 1$ es conocido y $\eta \in (0, 1)$, entonces, usando la función de distribución continua por la derecha,

$$G_*(w; \eta) = \eta^{r^*} \sum_{j=0}^w \binom{j+r^*-1}{r^*-1} (1-\eta)^j, \quad w \in \{0, 1, 2, \dots\},$$

la cual es monótonamente creciente como función de η , resulta que la distribución fiducial de η es $H_*(\eta; w) = G_*(w; \eta)$, que corresponde a una distribución $\beta(r^*, w+1)$. Si hacemos $w = t$ y $r^* = n - t$, entonces $\beta(r^*, w+1) = \beta(n - t, t + 1)$, de donde se sigue que tenemos una distribución $\beta(t+1, n - t)$ para p . Así, las dos distribuciones fiduciales son diferentes en cuanto a invarianza de lo que podríamos llamar "éxito". Observamos que en el primer experimento binomial negativo donde $U \sim BN(r, p)$, como función de p la función de distribución continua por la izquierda $G^-(u, p)$ es monótonamente creciente para todo u excepto para $u = 0$, por lo que la función de distribución fiducial propuesta sería la función

$$H(p; u) = \begin{cases} G(u; p), & \text{si } u = 0, \\ \frac{1}{2}G(u; p) + \frac{1}{2}G^-(u; p), & \text{si } u > 0, \end{cases}$$

la cual tiene densidad

$$h(p; u) = \begin{cases} \beta(r, u+1), & \text{si } u = 0, \\ \frac{1}{2}\beta(r, u+1) + \frac{1}{2}\beta(r+1, u), & \text{si } u > 0. \end{cases}$$

Siguiendo la misma construcción del segundo experimento binomial negativo, la densidad fiducial que se obtiene es, ahora en términos de p en lugar de η ,

$$h_*(p; w) = \begin{cases} \beta(w+1, r^*), & \text{si } w = 0, \\ \frac{1}{2}\beta(w+1, r^*) + \frac{1}{2}\beta(w, r^*+1), & \text{si } w > 0. \end{cases}$$

Así, con resultados de verosimilitud equivalentes, $w = r$ y $u = r^*$, la densidad fiducial que es equivalente para $t = w$ y $n - t = u$ es

$$h(p; t) = \begin{cases} \beta(t+1, n-t), & \text{si } t = 0, \\ \frac{1}{2}\beta(t, n-t+1) + \frac{1}{2}\beta(t+1, n-t), & \text{si } t = 1, 2, \dots, n-1, \\ \beta(t, n-t+1), & \text{si } t = n, \end{cases}$$

la cual es la misma que la densidad fiducial invariante inducida en el caso binomial. Así, si se pudiera pensar en un resultado binomial y en un resultado binomial negativo equivalente, esta densidad fiducial produce las mismas inferencias. Por lo tanto, la distribución fiducial propuesta para p satisface el Principio Fuerte de Verosimilitud.

Ejemplo 2.14 *Distribución de Poisson.* Sea $T \sim P(\lambda)$, es decir,

$$G(t; \lambda) = \sum_{j=0}^t \frac{e^{-\lambda} \lambda^j}{j!}, \quad t = 0, 1, 2, \dots, \lambda > 0.$$

En este caso

$$\frac{d}{d\lambda} G(t; \lambda) = -\frac{e^{-\lambda} \lambda^t}{t!},$$

luego $G(t; \lambda)$ es decreciente como función de λ ; más aún, para todo $t \geq 0$, $G(t; \lambda) \rightarrow 1$ cuando $\lambda \rightarrow 0$, y $G(t; \lambda) \rightarrow 0$ cuando $\lambda \rightarrow \infty$. En este caso G satisface las propiedades de contrastabilidad monótona y extrema; la distribución fiducial inducida es la distribución gamma $(t+1, 1)$, la cual es la que se usa para obtener el extremo superior del intervalo de confianza obtenido por Garwood (1936), como se menciona en Casella y Berger [5]. El procedimiento es similar al de Clopper-Pearson para la distribución binomial. En contraste con la distribución binomial, en este caso el uso de G^- produce una distribución constante cuando $t = 0$. Si entendemos a la distribución de Poisson como una aproximación a la binomial cuando $n \rightarrow \infty$ y $np \rightarrow \lambda$, entonces haciendo el paso al límite, la distribución fiducial obtenida en el caso Poisson corresponde a la combinación convexa, con ponderaciones iguales a $1/2$, de las correspondientes gamma $(t+1)$ y gamma (t) , con uso exclusivo de gamma $(t+1)$, si $t=0$.

2.5.1. Análisis fiducial por medio de distribuciones de series de potencias

Definición 2.15 Se dice (ver [18]) que una distribución de probabilidad discreta pertenece a la familia de distribuciones de serie de potencias si su función de probabilidad puede ser escrita en la forma

$$f(x; \eta) = \frac{a_x \eta^x}{b(\eta)}, \quad x = 0, 1, 2, \dots, \quad (2.6)$$

donde $\eta > 0$, $a_x \geq 0$ y $b(\eta) = \sum_{x=0}^{\infty} a_x \eta^x$.

En (2.6) η se llama parámetro potencia y $b(\cdot)$ función de serie.

Las distribuciones discretas consideradas son miembros de la familia de series de potencias, como lo vemos enseguida.

a) Distribución binomial $B(n, \theta)$. En este caso

$$\eta = \frac{\theta}{1 - \theta}, \quad a_x = \binom{n}{x} \quad \text{si } x = 0, 1, \dots, n, \quad \text{y } a_x = 0 \quad \text{si } x = n + 1, \dots$$

b) Distribución binomial negativa $BN(r, p)$. Aquí

$$\eta = 1 - p \quad \text{y} \quad a_x = \binom{r + x - 1}{x}, \quad \text{si } x = 0, 1, \dots$$

c) Distribución de Poisson $P(\lambda)$. En este caso

$$\eta = \lambda \quad \text{y} \quad a_x = \frac{1}{x!} \quad \text{si } x = 0, 1, \dots$$

Si $f(x; \eta)$ pertenece a la familia de serie de potencias, entonces el estadístico suficiente minimal de η es $T = \sum_{i=1}^n X_i$, el cual tiene distribución en la misma familia (ver [18]).

Dado que las distribuciones discretas consideradas pertenecen a la familia de serie de potencias, debe ser posible deducir tanto la distribución fiducial como las mismas conclusiones obtenidas, a partir de su representación como elementos de esta familia. Así, si el rango de valores de X es acotado, digamos $\{0, 1, \dots, K\}$, la inferencia que se obtiene para η usando T , (digamos $T = t$) debe ser la misma que la inferencia obtenida para η usando el estadístico $V = nK - T$, el cual tiene una distribución de serie de potencias con parámetro η^{-1} . Si la función de probabilidad de T es

$$g(t; \eta) = \frac{A_t \eta^t}{B(\eta)},$$

entonces $V = nK - T$ tiene función de probabilidad

$$g^*(v; \eta^*) = \frac{A_v^* (\eta^*)^v}{B^*(\eta^*)},$$

con $\eta^* = \frac{1}{\eta}$. La distribución fiducial de η obtenida de la función de distribución de T (en t) debe estar relacionada con la distribución fiducial de η^* obtenida de la función de distribución de V (en v), produciendo en ambos casos la misma distribución fiducial de η . Es decir, se debe satisfacer la propiedad de invarianza en la distribución fiducial de η si usamos T o V .

Sin embargo, el análisis fiducial utilizando este enfoque es material de investigaciones futuras.

Capítulo 3

Generadores de la distribución fiducial

Con la disposición cada vez mayor de poder de cómputo, el paradigma dentro de la investigación se ha ido modificando en el sentido de que se depende cada vez más de soluciones numéricas en los llamados procedimientos de cómputo intensivo.

En el problema de identificación de una distribución fiducial se puede recurrir a procedimientos numéricos que resultan equivalentes, desde el punto de vista práctico. Esto ha venido ocurriendo por ejemplo cuando en la estadística bayesiana no se puede dar una forma cerrada a una distribución final, pero sí se puede obtener un método que permita simular de esa distribución final, como se ilustra más adelante en este capítulo en el caso de la distribución $gamma(\alpha, \beta)$, cuando la distribución inicial es la no informativa de Jeffreys.

En este capítulo, para obtener la distribución fiducial se propone el uso del generador que sin proponérselo obtuvieron Engen y Lillegård [8]. Cuando se aplica a la distribución $gamma(\alpha, \beta)$ se observa que se reproduce lo señalado en [27], en 2013, sobre el estadístico $W = \tilde{X}/\bar{X}$, donde \tilde{X} y \bar{X} son, respectivamente, las medias geométrica y aritmética. También se ve que el uso de la distribución inicial no informativa de Jeffreys produce una distribución final bayesiana mucho más parecida a la distribución fiducial, que cuando se usa como distribución inicial la de referencia.

