

Lema de Neyman–Pearson para distribuciones de confianza basadas en estadísticas suficientes

Neyman–Pearson lemma for confidence distributions based on sufficient statistics

EDILBERTO NÁJERA RANGEL

Universidad Juárez Autónoma de Tabasco, Cunduacán, México

BRALY GUADALUPE PERALTA REYES

Instituto Tecnológico Superior de Comalcalco, México

AROLDO PÉREZ PÉREZ

Universidad Juárez Autónoma de Tabasco, Cunduacán, México

RESUMEN. En 2002 SCHWEDER y HJORT [9] obtuvieron una nueva versión del lema de Neyman–Pearson para distribuciones de confianza, junto con otro resultado optimal. Bajo ciertas hipótesis, esencialmente dice que la distribución de confianza basada en una estadística suficiente del parámetro de interés es menos dispersa en promedio que cualquier otra distribución de confianza basada en otra estadística de la muestra. El interés principal en este trabajo de divulgación es hacer un desarrollo detallado de las demostraciones dadas por SCHWEDER y HJORT, así como también ilustrar numéricamente tales resultados, lo cual no se hace en [9], y dar un algoritmo para realizar pruebas de hipótesis unilaterales y obtener intervalos de confianza.

Key words and phrases. Confidence distribution, sufficient statistic, hypothesis testing

ABSTRACT. SCHWEDER and HJORT in 2002 obtained a new version of the Neyman–Pearson lemma for confidence distributions, along with other optimal result. Under certain assumptions, essentially says that the confidence distribution based on a sufficient statistic of the parameter of interest is less dispersed on average than any other confidence distribution based on other statistics of the sample. The main interest

in this expository paper is to make a detailed development of the proofs given by SCHWEDER and HJORT, and also, give a numerical illustration of these results, which is not done in [9], and give an algorithm for testing unilateral hypothesis and obtain confidence intervals.

2010 AMS Mathematics Subject Classification. 62F03, 62B05

1. Introducción

El lema de Neyman-Pearson es uno de los resultados más famosos en pruebas de hipótesis estadísticas. Describe de manera clara cuál procedimiento de prueba debemos aplicar cuando la forma de la distribución de probabilidad de la muestra aleatoria depende sólo del valor de un único parámetro desconocido y las hipótesis nula y alternativa son simples. Sin embargo, en muchas aplicaciones, la distribución de interés tiene más de un parámetro desconocido, y suele suceder que nuestro interés se centre en sólo uno de ellos. Al vector formado por los demás parámetros desconocidos se le acostumbra llamar vector de ruido. También, en la mayoría de las aplicaciones, las pruebas estadísticas consideran no solamente hipótesis simples, sino hipótesis compuestas. En el 2002 SCHWEDER y HJORT [9] obtuvieron una versión nueva del lema de Neyman-Pearson para distribuciones de confianza que permite considerar hipótesis estadísticas compuestas unilaterales, cuando hay incluso un vector de ruido. Cabe destacar, como se verá en la sección 2, la relevancia de las distribuciones de confianza, pues éstas constituyen la base de la inferencia en la estadística clásica (ver [9]), debido a que los intervalos de confianza para un parámetro simple son generados por percentiles de una distribución de confianza y los valores p unilaterales son valores de distribuciones de confianza acumuladas; siendo los intervalos de confianza y los valores p la forma básica, en la tradición frecuentista, de presentar los reportes estadísticos. La relación cercana entre los valores p y los intervalos de confianza nos permite verlos de forma unificada a través de una distribución de confianza. Volviendo a la versión del lema de Neyman-Pearson obtenida por SCHWEDER y HJORT [9], ésta esencialmente dice, como se verá en la sección 3, que la distribución de confianza basada en una estadística suficiente del parámetro de interés con razón de verosimilitudes creciente, es menos dispersa en promedio que cualquier otra distribución de confianza basada en otra estadística de la muestra. De esta manera, esta generalización nos permite obtener mejores procedimientos para probar hipótesis estadísticas compuestas unilaterales, esto es, si se utiliza una distribución de confianza basada en una estadística suficiente con razón de verosimilitudes creciente, las posibilidades de cometer errores tanto del tipo I como del tipo II son menores a las que se obtendrían utilizando una distribución de confianza basada en otra estadística. Debido a la importancia ya señalada de la versión de SCHWEDER y HJORT del lema de Neyman-Pearson, en este trabajo de divulgación, en la sección 3, se enuncia y demuestra este resultado, junto con otro resultado optimal, y ambas demostraciones se enriquecen con detalles omitidos en [9]. Por último, en la sección 4 presentamos varios ejemplos numéricos que ilustran con claridad

las propiedades de tales resultados; damos además un algoritmo para hacer pruebas de hipótesis unilaterales y encontrar intervalos de confianza.

2. Distribuciones de confianza, pruebas de hipótesis e intervalos de confianza

Como ya se señaló en la introducción, las distribuciones de confianza constituyen la base de la inferencia en la estadística clásica, debido a que como se verá en esta sección, las distribuciones de confianza nos permiten ver de forma unificada las pruebas de hipótesis y los intervalos de confianza. Por otra parte, mientras más angostos sean los intervalos de confianza, mejores son, siempre y cuando mantengan la confianza que se afirma. De esta manera, es siempre deseable que la distribución de confianza tenga tan poca dispersión como sea posible.

2.1. Distribuciones de confianza. En la siguiente definición el contexto es un modelo paramétrico de una población descrita por una variable aleatoria X con distribución de probabilidad $P(x; \psi, \omega)$, donde ψ es un parámetro real de interés primario que pertenece a un intervalo finito o infinito, y ω es un vector formado por los demás parámetros reales referidos como de ruido.

Definición 1. Sean $\underline{X} = (X_1, \dots, X_n)$ una muestra aleatoria de la distribución continua $P(x; \psi, \omega)$ y \underline{x} una observación de \underline{X} . Una función de distribución continua $C(\psi; \underline{x})$ de ψ , es una distribución de confianza de ψ si

$$P(C(\psi; \underline{X}) \leq a; (\psi, \omega)) = a, \quad \text{con } 0 \leq a \leq 1,$$

para cualquier par (ψ, ω) .

Observación 1. La definición 1 se centra en el hecho de que la variable aleatoria $C(\psi; \underline{X})$ tiene distribución uniforme cuando los parámetros verdaderos son (ψ, ω) .

Observación 2. El término distribución de confianza es la interpretación de EFRON (ver [5]) de la distribución fiducial de FISHER (ver [6] y [8]).

Si $0 < \alpha < 1$, sea

$$C^{-1}(\alpha; \underline{x}) = \inf\{\psi : C(\psi; \underline{x}) \geq \alpha\}.$$

Para intervalos de confianza unilaterales $(-\infty, \psi_\alpha)$, donde $\psi_\alpha = C^{-1}(\alpha; \underline{X})$, la probabilidad de cobertura (o cobertura frecuentista) es α , porque

$$P(\psi \leq \psi_\alpha) = P_{\psi, \omega}(C(\psi; \underline{X}) \leq C(\psi_\alpha; \underline{X})) = P_{\psi, \omega}(C(\psi; \underline{X}) \leq \alpha) = \alpha.$$

De igual modo, para intervalos de confianza bilaterales $(\psi_\alpha, \psi_\beta)$ la probabilidad de cobertura es $\beta - \alpha$, porque

$$\begin{aligned} P(\psi_\alpha \leq \psi \leq \psi_\beta) &= P_{\psi, \omega}(C(\psi_\alpha; \underline{X}) \leq C(\psi; \underline{X}) \leq C(\psi_\beta; \underline{X})) \\ &= P_{\psi, \omega}(\alpha \leq C(\psi; \underline{X}) \leq \beta) \\ &= \beta - \alpha. \end{aligned}$$

Definición 2. Una familia de funciones de densidad de probabilidad o de funciones de probabilidad $\{g(t; \theta) : \theta \in \Theta \subseteq \mathbb{R}\}$ de una variable aleatoria T , tiene razón de verosimilitudes monótona si, siempre que $\theta_1 < \theta_2$,

$$\frac{g(t; \theta_2)}{g(t; \theta_1)}$$

es monótona como función de t en $\{t : g(t; \theta_1) > 0 \text{ ó } g(t; \theta_2) > 0\}$, donde por definición $\frac{c}{0} = \infty$ si $c > 0$.

Proposición 1. Sea T una variable aleatoria cuya familia de densidades de probabilidad o de funciones de probabilidad es $\{g(t; \theta) : \theta \in \Theta \subseteq \mathbb{R}\}$, tal que para $\theta_1 < \theta_2$,

$$\frac{g(t; \theta_2)}{g(t; \theta_1)}$$

es creciente como función de t . Entonces

$$P_{\theta_2}(T \leq t) \leq P_{\theta_1}(T \leq t).$$

Demostración. Sean $t_1 < t < t_2$. De aquí

$$\frac{g(t_1; \theta_2)}{g(t_1; \theta_1)} \leq \frac{g(t_2; \theta_2)}{g(t_2; \theta_1)} \implies g(t_1; \theta_2)g(t_2; \theta_1) \leq g(t_1; \theta_1)g(t_2; \theta_2).$$

Integrando primero sobre t_1 de $-\infty$ a t y después sobre t_2 de t a ∞ se tiene

$$P_{\theta_2}(T \leq t)(1 - P_{\theta_1}(T \leq t)) \leq P_{\theta_1}(T \leq t)(1 - P_{\theta_2}(T \leq t)),$$

de donde,

$$P_{\theta_2}(T \leq t) \leq P_{\theta_1}(T \leq t).$$

✓

Corolario 1. Si para $\theta_1 < \theta_2$,

$$\frac{g(t; \theta_2)}{g(t; \theta_1)}$$

es decreciente como función de t , entonces

$$P_{\theta_1}(T \leq t) \leq P_{\theta_2}(T \leq t).$$

Corolario 2. Sean T una variable aleatoria continua como en la proposición 1, $P_\theta(T \leq t)$ continua como función de θ , y $\Theta = (\theta_{\text{inf}}, \theta_{\text{sup}})$. Si $P_\theta(T \leq t) \rightarrow 1$ cuando $\theta \rightarrow \theta_{\text{inf}}$ y $P_\theta(T \leq t) \rightarrow 0$ cuando $\theta \rightarrow \theta_{\text{sup}}$, entonces $C(\theta; t) = 1 - P_\theta(T \leq t)$ es una distribución de confianza de θ .