Se muestra que el uso de la verosimilitud perfil también lleva al uso del estadístico W para obtener la distribución fiducial de α . Dicho estadístico ya había sido mencionado en [15], en 1976, incluyendo su independencia de \bar{X} y otras propiedades que también mencionan en [27], aunque sin citar a [15].

Para la distribución exponencial truncada, además de obtener la distribución fiducial con el generador (implícito) de Engen y Lillegård, también se considera el método de generarla directamente con la función de distribución del estadístico suficiente minimal. En Ambos casos las distribuciones fiduciales obtenidas se comparan con la distribución fiducial exacta (2.4).

Adicionalmente, siguiendo las ideas desarrolladas en las secciones 2.1 y 2.3, en la obser-

vación 3.7 se propone un método alternativo para generar la distribución fiducial, aunque su aplicabilidad es menos general que la del generador de Engen y Lillegård. Ambos procedimientos son comparados en dos ejemplos, en los que se ve que las distribuciones fiduciales simuladas correspondientes prácticamente coinciden.

3.1. El generador inducido por la función de distribución

Sea $\underline{X} = (X_1, X_2, \dots, X_n)$ un vector de variables aleatorias independientes de la distribución de probabilidad $F(\cdot; \theta)$, donde θ es el parámetro simple desconocido de la distribución, y sea $T(\underline{X})$ un estadístico suficiente minimal de θ . Sea $G(t; \theta)$ la función de distribución de T . Si $G(t; \theta)$ satisface las propiedad de contrastabilidad monótona, entonces $1 - G(t; \theta)$, o $G(t; \theta)$, es la función de distribución fiducial de θ , luego de manera natural $G(t; \theta)$ induce un generador de la distribución fiducial de θ , de la siguiente manera: puesto que $G(T; \theta) \sim U(0, 1)$, entonces para un valor observado t de $T(\underline{X})$ generamos una muestra de observaciones independientes u_1, u_2, \dots, u_m de la distribución $U(0, 1)$. Para cada $i = 1, \dots, m$, sea $\tilde{\theta}_i$ la solución única de la ecuación

$$G(t; \theta) = u_i.$$

Por lo tanto $\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_m$ constituyen una muestra de observaciones independientes de la distribución fiducial de θ . Con esta muestra podemos estimar dicha distribución fiducial.

Ejemplo 3.1 *Sea la función de distribución exponencial truncada*

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

Del ejemplo (2.11), sabemos que la estadística suficiente minimal de θ es $T = \sum_{i=1}^n X_i$, donde X_1, \dots, X_n es una muestra aleatoria de $F(x; \theta)$; también, la función de densidad de T es

$$g(t; \theta) = \left(\frac{\theta}{e^\theta - 1} \right)^n I_n(t) e^{\theta t}, \quad t \in (0, n),$$

donde

$$I_n(t) = \frac{1}{(n-1)!} \sum_{j=0}^n \binom{n}{j} (-1)^j (t-j)_+^{n-1},$$

(ver [23]); asimismo, $G(t; \theta)$ satisface las propiedades de contrastabilidad monótona y extrema. Por lo tanto, para un valor dado t de T , podemos resolver numéricamente la ecuación

$G(t; \theta) = u$, es decir, la ecuación

$$\int_0^t \left(\frac{\theta}{e^\theta - 1} \right)^n I_n(s) e^{\theta s} ds = u,$$

donde u es un valor observado de la distribución $U(0, 1)$. Para cada combinación de (θ, n, t) del ejemplo (2.11), se obtuvieron 10000 observaciones de θ con las que se estimaron numéricamente las densidades fiduciales de θ en cada caso. Sus gráficas son las que están en color rojo en las figuras 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7 y 3.8. En cada una de las figuras, la gráfica en color azul es la gráfica de la densidad fiducial, vista en el mismo ejemplo (2.11). Se observa que se replica el parecido ya visto en O'Reilly y Rueda [23]. De las réplicas se puede observar que el parecido es enorme, incluso para el tamaño de muestra $n = 1$.

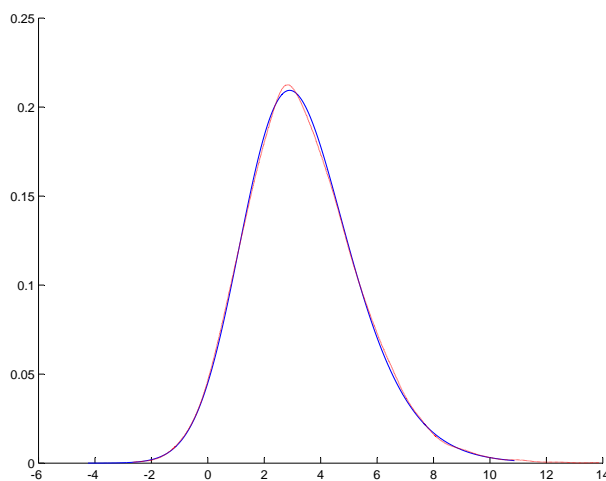


Figura 3.1: ($\theta = 1$, $n = 5$, $t = 3.6906$)

3.2. El generador de Engen y Lillegård

Sea $\underline{X} = (X_1, X_2, \dots, X_n)$ un vector de variables aleatorias independientes de la distribución de probabilidad $F(\cdot; \theta)$, donde θ es el parámetro desconocido de la distribución, y sea $T(\underline{X})$ un estadístico suficiente para θ . Sea U una variable aleatoria de una distribución conocida, tal que para θ dado, X , con distribución $F(\cdot; \theta)$, puede ser simulada por medio de U . Más precisamente, supongamos que existe una función Borel medible h tal que $h(U, \theta)$ tiene distribución $F(\cdot; \theta)$. Si U_1, U_2, \dots, U_n son variables aleatorias independientes con la misma distribución que U , denotemos por $\chi(\underline{U}; \theta)$ a $(h(U_1, \theta), h(U_2, \theta), \dots, h(U_n, \theta))$, el vector de variables aleatorias independientes generadas por este método, y por $T(\underline{U}; \theta)$ a $T(\chi(\underline{U}; \theta))$, el cual tiene la misma distribución que $T(\underline{X})$.

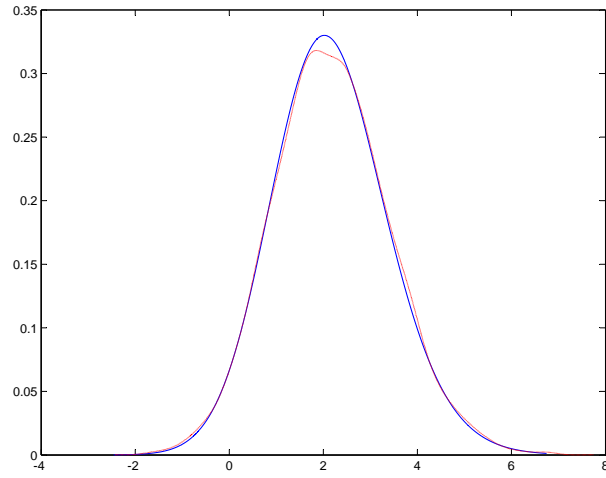


Figura 3.2: ($\theta = 2$, $n = 10$, $t = 6.6743$)

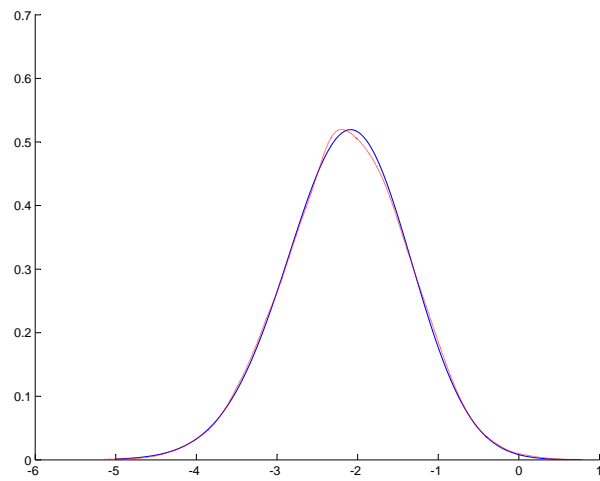


Figura 3.3: ($\theta = -2$, $n = 25$, $t = 8.3469$)

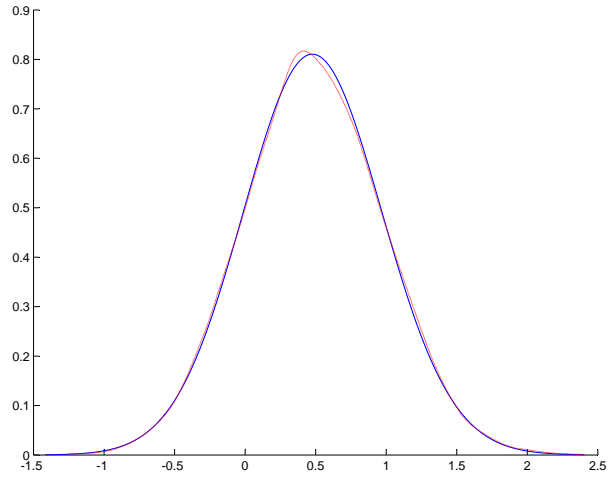


Figura 3.4: ($\theta = 0$, $n = 50$, $t = 27.0027$)