Definición 3. La familia $\{g(t; \theta) : \theta \in \Theta \subseteq \mathbb{R}\}$, de funciones de densidad de probabilidad o de funciones de probabilidad, se llama una **familia exponencial** si $g(t; \theta)$ se puede expresar en la forma

$$g(t; \theta) = h(t)c(\theta) \exp \left(\sum_{i=1}^k w_i(\theta)q_i(t) \right),$$

donde $h(t) \geq 0$ y q_1, \dots, q_k son funciones de valor real que no dependen de θ , y $c(\theta) \geq 0$ y w_1, \dots, w_k son funciones de valor real que no dependen de t .

Proposición 2. Sea

$$g(t; \theta) = h(t)c(\theta) \exp \left(\sum_{i=1}^k w_i(\theta)q_i(t) \right), \quad \theta \in \Theta \subseteq \mathbb{R},$$

una familia exponencial. Si todas las funciones w_i son monótonas crecientes o todas son monótonas decrecientes, y si todas las funciones q_i son monótonas crecientes o todas son monótonas decrecientes, entonces $g(t; \theta)$ tiene razón de verosimilitudes monótona.

Demostración. Sean $\theta_1 < \theta_2$. Entonces

$$\frac{g(t; \theta_2)}{g(t; \theta_1)} = \frac{c(\theta_2)}{c(\theta_1)} \exp \left(\sum_{i=1}^k (w_i(\theta_2) - w_i(\theta_1))q_i(t) \right)$$

es monótona como función de t . □

Ejemplo 1. Sea X_1, \dots, X_n una muestra aleatoria de una distribución gama con sólo el parámetro de forma $\alpha > 0$ desconocido, β conocido. Sea $S = \sum_{i=1}^n X_i$. Entonces S tiene una distribución gama con parámetro de forma $n\alpha$ y β conocido. Si f es su función de densidad,

$$f(s; n\alpha) = \frac{1}{\Gamma(n\alpha)\beta^{n\alpha}} s^{n\alpha-1} e^{-\frac{s}{\beta}},$$

la cual se puede escribir en la forma

$$f(s; n\alpha) = \frac{1}{\Gamma(n\alpha)\beta^{n\alpha}} e^{-\frac{s}{\beta}} e^{(n\alpha-1) \log s},$$

es decir, $\{f(s; n\alpha) : \alpha > 0\}$ es una familia exponencial. Puesto que $n\alpha - 1$ es creciente como función de α , y $\log s$ es creciente como función de s , por la proposición 2, $f(s; n\alpha)$ tiene razón de verosimilitudes creciente. Si $F(s; n\alpha)$ es su función de distribución, entonces $F(s; n\alpha) \rightarrow 1$ cuando $\alpha \rightarrow 0$, y $F(s; n\alpha) \rightarrow 0$ cuando $\alpha \rightarrow \infty$. Por el corolario 2

$$C(\alpha; s) = 1 - F(s; n\alpha)$$

es una distribución de confianza de α . La función de densidad de confianza de α , una vez que se observó s , es

$$c(\alpha; s) = -(d/d\alpha)F(s; n\alpha) = n \int_0^s (\psi(n\alpha) - \log(t/\beta))f(t; n\alpha)dt,$$

donde $\psi(\alpha)$ es la función digama, $\psi(\alpha) = (\frac{d}{d\alpha})(\log \Gamma(\alpha))$.

Ejemplo 2. Sea X_1, \dots, X_n una muestra aleatoria de una población $N(\mu, \sigma^2)$, con μ desconocida y σ^2 conocida. Sea $T = \frac{1}{n} \sum_{i=1}^n X_i$. Entonces $T \sim N(\mu, \frac{\sigma^2}{n})$, la cual es una familia exponencial. Si $f(t; \mu)$ es la función de densidad de T , resulta que $\frac{f(t; \mu_2)}{f(t; \mu_1)}$ es creciente como función de t si $\mu_1 < \mu_2$. De aquí, si $F(t; \mu)$ es la función de distribución de T , se tiene que $F(t; \mu)$ es monótonamente decreciente como función de μ . Por otro lado, si Φ es la función de distribución normal estándar, $F(t; \mu) = P(T \leq t) = \Phi\left(\frac{\sqrt{n}(t-\mu)}{\sigma}\right)$, de donde también se ve que la función de distribución de T es monótonamente decreciente como función de μ . Puesto que $F(t; \mu) \rightarrow 1$ cuando $\mu \rightarrow -\infty$, y $F(t; \mu) \rightarrow 0$ cuando $\mu \rightarrow \infty$, entonces la distribución de confianza de μ es

$$C(\mu; t) = 1 - F(t; \mu) = 1 - \Phi\left(\frac{\sqrt{n}(t-\mu)}{\sigma}\right).$$

Una vez que se observó t , la densidad de confianza de μ es

$$c(\mu; t) = -(d/d\mu)F(t; \mu) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n(\mu-t)^2}{2\sigma^2}\right)$$

es decir, $\mu \sim N(t, \sigma^2/n)$.

2.2. Pruebas de hipótesis e intervalos de confianza.

Definición 4. Sean $\Theta \subseteq \mathbb{R}$ el conjunto de valores posibles del parámetro θ , $\Theta_0 \subset \Theta$, y Θ_0^c el complemento de Θ_0 en Θ . En un problema de prueba de hipótesis

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_0^c,$$

un valor p es un estadístico de prueba $p(\underline{X})$ que satisface $0 \leq p(\underline{x}) \leq 1$, para toda observación \underline{x} de \underline{X} , de modo que valores pequeños de $p(\underline{X})$ den evidencia de que H_1 es cierta. Se dice que $p(\underline{x})$ es el nivel de significancia alcanzado por la prueba.

La siguiente proposición muestra que las distribuciones de confianza permiten ver de forma unificada las pruebas de hipótesis y los intervalos de confianza.

Proposición 3. La confianza de la afirmación $\psi \leq \psi_0$ es el grado de confianza $C(\psi_0; \underline{x})$ del intervalo de confianza $(-\infty, C^{-1}(C(\psi_0; \underline{x})))$, y es igual al valor p observado de la prueba de hipótesis

$$H_0 : \psi \leq \psi_0 \quad \text{vs} \quad H_1 : \psi > \psi_0.$$

Demostración. Puesto que $C(\psi; \underline{x})$ es creciente como función de ψ , entonces $C(\psi_0; \underline{X})$ es un estadístico de prueba para probar

$$H_0 : \psi \leq \psi_0 \quad \text{vs} \quad H_1 : \psi > \psi_0,$$

de modo que valores pequeños de $C(\psi_0; \underline{X})$ dan evidencia de que H_1 es cierta; como $0 \leq C(\psi; \underline{x}) \leq 1$, entonces $C(\psi; \underline{X})$ es un valor p de la prueba de hipótesis. Si \underline{x} es una observación de \underline{X} , el nivel de significancia alcanzado por la prueba de hipótesis, o la confianza de la afirmación $\psi \leq \psi_0$, es

$$p(\underline{x}) = C(\psi_0; \underline{x}) = P(C(\psi; \underline{X}) \leq C(\psi_0; \underline{x})).$$

Además,

$$\begin{aligned} P(\psi \leq C^{-1}(C(\psi_0; \underline{X}))) &= P(C(\psi; \underline{X}) \leq C(C^{-1}(C(\psi_0; \underline{X})))) \\ &= P(C(\psi; \underline{X}) \leq C(\psi_0; \underline{X})), \end{aligned}$$

o sea, $C(\psi_0; \underline{x})$ también es el grado de confianza del intervalo de confianza $(-\infty, C^{-1}(C(\psi_0; \underline{x})))$. \square

Ejemplo 3. Sea X_1, \dots, X_n una muestra aleatoria de una población $N(\mu, \sigma^2)$, donde los parámetros μ y σ^2 son desconocidos, μ es el parámetro de interés y σ^2 es de ruido. Sean $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ y $S_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Para realizar la prueba de hipótesis

$$H_0 : \mu \leq \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0,$$

el estadístico de prueba es $\frac{\sqrt{n}(\bar{X} - \mu_0)}{S_{n-1}}$. En este caso el valor p es

$$p(\underline{X}) = 1 - F_{t, n-1} \left(\frac{\sqrt{n}(\bar{X} - \mu_0)}{S_{n-1}} \right) = C(\mu_0; \underline{X}),$$

donde $F_{t, n-1}$ denota a la función de distribución t de Student con $n-1$ grados de libertad, y $C(\mu; \underline{x})$ es la distribución de confianza de μ .

Ejemplo 4. Sea $\underline{x} = (x_1, \dots, x_n)$ una observación del vector $\underline{X} = (X_1, \dots, X_n)$, donde X_1, \dots, X_n es una muestra aleatoria de la distribución $N(\mu, \sigma^2)$ con los parámetros μ y σ^2 desconocidos, μ es el parámetro de interés y σ^2 es de ruido. Sean \bar{X} y S_{n-1}^2 definidas como en el ejemplo 3. Si $F_{t, n-1}$ denota a la función de distribución t de Student con $n-1$ grados de libertad, sabemos que,

$$C(\mu; \underline{x}) = 1 - F_{t, n-1} \left(\frac{\sqrt{n}(\bar{x} - \mu)}{s_{n-1}} \right)$$

es una distribución de confianza para μ . Podemos determinar un intervalo de $(1 - \alpha)100\%$ de confianza para μ como sigue: Sean $\mu_{\frac{\alpha}{2}}$ y $\mu_{1-\frac{\alpha}{2}}$ tales que

$$C(\mu_{\frac{\alpha}{2}}; \underline{x}) = \frac{\alpha}{2} \quad \text{y} \quad C(\mu_{1-\frac{\alpha}{2}}; \underline{x}) = 1 - \frac{\alpha}{2},$$

o sea,

$$1 - F_{t, n-1} \left(\frac{\sqrt{n}(\bar{x} - \mu_{\frac{\alpha}{2}})}{s_{n-1}} \right) = \frac{\alpha}{2} \quad \text{y} \quad 1 - F_{t, n-1} \left(\frac{\sqrt{n}(\bar{x} - \mu_{1-\frac{\alpha}{2}})}{s_{n-1}} \right) = 1 - \frac{\alpha}{2}.$$

De la primera de estas dos últimas igualdades tenemos

$$\mu_{\frac{\alpha}{2}} = \bar{x} - t_{1-\frac{\alpha}{2}, n-1} \frac{s_{n-1}}{\sqrt{n}},$$

y de la segunda

$$\mu_{1-\frac{\alpha}{2}} = \bar{x} + t_{1-\frac{\alpha}{2}, n-1} \frac{s_{n-1}}{\sqrt{n}},$$

donde $t_{1-\frac{\alpha}{2}, n-1}$ es el percentil $(1 - \frac{\alpha}{2})100$ de la distribución t de Student con $n - 1$ grados de libertad. Por lo tanto, un intervalo de $(1 - \alpha)100\%$ de confianza para μ es

$$\bar{x} - t_{1-\frac{\alpha}{2}, n-1} \frac{s_{n-1}}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{1-\frac{\alpha}{2}, n-1} \frac{s_{n-1}}{\sqrt{n}},$$

el cual coincide con el intervalo de confianza respectivo que se obtiene usando el método común de estimación por intervalo.