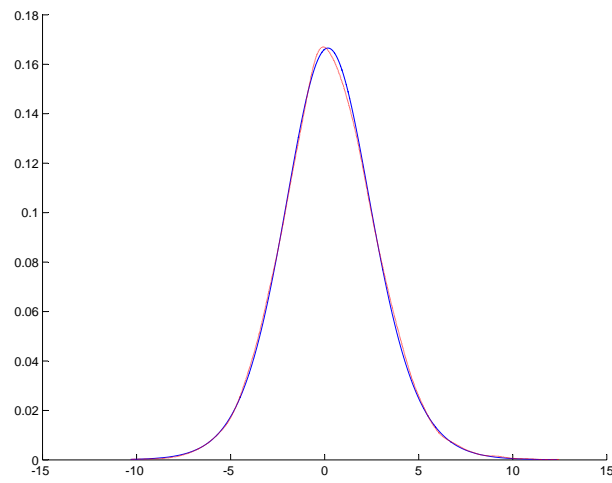


Figura 3.5: ($\theta = 4$, $n = 2$, $t = 1.0360$)

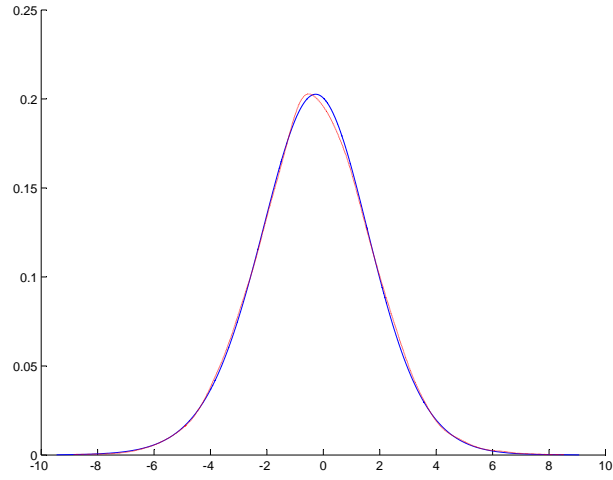


Figura 3.6: ($\theta = 0, n = 3, t = 1.4168$)

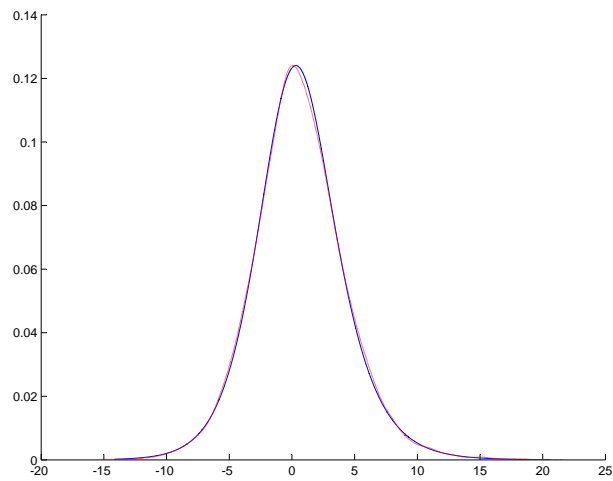


Figura 3.7: ($\theta = 2, n = 1, t = 0.5585$)

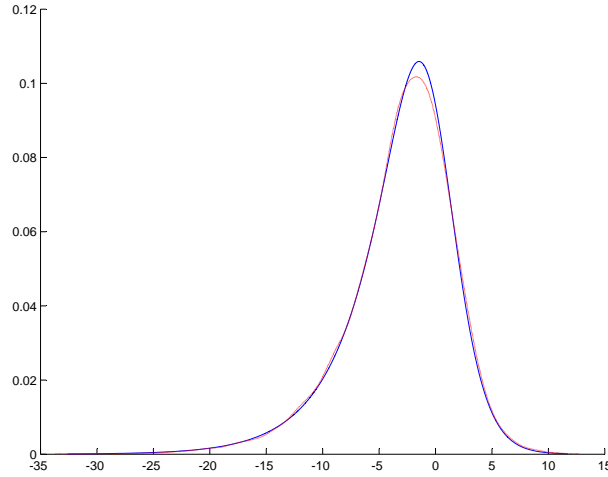


Figura 3.8: ($\theta = -3$, $n = 1$, $t = 0.2519$)

A pesar de que el modelo $N(\mu, \sigma^2)$ es de grupo y por ello existe una cantidad pivotal genérica, denotada por $\theta^{-1}T$, a partir de la cual se induce la distribución fiducial para θ , habiendo observado t , para entender cabalmente la idea del generador (implícito) de la distribución fiducial, utilizado en Engen y Lillegård [8], es de interés revisar de nuevo este caso.

Ejemplo 3.2 *Distribución normal.* Sea $\underline{X} = (X_1, X_2, \dots, X_n)$ un vector de variables aleatorias independientes con distribución $N(\mu, \sigma^2)$. Si $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ y $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, entonces $T(X_1, X_2, \dots, X_n) = (\bar{X}, S)$ es una estadística suficiente para $\theta = (\mu, \sigma)$. Si $\underline{U} = (U_1, U_2, \dots, U_n)$ es un vector de variables aleatorias independientes con distribución $N(0, 1)$, entonces las variables aleatorias $h(U_i, \theta) = \mu + \sigma U_i$ son independientes y tienen distribución $N(\mu, \sigma^2)$, $i = 1, \dots, n$. De aquí

$$\chi(\underline{U}; \theta) = (\mu + \sigma U_1, \mu + \sigma U_2, \dots, \mu + \sigma U_n),$$

y

$$T(\underline{U}, \theta) = (\mu + \sigma \bar{U}, \sigma S_U), \quad (3.1)$$

donde \bar{U} y S_U están definidos análogamente como \bar{X} y S .

Ahora supongamos que se observa un valor de T , o sea, se observa que $T = t$. Para \underline{U} fijo, la ecuación

$$T(\underline{U}; \theta) = t \quad (3.2)$$

puede o no tener solución única para θ . Si tiene solución única, entonces generando observaciones $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_m$ de \underline{U} , obtenemos valores $\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_m$ que satisfacen la ecuación (3.2).

Estos valores de θ así obtenidos son una muestra de una cierta distribución del parámetro, habiendo observado t . Esta distribución así inducida es la distribución fiducial.

Ejemplo 3.3 *Distribución normal (continuación).* Si $t = (\bar{x}, s)$, de la ecuación (3.1) se tiene

$$(\mu + \sigma \bar{u}_i, \sigma s_{u_i}) = (\bar{x}, s), \quad i = 1, \dots, m,$$

de donde

$$(\mu_i, \sigma_i) = \left(\bar{x} - \frac{\bar{u}_i}{s_{u_i}} s, \frac{1}{s_{u_i}} s \right), \quad i = 1, \dots, m. \quad (3.3)$$

De aquí se tiene un generador de la distribución fiducial de μ y σ .

Observación 3.4 Si (U_1, U_2, \dots, U_n) es un vector de variables aleatorias independientes de la distribución $N(0, 1)$, entonces

$$\bar{U} \sim N\left(0, \frac{1}{n}\right),$$

y

$$(n-1)S_U^2 \sim \chi_{n-1}^2,$$

donde \bar{U} y S_U^2 son independientes. Por lo tanto, de (3.3) vemos que para simular cada par (μ_i, σ_i) basta generar, de manera independiente, una observación de la distribución $N(0, 1)$, digamos z_i , y una observación de la distribución χ_{n-1}^2 , digamos ξ_i^2 . Entonces en (3.3) sustituimos \bar{u}_i y s_{u_i} por $\frac{z_i}{\sqrt{n}}$ y $\frac{\xi_i}{\sqrt{n-1}}$, respectivamente, obteniendo

$$\begin{aligned} (\mu_i, \sigma_i) &= \left(\bar{x} - \frac{\frac{z_i}{\sqrt{n}}}{\frac{\xi_i}{\sqrt{n-1}}} s, \frac{1}{\frac{\xi_i}{\sqrt{n-1}}} s \right) \\ &= \left(\bar{x} - \frac{z_i}{\frac{\xi_i}{\sqrt{n-1}}} \frac{s}{\sqrt{n}}, \frac{\sqrt{n-1}}{\xi_i} s \right). \end{aligned} \quad (3.4)$$

Así, para generar m pares de (μ, σ) es suficiente generar m pares de observaciones independientes, cada par formado por una observación de la distribución $N(0, 1)$ y una de la distribución χ_{n-1}^2 . Este procedimiento es mucho más corto que el obtenido al aplicar directamente el generador de Engen y Lillegård, ya que en este último para generar cada par de (μ, σ) primero se simulan n observaciones independientes de la distribución normal estándar.

Observación 3.5 Para el caso de la distribución normal, una vez que vimos que en (3.3) \bar{u}_i y s_{u_i} pueden simularse directamente, entonces de (3.4), (1.29) y (1.30) observamos que

a partir de el generador de Engen y Lillegård reconstruimos el generador de la distribución fiducial que obtuvimos en la sección 1.3, lo que ya sabíamos porque la distribución normal es un modelo de grupo.

3.3. Distribución Gamma (α, β)

Un caso en el que la familia no es de grupo, pero en el que podemos utilizar esta idea para generar muestras de la distribución fiducial, es el de la distribución Gamma (α, β) . Aunque el procedimiento de Engen y Lillegård que propusieron en [8] no se diseñó para simular valores de los parámetros, sino para simular muestras condicionales, discutieron allí el caso Gamma (α, β) y obtuvieron el algoritmo correspondiente. La siguiente proposición está basada en los resultados que derivaron para este caso.

Proposición 3.6 Sean x_1, x_2, \dots, x_n observaciones de la muestra aleatoria X_1, X_2, \dots, X_n de la distribución Gamma (α, β) . Entonces existe el generador de la densidad fiducial conjunta de α y β .

Demostración. La función de densidad conjunta de (X_1, X_2, \dots, X_n) es

$$f(x_1, x_2, \dots, x_n; \alpha, \beta) = \frac{1}{[\Gamma(\alpha)]^n \beta^{n\alpha}} (\prod_{i=1}^n x_i)^{\alpha-1} e^{-\frac{1}{\beta} \sum_{i=1}^n x_i}.$$

El estadístico

$$T(\underline{X}) = \left(\prod_{i=1}^n X_i, \sum_{i=1}^n X_i \right)$$

es suficiente minimal para (α, β) . Ahora, si

$$Y = \frac{X}{\beta},$$

entonces la función de densidad de Y es

$$f_Y(y; \alpha) = \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y}.$$

Si

$$F_Y(Y; \alpha) = U,$$

donde U tiene distribución uniforme en $(0, 1)$ y F_Y es la función de distribución de Y , se tiene

$$X = \beta Y = \beta F_Y^{-1}(U, \alpha).$$

De la ecuación anterior se tiene

$$T(\underline{U}; \alpha, \beta) = \left(\beta^n \prod_{i=1}^n F_Y^{-1}(U_i, \alpha), \beta \sum_{i=1}^n F_Y^{-1}(U_i, \alpha) \right),$$

donde (U_1, U_2, \dots, U_n) es un vector de variables aleatorias independientes con distribución uniforme en $(0, 1)$. Si

$$t = \left(\prod_{i=1}^n x_i, \sum_{i=1}^n x_i \right) = (t_1, t_2)$$

es un valor observado del estadístico suficiente, entonces, de las ecuaciones

$$\begin{aligned} \beta^n \prod_{i=1}^n F_Y^{-1}(U_i, \alpha) &= t_1 \\ \beta \sum_{i=1}^n F_Y^{-1}(U_i, \alpha) &= t_2, \end{aligned}$$

resulta

$$\frac{\sum_{i=1}^n F_Y^{-1}(U_i, \alpha)}{\left(\prod_{i=1}^n F_Y^{-1}(U_i, \alpha) \right)^{\frac{1}{n}}} = \frac{t_2}{t_1^{1/n}}. \quad (3.5)$$

Si (u_1, u_2, \dots, u_n) es una observación de (U_1, U_2, \dots, U_n) , entonces

$$\frac{\sum_{i=1}^n F_Y^{-1}(u_i, \alpha)}{\left(\prod_{i=1}^n F_Y^{-1}(u_i, \alpha) \right)^{\frac{1}{n}}}$$

es decreciente como función de α , y la ecuación

$$\frac{\sum_{i=1}^n F_Y^{-1}(u_i, \alpha)}{\left(\prod_{i=1}^n F_Y^{-1}(u_i, \alpha) \right)^{\frac{1}{n}}} = \frac{t_2}{t_1^{1/n}}$$

tiene solución única. Aunque no la podemos resolver analíticamente, podemos aproximar numéricamente la solución. El valor de β lo encontramos, por ejemplo, por medio de la ecuación

$$\beta \sum_{i=1}^n F_Y^{-1}(u_i, \alpha) = t_2.$$

De esta forma podemos generar tantos valores de (α, β) como queramos, con los cuales podemos encontrar una estimación de la densidad fiducial conjunta de α y β , así como de las densidades marginales correspondientes. ■

En el artículo *Fiducial theory and optimal inference* [27], publicado en 2013 y cuya lectura fue hecha después de escribir la proposición 3.6, Taraldsen y Lindqvist dan el modelo fiducial para $\theta = (\alpha, \beta)$,

$$\chi_1(u, \theta) = \beta \overline{F^{-1}(u, \alpha)}, \quad \chi_2(u, \theta) = \widetilde{F^{-1}(u, \alpha)} / \overline{F^{-1}(u, \alpha)},$$

donde \overline{X} y \widetilde{X} denotan las medias aritmética y geométrica, respectivamente, a partir del cual se reconstruye la proposición 3.6.

También dan un modelo fiducial alternativo para $\theta = (\alpha, \beta)$ basados en los hechos de que la distribución del estadístico $W = \widetilde{X}/\overline{X}$, el cual sólo depende de α (lo que también puede ser visto de la ecuación (3.5)), es conocida, y que W y \overline{X} son independientes. Estos hechos fueron establecidos en [15], en 1976. Entonces proponen invertir la función de distribución de W , y como una alternativa a $\overline{F^{-1}(u, \alpha)}$ proponen invertir la función de distribución gamma con parámetro de forma $n\alpha$ y escala $1/n$. Esto último se debe a que si n variables aleatorias son independientes con distribución $gamma(\alpha, 1)$, entonces su media aritmética, la cual no es una estadística suficiente de α , tiene distribución $gamma(n\alpha, 1/n)$. El modelo fiducial alternativo tiene el inconveniente adicional de que en general la distribución de W es complicada de manipular, incluso numéricamente, como puede verse en [15].

Observación 3.7 *A partir de la verosimilitud perfil de α también es posible darse cuenta de que el estadístico W sólo depende de α , como se muestra en seguida. Sea $\underline{x} = (x_1, x_2, \dots, x_n)$ una observación de la muestra aleatoria $\underline{X} = (X_1, X_2, \dots, X_n)$. Puesto que la función de verosimilitud de α y β es*

$$f(x_1, x_2, \dots, x_n; \alpha, \beta) = \frac{1}{[\Gamma(\alpha)]^n \beta^{n\alpha}} (\prod_{i=1}^n x_i)^{\alpha-1} e^{-\frac{1}{\beta} \sum_{i=1}^n x_i},$$

entonces la verosimilitud perfil de α es

$$L_p(\alpha; \underline{x}) = \frac{\alpha^{n\alpha}}{(\Gamma(\alpha))^n (\overline{x})^{n\alpha}} \left(\prod_{i=1}^n x_i \right)^{\alpha-1} e^{-n\alpha} = \frac{\alpha^{n\alpha} e^{-n\alpha}}{(\Gamma(\alpha))^n} \left(\frac{(\prod_{i=1}^n x_i)^{\frac{1}{n}}}{\overline{x}} \right)^{n\alpha} \prod_{i=1}^n x_i,$$

de donde tenemos que

$$T = \frac{\left(\prod_{i=1}^n X_i \right)^{\frac{1}{n}}}{\overline{X}}$$

es un estadístico suficiente minimal de α , es decir, T es sólo función de α . Así, $W = 1/T$ sólo depende de α . De acuerdo a lo expuesto en la secciones 2.1 y 2.3, la función de distribución fiducial de α está dada por

$$H(\alpha; t) = 1 - P_\alpha(T \leq t),$$

de donde

$$H(\alpha; t) = P_\alpha(T \geq t).$$

De aquí obtenemos un método alternativo para estimar la distribución fiducial de α . Si t_0 es un valor de T , la función de distribución fiducial de α , $H(\alpha; t_0)$, se estima como sigue:

1. Se generan m observaciones, $(x_{11}, \dots, x_{1n}), \dots, (x_{m1}, \dots, x_{mn})$, de una muestra aleatoria de tamaño n de la distribución Gamma ($\alpha, \beta = 1$).
2. Para $j = 1, \dots, m$, sea

$$t_j = \frac{\left(\prod_{i=1}^n x_{ji} \right)^{\frac{1}{n}}}{\frac{1}{n} \sum_{i=1}^n x_{ji}}.$$

Sea $k = \text{cardinalidad}\{j : t_j > t_0\}$. Entonces

$$H(\alpha; t_0) \approx \frac{k}{n}.$$

De lo anterior se tiene el algoritmo que sigue para estimar la densidad fiducial de α :

Algoritmo 3.8

1. Dada la observación $\underline{x} = (x_1, x_2, \dots, x_n)$ de la muestra aleatoria $\underline{X} = (X_1, X_2, \dots, X_n)$, sea

$$t_0 = \frac{\left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}}{\bar{x}}.$$

2. Generar una partición $\alpha_{\text{inf}} = \alpha_0 < \alpha_1 < \dots < \alpha_r = \alpha_{\text{sup}}$, de $[\alpha_{\text{inf}}, \alpha_{\text{sup}}]$, donde α_{inf} es el alfa tal que el área bajo la densidad fiducial a la izquierda de α_{inf} es un valor predeterminado, por ejemplo .01, y α_{sup} es el alfa tal que el área bajo la densidad fiducial a la derecha de α_{sup} es un valor predeterminado, por ejemplo .01.
3. Obtener estimaciones de $H(\alpha_i; t_0)$, $i = 0, \dots, r$.
4. Obtener una estimación de la densidad fiducial $h(\alpha; t_0)$.