2.3. Lema clásico de Neyman-Pearson. Como referencia, enunciamos en esta última subsección de la presente sección, el lema de Neyman-Pearson clásico, cuya demostración puede ser consultada en varios textos sobre estadística matemática, como por ejemplo [2] y [7].

Definición 5. Dada una prueba de hipótesis con región de rechazo R , la función $\beta : \Theta \rightarrow [0, 1]$ dada por $\beta(\theta) = P_{\theta}(\underline{X} \in R)$ se llama la función potencia de la prueba.

Definición 6. Consideremos el problema de prueba de hipótesis $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_0^c$.

(a) Para $0 \leq \alpha \leq 1$, una prueba con función potencia $\beta(\theta)$ es una prueba de tamaño α si $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$.

(b) Para $0 \leq \alpha \leq 1$, una prueba con función potencia $\beta(\theta)$ es una prueba de nivel α si $\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$.

(c) Sea C una clase de pruebas para probar $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_0^c$. Se dice que una prueba en la clase C , con función potencia $\beta(\theta)$, es una prueba uniformemente más potente (UMP) si para cada función potencia $\beta'(\theta)$ de cada prueba en C se cumple que $\beta(\theta) \geq \beta'(\theta)$ para toda $\theta \in \Theta_0^c$.

Teorema 1. (Lema de Neyman-Pearson) Sea la prueba de hipótesis $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$, donde $f(\underline{x}; \theta_i)$, $i = 0, 1$, es la función de densidad o de probabilidad, con región de rechazo R tal que

$$\underline{x} \in R \text{ si } f(\underline{x}; \theta_1) > k f(\underline{x}; \theta_0) \text{ y } \underline{x} \in R^c \text{ si } f(\underline{x}; \theta_1) < k f(\underline{x}; \theta_0) \quad (1)$$

para algún $k \geq 0$, y

$$\alpha = P_{\theta_0}(\underline{X} \in R). \quad (2)$$

Entonces,

(a) (Suficiencia) Cualquier prueba que satisface las condiciones anteriores es una prueba de nivel α uniformemente más potente.

(b) (Necesidad) Si existe una prueba que satisface las condiciones anteriores con $k > 0$, entonces toda prueba de nivel α uniformemente más potente es una prueba de tamaño α , y toda prueba de nivel α uniformemente más potente satisface las dos primeras condiciones, excepto quizá en un conjunto A que satisface $P_{\theta_0}(\underline{X} \in A) = P_{\theta_1}(\underline{X} \in A) = 0$.

3. Lemas de Neyman–Pearson para valores p y distribuciones de confianza

Cuando probamos las hipótesis $H_0 : \psi \leq \psi_0$ vs $H_1 : \psi > \psi_0$, y la distribución de confianza de ψ es $C(\psi)$, rechazamos H_0 al nivel de confianza α si $C(\psi_0) < \alpha$. La potencia de la prueba es $P_\psi(C(\psi_0) < \alpha)$. Cuando se utiliza una distribución de confianza basada en una estadística suficiente con razón de verosimilitudes creciente, el siguiente teorema establece que cuando $\psi < \psi_0$ es menos probable rechazar H_0 , que cuando se utiliza una distribución de confianza basada en otra estadística; y cuando $\psi > \psi_0$ es más probable rechazarla. Es decir, afirma que si el estadístico de prueba es $C(\psi_0; S)$, donde $C(\psi; s)$ es la distribución de confianza de ψ basada en la estadística suficiente S con razón de verosimilitudes creciente, entonces la prueba tiene errores de tipo I y II menores que los respectivos errores que se tienen cuando el estadístico de prueba es $C(\psi_0; T)$, donde $C(\psi; t)$ es la distribución de confianza de ψ basada en otra estadística T . Así pues, este teorema es una generalización no sólo del lema de Neyman–Pearson, sino también del teorema de KARLIN RUBIN (ver [2]).

Los conceptos y resultados de esperanza condicional que se usan a continuación pueden ser consultados, por ejemplo, en [1] y [4].

Teorema 2. (Lema de Neyman–Pearson para valores p). *Sea S una estadística suficiente unidimensional con razón de verosimilitudes $\frac{L(\psi_2; s)}{L(\psi_1; s)}$ creciente como función de s , siempre que $\psi_1 < \psi_2$. Sea C^S la distribución de confianza basada en S , y sea C^T la distribución de confianza de ψ basada en otra estadística T . Entonces*

$$P_\psi(C^S(\psi_0) \leq u) \leq P_\psi(C^T(\psi_0) \leq u), \text{ si } \psi < \psi_0,$$

y

$$P_\psi(C^T(\psi_0) \leq u) \leq P_\psi(C^S(\psi_0) \leq u), \text{ si } \psi_0 < \psi.$$