Ejemplo 3.9 Se simuló una observación $(x_1, x_2, \dots, x_{30})$ de la muestra aleatoria $(X_1, X_2, \dots, X_{30})$ de una distribución Gamma(3, 5). Con élla se obtuvieron 10000 observaciones $(\tilde{\alpha}_1, \tilde{\beta}_1)$, $(\tilde{\alpha}_2, \tilde{\beta}_2)$, \dots , $(\tilde{\alpha}_{10000}, \tilde{\beta}_{10000})$ por medio del generador de Engen y Lillegård, con las que se

estimaron numéricamente las densidades fiduciales de α y β . La gráfica de la densidad fiducial de α estimada de este modo es la de color azul en la figura 3.9. Con la misma muestra simulada, con el algoritmo 3.8 también se estimó la densidad fiducial de α , cuya gráfica es la de color rojo en la misma figura 3.9. Vemos que ambas gráficas prácticamente coinciden.

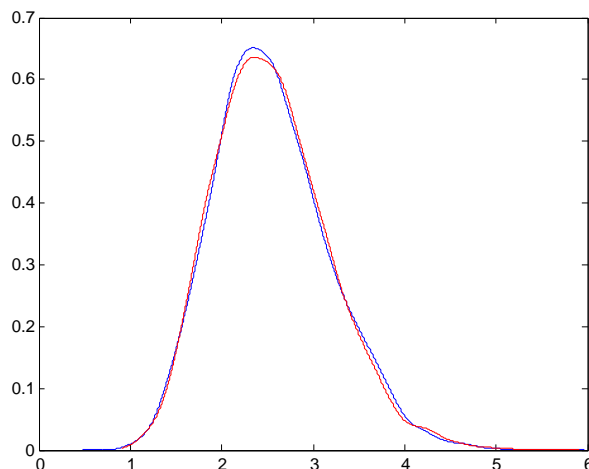


Figura 3.9: Densidades fiduciales del parámetro α de la distribución gamma.¹

Con el propósito de comparar las densidades fiduciales, con los mismos valores simulados x_1, x_2, \dots, x_{30} de la distribución $Gamma(3, 5)$, se estimarán las densidades finales bayesianas de α y β , cuando la distribución inicial es la no informativa de Jeffreys (ver [2] y [21]),

$$\pi(\alpha, \beta) \propto \frac{1}{\beta} \left(\frac{d\psi(\alpha)}{d\alpha} \right)^{\frac{1}{2}},$$

donde

$$\psi(\alpha) = \frac{d}{d\alpha} (\log \Gamma(\alpha)).$$

Por lo tanto,

$$p(\alpha, \beta | x_1, x_2, \dots, x_n) \propto \frac{1}{[\Gamma(\alpha)]^n \beta^{n\alpha}} (\prod_{i=1}^n x_i)^{\alpha-1} \exp\left(-\frac{1}{\beta} \sum_{i=1}^n x_i\right) \frac{1}{\beta} \left(\frac{d\psi(\alpha)}{d\alpha} \right)^{\frac{1}{2}}.$$

¹La gráfica en color azul es la obtenida con el generador de Engen y Lillegard; la gráfica en color rojo es la obtenida por medio del algoritmo 3.8

Una forma de encontrar $p(\alpha, \beta \mid x_1, \dots, x_{30})$ es resolviendo la integral doble

$$\int f(x_1, \dots, x_{30} \mid \alpha, \beta) \pi(\alpha, \beta) d\alpha d\beta. \quad (3.6)$$

Por último, la densidad posterior de α dados x_1, \dots, x_{30} se obtiene como

$$h(\alpha \mid x_1, \dots, x_{30}) = \int p(\alpha, \beta \mid x_1, \dots, x_{30}) d\beta. \quad (3.7)$$

Sin embargo, las integrales (3.6) y la del lado derecho de (3.7) no se pueden resolver analíticamente. En este caso lo mejor es obtener una muestra simulada de $p(\alpha, \beta \mid x_1, \dots, x_{30})$, lo cual hacemos a través de la aproximación normal asintótica y del algoritmo de Metropolis-Hastings (ver [14] y [16]). Por medio del cual se genera una cadena de Markov cuya distribución de equilibrio sea la distribución final. Para ello primero efectuamos la reparametrización

$$a = \log(\alpha), \quad b = \log(\beta),$$

luego, si $q(a, b \mid x_1, x_2, \dots, x_n)$ es la densidad final de a y b dados x_1, x_2, \dots, x_n , se tiene

$$\pi^*(a, b \mid x_1, x_2, \dots, x_n) \propto \frac{1}{[\Gamma(e^a)]^n e^{(ne^a+1)b}} (\prod_{i=1}^n x_i)^{e^a-1} \exp\left(-\frac{1}{e^b} \sum_{i=1}^n x_i + a + b\right) \left(\frac{d\psi(\alpha)}{d\alpha}\right)^{\frac{1}{2}},$$

con la que obtenemos la aproximación normal asintótica de $\pi^*(a, b \mid x_1, x_2, \dots, x_n)$, dada por

$$\tilde{g}(a, b \mid (\hat{a}, \hat{b})) = N(a, b \mid (\hat{a}, \hat{b}), V(\hat{a}, \hat{b})),$$

donde (\hat{a}, \hat{b}) y $V(\hat{a}, \hat{b})$ denotan la media y la matriz de varianza-covarianza de la distribución normal bivariada de a y b , y (\hat{a}, \hat{b}) es la moda estimada de $\pi^*(a, b \mid x_1, x_2, \dots, x_n)$. Para mayor referencia ver [16]. La probabilidad de transición de la cadena de Markov está dada por

$$\gamma((a^*, b^*), (a, b)) = \min\left\{\frac{w(a^*, b^*)}{w(a, b)}, 1\right\},$$

con $w(a, b) = \pi^*(a, b \mid \underline{x}) \tilde{g}(a, b \mid (\hat{a}, \hat{b}))$. El algoritmo es el siguiente:

Algoritmo 3.10

1. Dado un valor inicial (a^0, b^0) , la i -ésima iteración consiste en:
2. generar una observación (a^*, b^*) de $\tilde{g}(a, b \mid (\hat{a}, \hat{b}))$;

3. generar una observación u de la distribución $U(0, 1)$;
4. si $u \leq \gamma((a^*, b^*), (a^i, b^i))$, hacer $(a^{i+1}, b^{i+1}) = (a^*, b^*)$; en caso contrario, hacer $(a^{i+1}, b^{i+1}) = (a^i, b^i)$.

Ejemplo 3.11 Se generaron 10000 observaciones $(\tilde{\alpha}_1, \tilde{\beta}_1), (\tilde{\alpha}_2, \tilde{\beta}_2), \dots, (\tilde{\alpha}_{10000}, \tilde{\beta}_{10000})$ de (α, β) con las que se estimaron numéricamente las densidades finales de α y β , cuando la distribución inicial es la no informativa de Jeffreys. Las gráficas son, respectivamente, las que están en color rojo en las figuras 3.10 y 3.11. En estas figuras, las gráficas en color azul son las gráficas de las densidades fiduciales correspondientes estimadas con el generador de Engen y Lillegård, en el ejemplo 3.9.