Demostración. Sea $\psi < \psi_0$. Supongamos que

$$(C^T(\psi_0) \leq u) = (T \in A_{\psi_0}),$$

donde $(T \in A_{\psi_0})$ es Borel-medible. Se tiene

$$\begin{aligned} P_{\psi} (C^T(\psi_0) \leq u) &= P_{\psi} (T \in A_{\psi_0}) = E_{\psi} \left(\mathbb{1}_{(T \in A_{\psi_0})} \right) \\ &= \int P_{\psi} (T \in A_{\psi_0} | S = s) dP_{\psi}^S(s) \\ &= \int P_{\psi} (T \in A_{\psi_0} | S = s) \frac{L(\psi; s)}{L(\psi_0; s)} dP_{\psi_0}^S(s) \\ &= \int P_{\psi_0} (T \in A_{\psi_0} | S = s) \frac{L(\psi; s)}{L(\psi_0; s)} dP_{\psi_0}^S(s). \end{aligned}$$

Por hipótesis, puesto que $u = P_{\psi_0} (C^T(\psi_0) \leq u) = P_{\psi_0} (C^S(\psi_0) \leq u)$,

$$\begin{aligned} P_{\psi_0} (C^T(\psi_0) \leq u) &= P_{\psi_0} (T \in A_{\psi_0}) = \int P_{\psi_0} (T \in A_{\psi_0} | S = s) dP_{\psi_0}^S(s) \\ &= P_{\psi_0} (S \geq s_{\psi_0}), \end{aligned}$$

donde, dado que $\frac{L(\psi_2; s)}{L(\psi_1; s)}$ es creciente como función de s si $\psi_1 < \psi_2$, entonces por el corolario 2,

$$(C^S(\psi_0) \leq u) = (S \geq s_{\psi_0}).$$

Como $\frac{L(\psi; s)}{L(\psi_0; s)}$ es decreciente en s , la función g tal que $0 \leq g \leq 1$ y que minimiza a

$$\int g(s) \frac{L(\psi; s)}{L(\psi_0; s)} dP_{\psi_0}^S(s)$$

sujeta a

$$\int g(s) dP_{\psi_0}^S(s) = P_{\psi_0} (S \geq s_{\psi_0})$$

es la indicadora $\mathbb{1}_{(s_{\psi_0}, \infty)}$, para la que el mínimo es

$$\int \mathbb{1}_{(s_{\psi_0}, \infty)} \frac{L(\psi; s)}{L(\psi_0; s)} dP_{\psi_0}^S(s) = \int \mathbb{1}_{(s_{\psi_0}, \infty)} dP_{\psi}^S(s) = P_{\psi} (S \geq s_{\psi_0}),$$

de donde,

$$P_{\psi} (C^T(\psi_0) \leq u) \geq P_{\psi} (C^S(\psi_0) \leq u),$$

o sea,

$$P_{\psi} (C^S(\psi_0) \leq u) \leq P_{\psi} (C^T(\psi_0) \leq u), \text{ si } \psi < \psi_0.$$

Análogamente se prueba que

$$P_{\psi} (C^T(\psi_0) \leq u) \leq P_{\psi} (C^S(\psi_0) \leq u), \text{ si } \psi_0 < \psi.$$

✓

En el sentido clásico de Neyman-Pearson, el enfoque consiste en rechazar o no rechazar una hipótesis, lo cual equivale a usar para C una medida de

dispersión de tipo indicadora. Más generalmente, diremos que $\gamma(C)$ es una medida de dispersión de C , alrededor del valor verdadero ψ_0 del parámetro, si

$$\gamma(C) = \int_{-\infty}^{\infty} \Gamma(\psi - \psi_0) dC(\psi),$$

donde $\Gamma(0) = 0$, Γ es decreciente en $(-\infty, 0)$ y creciente en $(0, \infty)$,

$$\lim_{\psi \rightarrow \infty} (1 - C(\psi))\Gamma(\psi - \psi_0) = 0,$$

y

$$\lim_{\psi \rightarrow \infty} C(-\psi)\Gamma(-\psi - \psi_0) = 0.$$

En el teorema 3, el cual establece, como ya se mencionó en la introducción, que la distribución de confianza basada en una estadística suficiente con razón de verosimilitudes creciente es menos dispersa en promedio que cualquier otra distribución de confianza basada en otra estadística, utilizaremos la proposición siguiente, cuya demostración se puede ver en [3].

Proposición 4. Si X es una v.a. no - negativa absolutamente continua con respecto a la medida de Lebesgue, entonces

$$E(X) = \int_0^{\infty} (1 - F(x)) dx,$$

donde F es la función de distribución de X .

Teorema 3. (Lema de Neyman–Pearson para distribuciones de confianza). Si S es una estadística suficiente unidimensional y la razón de verosimilitudes $\frac{L(\psi_2; s)}{L(\psi_1; s)}$ es creciente en s siempre que $\psi_1 < \psi_2$, entonces la distribución de confianza basada en S es uniformemente más potente en promedio, es decir, se cumple $E_{\psi_0} [\gamma(C^S)] \leq E_{\psi_0} [\gamma(C^T)]$ para toda función de dispersión γ y en todos los valores verdaderos del parámetro ψ_0 .