Con los mismos valores simulados x_1, x_2, \dots, x_{30} de la distribución $Gamma(3, 5)$, ahora se estimarán las densidades finales bayesianas de α y β , pero ahora cuando la distribución inicial es la no informativa de referencia (ver [2]),

$$\pi(\alpha, \beta) \propto \beta \left(\alpha \frac{d\psi(\alpha)}{d\alpha} - 1 \right)^{\frac{1}{2}},$$

donde, como antes,

$$\psi(\alpha) = \frac{d}{d\alpha} (\log \Gamma(\alpha)).$$

Por lo tanto,

$$p(\alpha, \beta \mid x_1, x_2, \dots, x_n) \propto \frac{1}{[\Gamma(\alpha)]^n \beta^{n\alpha}} (\prod_{i=1}^n x_i)^{\alpha-1} \exp\left(-\frac{1}{\beta} \sum_{i=1}^n x_i\right) \beta \left(\alpha \frac{d\psi(\alpha)}{d\alpha} - 1 \right)^{\frac{1}{2}}.$$

Otra vez, por medio del algoritmo de Metropolis-Hastings, generaremos una cadena de Markov cuya distribución de equilibrio sea la distribución final. Para ello, de nuevo primero efectuamos la reparametrización

$$a = \log(\alpha), \quad b = \log(\beta),$$

luego, si $\pi^{**}(a, b \mid x_1, x_2, \dots, x_n)$ es la densidad final de a y b dados x_1, x_2, \dots, x_n , se tiene

$$\pi^{**}(a, b \mid x_1, x_2, \dots, x_n) \propto \frac{1}{[\Gamma(e^a)]^n e^{(ne^a-1)b}} (\prod_{i=1}^n x_i)^{e^a-1} e^{(-\frac{1}{e^b} \sum_{i=1}^n x_i + a+b)} \left(e^a \frac{d\psi(e^a)}{d\alpha} - 1 \right)^{\frac{1}{2}},$$

con la cual obtenemos la aproximación normal asintótica de $\pi^{**}(a, b \mid x_1, x_2, \dots, x_n)$, la cual es

$$\tilde{g}(a, b \mid (\hat{a}, \hat{b})) = N(a, b \mid (\hat{a}, \hat{b}), V(\hat{a}, \hat{b})),$$

donde (\hat{a}, \hat{b}) y $V(\hat{a}, \hat{b})$ denotan la media y la matriz de varianza-covarianza de la distribución normal bivariada de a y b , y (\hat{a}, \hat{b}) es la moda estimada de $\pi^{**}(a, b | x_1, x_2, \dots, x_n)$.

Ejemplo 3.12 *Se generaron 10000 observaciones de (α, β) con las que se estimaron numéricamente las densidades finales de α y β , cuando la distribución inicial es la no informativa de referencia. Las gráficas son, respectivamente, las que aparecen en color negro en las figuras 3.10 y 3.11. La gráfica en color azul es la gráfica de la densidad fiducial de β estimada con el generador de Engen y Lillegård.*

De las figuras 3.10 y 3.11, obsérvese que el parecido entre la densidad fiducial (simulada) y la densidad final es grande cuando se utiliza la distribución inicial de Jeffreys, y menor cuando se utiliza la de referencia. Desde el punto de vista de la Inferencia Bayesiana, esto es muy importante porque en general sugiere que en ausencia de información inicial sobre el parámetro, la distribución inicial del parámetro que debería usarse es la no informativa de Jeffreys. También, usando como respaldo la Inferencia Bayesiana, esto sugiere que Fisher tenía razón en cuanto a su posición de obtener una distribución del parámetro basándose únicamente en los datos, la cual se puede usar para hacer inferencia desde el punto de vista frecuentista.

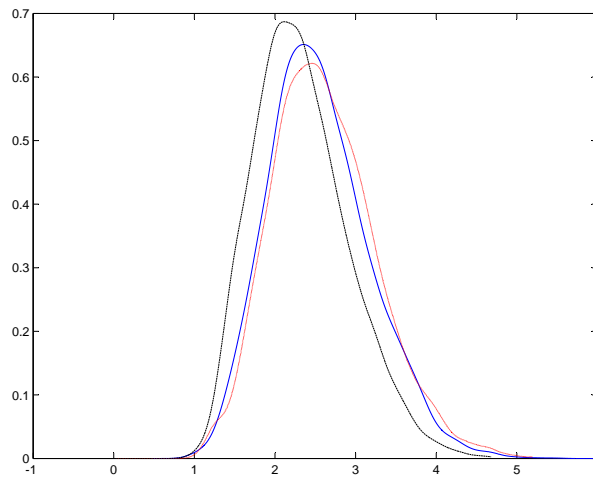


Figura 3.10: Densidades fiducial (en color azul) y finales (en rojo y negro) del parámetro α de la distribución gamma.²

²La gráfica en color rojo se obtuvo usando como distribución inicial la no informativa de Jeffreys; la gráfica en color negro se obtuvo usando como distribución inicial la de referencia de Bernardo

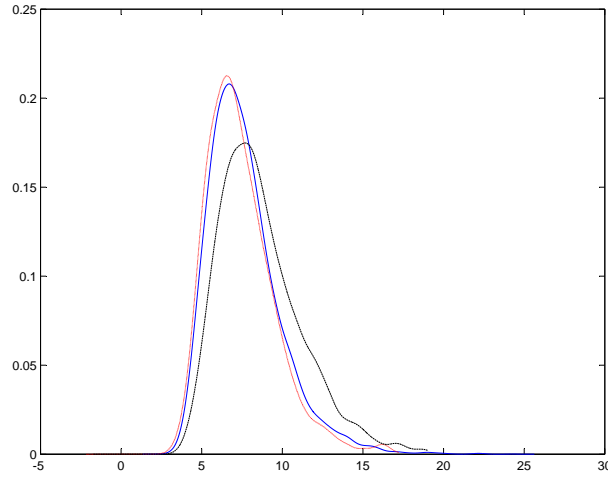


Figura 3.11: Densidades fiducial (en azul) y finales (en rojo y negro) del parámetro β de la distribución gamma. ²

3.4. Distribución Exponencial Truncada

Como ya se mencionó, otro caso que no es de grupo pero es de la familia exponencial, y en el que también podemos encontrar el generador de la densidad fiducial, es el de la Exponencial Truncada.

Proposición 3.13 Sean x_1, x_2, \dots, x_n observaciones de la muestra aleatoria X_1, X_2, \dots, X_n de la distribución Exponencial Truncada

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

Entonces existe el generador de la densidad fiducial de θ .

Demostración. (La demostración que damos esencialmente fue tomada de Lindqvist y Taraldsen [20].) La función de densidad conjunta es

$$f(x_1, x_2, \dots, x_n; \theta) = \left(\frac{\theta}{e^{\theta} - 1} \right)^n e^{\theta \sum_{i=1}^n x_i}.$$

El estadístico

$$T(\underline{X}) = \sum_{i=1}^n X_i$$

es suficiente minimal para θ . Ahora, puesto que $F(X; \theta)$ tiene distribución uniforme en $(0, 1)$, entonces de la igualdad,

$$\frac{e^{\theta X} - 1}{e^{\theta} - 1} = U,$$

donde U tiene distribución uniforme en $(0, 1)$, se tiene

$$X = \frac{\log [U (e^\theta - 1) + 1]}{\theta}. \quad (3.8)$$

De la ecuación (3.8) tenemos

$$T(\underline{U}; \theta) = \left(\sum_{i=1}^n \frac{\log [U_i (e^\theta - 1) + 1]}{\theta} \right),$$

donde (U_1, U_2, \dots, U_n) es una muestra aleatoria de la distribución uniforme en $(0, 1)$. Si

$$t = \sum_{i=1}^n x_i$$

es un valor observado de $T(\underline{X})$, entonces

$$\sum_{i=1}^n \frac{\log [U_i (e^\theta - 1) + 1]}{\theta} = t.$$

Si (u_1, u_2, \dots, u_n) es una observación de la muestra aleatoria (U_1, U_2, \dots, U_n) de la distribución uniforme en $(0, 1)$, entonces

$$\sum_{i=1}^n \frac{\log [u_i (e^\theta - 1) + 1]}{\theta}$$

es decreciente como función de θ , y la ecuación

$$\sum_{i=1}^n \frac{\log [u_i (e^\theta - 1) + 1]}{\theta} = t$$

tiene solución única. Aunque no la podemos resolver analíticamente, podemos aproximar numéricamente la solución. De esta forma podemos generar tantos valores de θ como queramos, con los cuales podemos estimar su densidad fiducial. ■

Ejemplo 3.14 Como ilustración numérica se hicieron las simulaciones siguientes: Otra vez, para cada combinación de (θ, n, t) del ejemplo (2.11), se obtuvieron 10000 observaciones de θ con las que se estimaron a través del generador de Engen y Lillegård las densidades fiduciales de θ en cada caso. Sus gráficas son las que aparecen en color rojo en las figuras 3.12, 3.13, 3.14, 3.15, 3.16, 3.17, 3.18 y 3.19. En cada una de estas figuras, la gráfica en color azul es la gráfica de la densidad fiducial, vista en el mismo ejemplo (2.11). Nuevamente observamos que se replica el parecido ya visto en O'Reilly y Rueda [23]. De las réplicas se puede observar que el parecido es enorme, incluso para el tamaño de muestra $n = 1$.

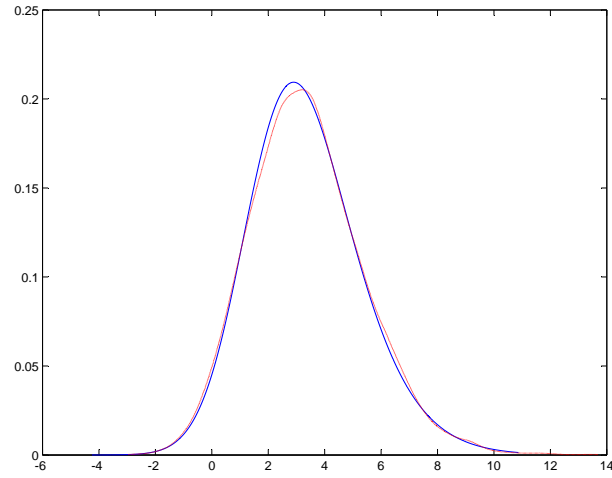


Figura 3.12: ($\theta = 1, n = 5, t = 3.6906$)

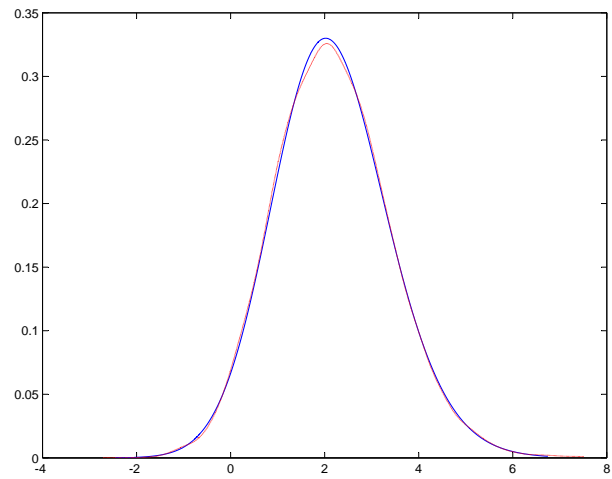


Figura 3.13: ($\theta = 2, n = 10, t = 6.6743$)

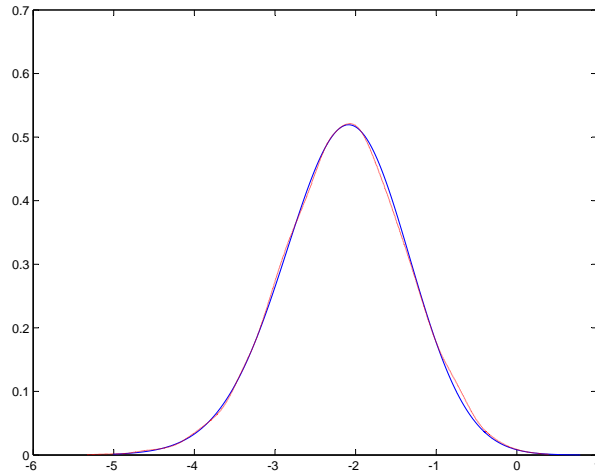


Figura 3.14: ($\theta = -2, n = 25, t = 8.3469$)

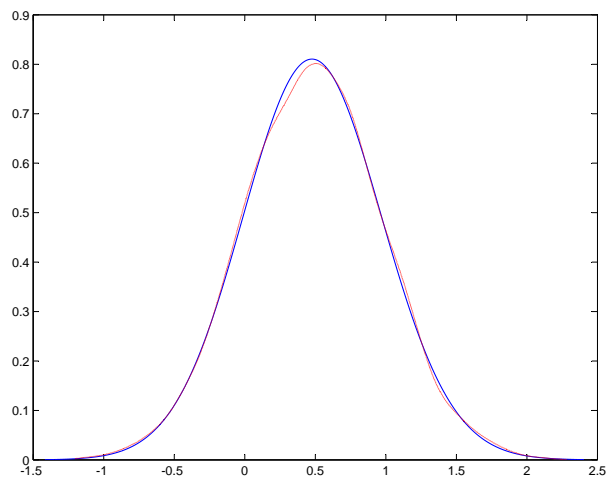


Figura 3.15: ($\theta = 0, n = 50, t = 27.0027$)

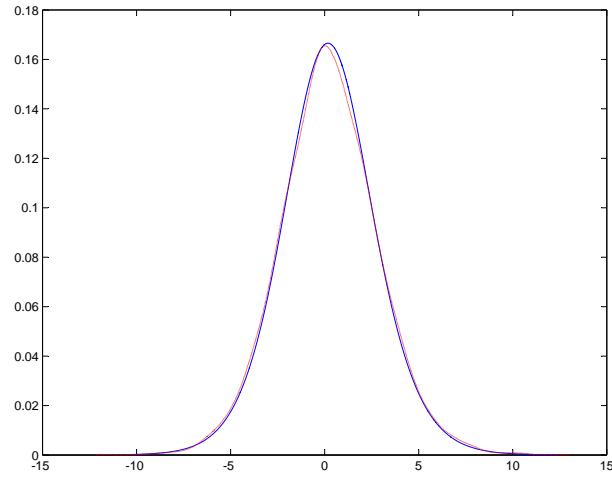


Figura 3.16: ($\theta = 4$, $n = 2$, $t = 1.0360$)

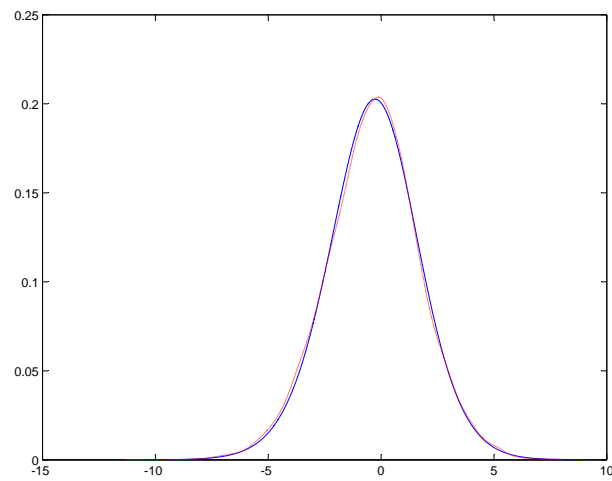


Figura 3.17: ($\theta = 0$, $n = 3$, $t = 1.4168$)

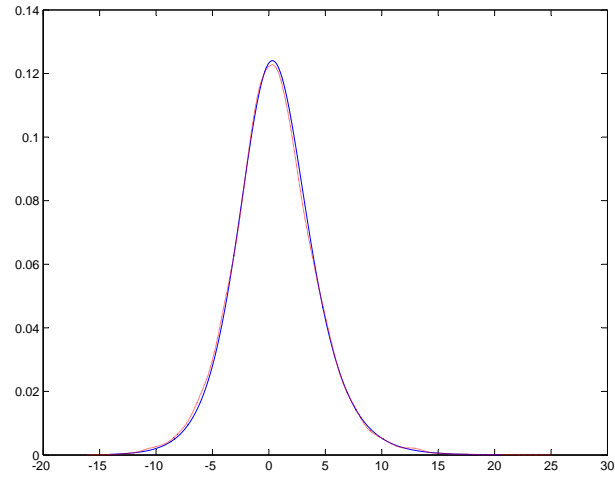


Figura 3.18: ($\theta = 2$, $n = 1$, $t = 0.5585$)

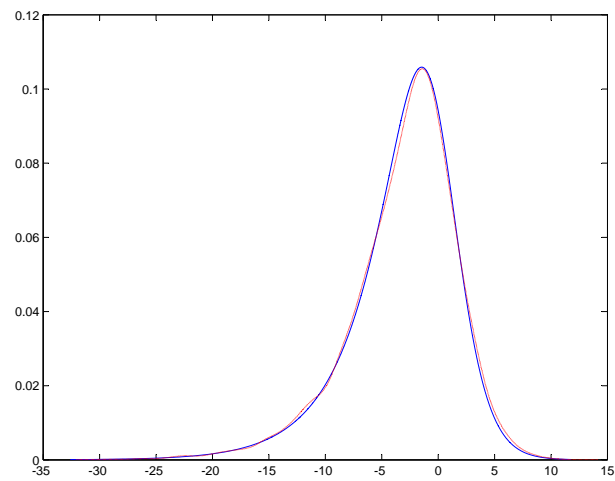


Figura 3.19: ($\theta = -3$, $n = 1$, $t = 0.2519$)

Con este ejemplo y otros similares, se ha encontrado que si bien algebraicamente no puede haber una distribución final bayesiana que coincida con la distribución fiducial en los casos que no son de grupo, en la práctica la densidad fiducial obtenida con el generador de Engen y Lillegård o directamente, como se mostró en el Capítulo 2, es muy parecida a la distribución final que se obtiene cuando la distribución inicial es la no informativa de Jeffreys.

Observación 3.15 *Como ya sabemos, puesto que en la distribución exponencial truncada el estadístico*

$$T(\underline{X}) = \sum_{i=1}^n X_i$$

es suficiente minimal para θ , otra vez, de acuerdo a lo expuesto en la secciones 2.1 y 2.3, la función de distribución fiducial de θ está dada por

$$H(\theta; t) = 1 - P_{\theta}(T \leq t),$$

de donde

$$H(\theta; t) = P_{\theta}(T \geq t).$$

Por lo tanto, la distribución fiducial de θ la podemos estimar también por el método alternativo dado en la observación 3.7.