Demostración. Notemos que

$$\begin{aligned} & \int_{\psi_0}^b \Gamma(\psi - \psi_0) dC(\psi) \\ &= - \int_{\psi_0}^b \Gamma(\psi - \psi_0) d(1 - C(\psi)) \\ &= - [(1 - C(b)) \Gamma(b - \psi_0) - (1 - C(\psi_0)) \Gamma(0)] \\ & \quad + \int_{\psi_0}^b (1 - C(\psi)) d\Gamma(\psi - \psi_0) \\ &= - (1 - C(b)) \Gamma(b - \psi_0) + \int_{\psi_0}^b (1 - C(\psi)) d\Gamma(\psi - \psi_0), \end{aligned}$$

y

$$\int_{-b}^{\psi_0} \Gamma(\psi - \psi_0) dC(\psi) = C(\psi_0)\Gamma(0) - C(-b)\Gamma(-b - \psi_0) - \int_{-b}^{\psi_0} C(\psi) d\Gamma(\psi - \psi_0).$$

Así, puesto que $\lim_{b \rightarrow \infty} (1 - C(b))\Gamma(b - \psi_0) = \lim_{b \rightarrow \infty} C(-b)\Gamma(-b - \psi_0) = 0$,

$$\int_{\psi_0}^{\infty} \Gamma(\psi - \psi_0) dC(\psi) = \int_{\psi_0}^{\infty} (1 - C(\psi)) d\Gamma(\psi - \psi_0)$$

y

$$\int_{-\infty}^{\psi_0} \Gamma(\psi - \psi_0) dC(\psi) = - \int_{-\infty}^{\psi_0} C(\psi) d\Gamma(\psi - \psi_0).$$

Sea $\alpha = \psi - \psi_0$. Entonces

$$\int_{\psi_0}^{\infty} \Gamma(\psi - \psi_0) dC(\psi) = \int_0^{\infty} (1 - C(\psi + \psi_0)) d\Gamma(\psi)$$

y

$$\int_{-\infty}^{\psi_0} \Gamma(\psi - \psi_0) dC(\psi) = - \int_{-\infty}^0 C(\psi + \psi_0) d\Gamma(\psi).$$

Por lo tanto,

$$\begin{aligned} \gamma(C) &= \int_{-\infty}^{\psi_0} \Gamma(\psi - \psi_0) dC(\psi) + \int_{\psi_0}^{\infty} \Gamma(\psi - \psi_0) dC(\psi) \\ &= - \int_{-\infty}^0 C(\psi + \psi_0) d\Gamma(\psi) + \int_0^{\infty} (1 - C(\psi + \psi_0)) d\Gamma(\psi) \\ &= \int_{-\infty}^0 C(\psi + \psi_0) [d(-\Gamma(\psi))] + \int_0^{\infty} (1 - C(\psi + \psi_0)) d\Gamma(\psi). \end{aligned}$$

Ahora,

$$\begin{aligned} &E_{\psi_0} [\gamma(C^S)] \\ &= \int \int \Gamma(\psi - \psi_0) dC(\psi) dP_{\psi_0} (C^S(\psi) \leq u) \\ &= \int \left[\int_{-\infty}^0 C^S(\psi + \psi_0) [d(-\Gamma(\psi))] + \int_0^{\infty} (1 - C^S(\psi + \psi_0)) d\Gamma(\psi) \right] \\ &\times dP_{\psi_0} (C^S(\psi) \leq u). \end{aligned}$$

Por el teorema de Fubini,

$$\begin{aligned} E_{\psi_0} [\gamma(C^S)] &= \int_{-\infty}^0 \int C^S(\psi + \psi_0) dP_{\psi_0} (C^S(\psi) \leq u) d(-\Gamma(\psi)) \\ &+ \int_0^{\infty} \int (1 - C^S(\psi + \psi_0)) dP_{\psi_0} (C^S(\psi) \leq u) d\Gamma(\psi). \end{aligned}$$

Por el teorema 2 y la proposición 1,

$$\int C^S(\psi + \psi_0) dP_{\psi_0} (C^S(\psi) \leq u) \leq \int C^T(\psi + \psi_0) dP_{\psi_0} (C^T(\psi) \leq u),$$

si $\psi < 0$, y

$$\begin{aligned} & \int (1 - C^S(\psi + \psi_0)) dP_{\psi_0} (C^S(\psi) \leq u) \leq \\ & \int (1 - C^T(\psi + \psi_0)) dP_{\psi_0} (C^T(\psi) \leq u), \end{aligned}$$

si $0 < \psi$. Por lo tanto,

$$\begin{aligned} E_{\psi_0} [\gamma(C^S)] & \leq \int_{-\infty}^0 \int C^T(\psi + \psi_0) dP_{\psi_0} (C^T(\psi) \leq u) d(-\Gamma(\psi)) \\ & + \int_0^{\infty} \int (1 - C^T(\psi + \psi_0)) dP_{\psi_0} (C^T(\psi) \leq u) d\Gamma(\psi) = E_{\psi_0} [\gamma(C^T)]. \end{aligned}$$

✓

4. Ejemplos numéricos de las nuevas versiones del lema de Neyman–Pearson

En esta sección X_1, \dots, X_n será una muestra aleatoria de la distribución *gama* con sólo el parámetro de forma α desconocido. Del ejemplo 1 sabemos que la distribución de confianza de α basada en $S = X_1 + \dots + X_n$ es

$$C(\alpha; s) = 1 - F(s; n\alpha), \quad (3)$$

donde F es la función de distribución *gama*($n\alpha, \beta$), con β conocido.