Ejemplo 3.16 *Para la misma combinación ($\theta = -2$, $n = 25$, $t = 8.3469$) con la que se obtuvieron las gráficas de la figura 3.14, y con las mismas observaciones x_1, x_2, \dots, x_{25} para las que $t = 8.3469$, mediante el método alternativo (algoritmo 3.8) se estimó la densidad fiducial de θ . Su gráfica es la de color negro en la figura 3.20. La gráfica roja es la misma de la figura 3.14, o sea la gráfica de la densidad fiducial estimada mediante el generador de Engen y Lillegård. Nuevamente vemos que ambas gráficas prácticamente coinciden.*

³La gráfica en color rojo es la obtenida con el generador de Engen y Lillegård; la gráfica en color negro es la obtenida por medio del algoritmo 3.8

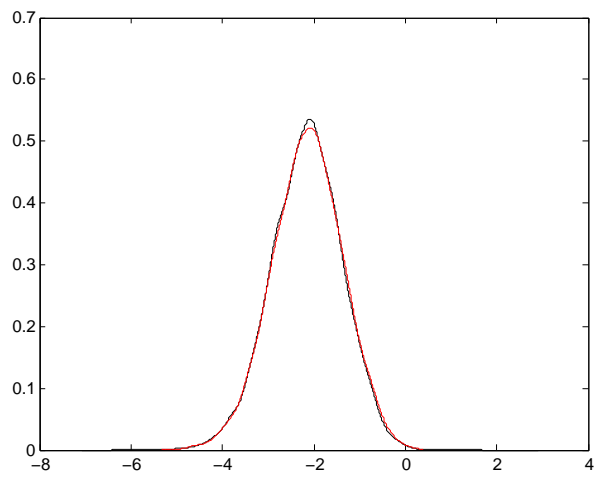


Figura 3.20: Densidades fiduciales del parámetro θ de la distribución exponencial truncada.

Capítulo 4

Conclusiones

En el capítulo 1 se expuso el argumento fiducial dado a conocer por Fisher en 1930 [9]. También se presentaron algunos resultados que pensamos que son importantes para entender la controversia que se originó en torno del argumento fiducial desde que fue expuesto.

En el capítulo 2, se presentó la revisión del enfoque que utiliza la significación en pruebas de hipótesis para obtener la distribución fiducial de un parámetro simple, mismo que fue desarrollado por O'Reilly [22]. Dicho enfoque contempla el caso de un parámetro simple, cuando la distribución pertenece a la familia exponencial natural. También incluye el caso de las distribuciones asociadas a la Bernoulli (binomial, binomial negativa y Poisson), en el que presenta una distribución fiducial única invariante en cuanto al intercambio entre “éxito” o “fracaso”. Además, la densidad fiducial resulta ser la misma en los casos binomial y binomial negativa. Así, la distribución fiducial propuesta satisface el principio fuerte de verosimilitud. Adicionalmente, en este capítulo también se presentaron dos resultados nuevos sobre la distribución fiducial cuando la distribución de la estadística pertenece a la familia exponencial, y se propuso utilizar las distribuciones de series de potencias para deducir la distribución fiducial en el caso de las distribuciones asociadas a la Bernoulli.

En el capítulo 3 se propuso el generador implícito de Engen y Lillegård como generador de la distribución fiducial. También se propuso un método alternativo para generar la distribución fiducial, el cual se dedujo a partir de las ideas desarrolladas en [22]. Además se consideró el generador fiducial obvio derivado de la expresión de la función de distribución fiducial. Los resultados numéricos sugieren que con los tres métodos se genera la misma distribución, lo que en particular nos permite afirmar que aún en modelos que no son de grupo el generador de Engen y Lillegård en efecto genera la distribución fiducial. Con este generador se obtuvo la distribución fiducial conjunta de (α, β) de la distribución gamma, así como la distribución fiducial del parámetro θ de la distribución exponencial truncada. En estos dos casos en que no existe grupo, se mostró numéricamente que la distribución fiducial y la distribución final Bayesiana resultan muy parecidas cuando la distribución inicial es la de Jeffreys. Desde el punto de vista de la inferencia Bayesiana este hecho es muy significativo

porque sugiere que en ausencia de información previa respecto al parámetro, la distribución inicial que debería de usarse es la no informativa de Jeffreys.

A partir de la distribución fiducial generada con cualquiera de los métodos considerados, también es posible implementar procedimientos numéricos para obtener intervalos de confianza o probar hipótesis.

Como trabajos pendientes quedan los siguientes:

- a) Obtener algún argumento formal con el cual se pueda concluir que con los tres métodos vistos en este trabajo se genera la misma distribución fiducial.
- b) Para el caso de estadística suficiente minimal T discreta y parámetro real θ continuo, llevar a cabo el análisis fiducial utilizando el enfoque de distribuciones de series de potencias.

Bibliografía

- [1] Beltrán, B. J. I. (2012). Cálculo del p-valor en pruebas de bondad de ajuste. Tesis de Maestría en Ciencias Matemáticas e Industriales. Universidad Autónoma Metropolitana, México, D. F.
- [2] Bernardo, J. M. and A. F. M. Smith (1994) Bayesian Theory. John Wiley & Sons, England.
- [3] Brillinger, D. R. (1962). Examples Bearing on the Definition of Fiducial Probability with a Bibliography. *The Annals of Mathematical Statistics*, 33(4), 1349-1355.
- [4] Brown, L. D., T. T. Cai and A. Das Gupta (2001). Interval estimation for a binomial proportion. *Statistical Science*, 16(2), 101-133.
- [5] Casella, G. and R. L. Berger (2002). *Statistical Inference*. Duxbury, USA.
- [6] Deemer, W. L. and D. F. Votaw (1955). Estimation of parameters of truncated or censored exponential distributions. *Annals of Mathematical Statistics*, 26, 498-504.
- [7] Efron, B. (1998). R. A. Fisher in the 21st Century. *Statistical Science*, 13, 95-122.
- [8] Engen, S. and M. Lillegård (1997). Stochastic simulations conditioned on sufficient statistics. *Biometrika*, 84(1), 235-240.
- [9] Fisher, R. A. (1930). Inverse Probability. *Proceedings of the Cambridge Philosophical Society*, 26, 528-535.
- [10] Fisher, R. A. (1935). The Fiducial Argument in Statistical Inference. *Annals of Eugenics*, 6, 391-398.
- [11] Fisher, R. A. (1945). The Logical Inversion of the Notation of the Random Variable. *Sankhya*, 7, 182-187.
- [12] Fisher, R. A. (1956). *Statistical Methods and Scientific Inference*. Oliver and Lloyd, Edinburgh and London.

- [13] Fraser. D. A. S. (1961). The fiducial method and invariance. *Biometrika*, 48, 261-280.
- [14] Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference*. Chapman & Hall, Great Britain.
- [15] Glaser, R. E. (1976). The ratio of the geometric mean to the arithmetic mean for a random sample from a gamma distribution. *Journal of the American Statistical Association*. 71(354), 480-487.
- [16] Gutiérrez Peña, E. (1997). *Métodos Computacionales en la Inferencia Bayesiana*. Serie Monografías, IIMAS, UNAM.
- [17] Hannig, J. (2009). On generalized fiducial inference. *Statistica Sinica*, 19, 491-544.
- [18] Johnson, N. L., S. Kotz and A. W. Kemp (1992) *Univariate discrete distributions*. John Wiley & Sons, New York.
- [19] Lindley, D. V. (1958). Fiducial Distributions and Bayes' Theorem. *Journal of the Royal Statistical Society, Series B (Methodological)*, 20, 1, 102-107.
- [20] Lindqvist, B. H. and G. Taraldsen (2005). Monte Carlo conditioning on a sufficient statistic. *Biometrika*, 92(2), 451-464.
- [21] O'Hagan, A. (1994). *Kendall's Advanced Theory of Statistics, Volume 2B, Bayesian Inference*. John Wiley & Sons, USA.
- [22] O'Reilly, F. (2003). *Significance Distributions*. Reporte Técnico, Unidad de Publicaciones y Difusión, IIMAS, UNAM.
- [23] O'Reilly, F. and R. Rueda (2007). Fiducial Inference for the Exponential Distribution. *Communications in Statistics-Theory and Methods*, 36, 2207-2212.
- [24] Pedersen, J.G. (1978). Fiducial inference. *International Statistical Review*, 46, 147-170.
- [25] Quenoille, M. H. ((1958). *The Fundamentals of Statistical Reasoning*. Charles Griffin, London.
- [26] Rosenbaltt, M. (1952). Remarks on a mulyivariate transformation. *Annals of Mathematical Statistics*, 23, 470-472.
- [27] Taraldsen G. and B. H. Lindqvist (2013). Fiducial theory and optimal inference. *The Annals of Statistics*, 41(1), 323-341.
- [28] Tukey, J. W. (1947). Some examples with fiducial relevance. *Annals of Mathematical Statistics*, 28, 687-695.

- [29] Wang, Y. H. (2000). Fiducial intervals: what are they? *The American Statistician*, 54(2), 105-111.