Por otro lado, la función de verosimilitud de α es

$$f(\alpha; x_1, \dots, x_n) = \frac{1}{(\Gamma(\alpha))^n} (\prod_{i=1}^n x_i)^{\alpha-1} \exp\left(-\frac{1}{\beta} \sum_{i=1}^n x_i\right),$$

de donde se tiene que

$$T = \prod_{i=1}^n X_i \quad (4)$$

es una estadística suficiente (minimal) de α . Se puede demostrar que la función de densidad de T tiene razón de verosimilitudes creciente, y aunque esta densidad es difícil de manipular, al ser T una estadística suficiente de α , si observamos un valor pequeño de T entonces se espera que el valor del parámetro α también sea pequeño, y si observamos un valor grande de T entonces se espera que el valor del parámetro α también sea grande. Esto nos permite afirmar que si $\alpha_1 < \alpha_2$, entonces

$$P_{\alpha_2}(T \leq t) \leq P_{\alpha_1}(T \leq t),$$

de donde se tiene que la distribución de confianza de α basada en la estadística suficiente T es

$$C(\alpha; t) = 1 - P_\alpha(T \leq t) = P_\alpha(T > t). \quad (5)$$

Una vez que observamos un valor t de T , de (5) tenemos el siguiente algoritmo para estimar $C(\alpha, t)$, el valor en α de la distribución de confianza basada en T .

Algoritmo 1

- a) 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 gama(α, β).
- b) Para $j = 1, \dots, m$, sea $t_j = \prod_{i=1}^n x_{ji}$.
- c) Sea $k = \text{cardinalidad} \{j : t_j > t\}$. Entonces

$$C(\alpha; t) \approx \frac{k}{m}.$$

Observación 3. Puesto que $C(\alpha_0, t)$ es el valor p de la prueba de hipótesis

$$H_0 : \alpha \leq \alpha_0 \quad \text{vs} \quad H_1 : \alpha > \alpha_0,$$

el algoritmo 1 es un algoritmo para llevar a cabo esta prueba de hipótesis. También se puede emplear para probar las hipótesis

$$H_0 : \alpha \geq \alpha_0 \quad \text{vs} \quad H_1 : \alpha < \alpha_0,$$

Este algoritmo se puede utilizar siempre que se tenga una estadística suficiente cuya distribución sea difícil de obtener o de manipular.

Observación 4. Con el algoritmo 1 también se pueden construir intervalos de confianza del parámetro de interés, una vez que se observó un valor de la estadística suficiente.

Ejemplo 5. Para ilustrar numéricamente el lema de Neyman-Pearson para valores p (Teorema 2), sea $\alpha_i = 3 + i(1/10)$, $i = 0, 1, \dots, 40$. Para cada α_i generamos 10000 muestras de tamaño 10 de la distribución gama(α_i, β), con β constante. Dado $i \in \{0, \dots, 40\}$, para cada muestra generada de la distribución gama(α_i, β), de (3) obtenemos el valor correspondiente de $C(\alpha_0, s)$, donde $\alpha_0 = 5$ y s es la suma de las observaciones de la muestra, y de (5), a través del algoritmo 1, obtenemos el respectivo valor de $C(\alpha_0, t)$, donde t es el producto de las observaciones de la muestra. Así, si $u \in (0, 1)$, la proporción de valores $C(\alpha_0, s)$ tales que $C(\alpha_0, s) \leq u$ nos da una estimación de $P_{\alpha_i}(C(\alpha_0; S) \leq u)$, y la proporción de valores $C(\alpha_0, t)$ tales que $C(\alpha_0, t) \leq u$ nos da una estimación de $P_{\alpha_i}(C(\alpha_0; T) \leq u)$. Obtenemos estas estimaciones para cada $\alpha_i = 3 + i(1/10)$, $i = 0, 1, \dots, 40$, y para cada $u = 0.3, 0.4, 0.5, 0.6$. Los resultados se ilustran en las figuras 1, 2, 3 y 4, respectivamente, en las que la curva continua es la gráfica de $P_{\alpha_i}(C(\alpha_0; T) \leq u)$, y la otra curva es la gráfica de $P_{\alpha_i}(C(\alpha_0; S) \leq u)$, $u = 0.3, 0.4, 0.5, 0.6$; en este caso T es la estadística suficiente. En cada figura se observa que antes de $\alpha_0 = 5$ la curva continua está abajo de la curva

segmentada, y que después de $\alpha_0 = 5$ la curva continua está arriba de la curva segmentada. Así, vemos claramente que se cumple lo que establece el Lema de Neyman–Pearson para valores p .

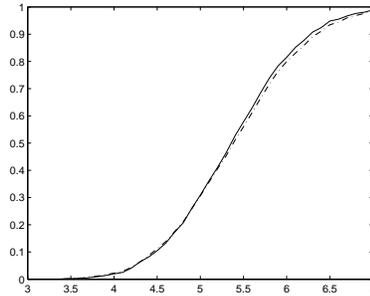


FIGURA 1. Gráficas de $P_\alpha(C(\alpha_0; T) \leq u)$ y $P_\alpha(C(\alpha_0; S) \leq u)$, $u = 0.3$.

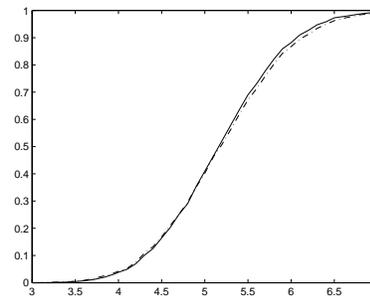


FIGURA 2. Gráficas de $P_\alpha(C(\alpha_0; T) \leq u)$ y $P_\alpha(C(\alpha_0; S) \leq u)$, $u = 0.4$.

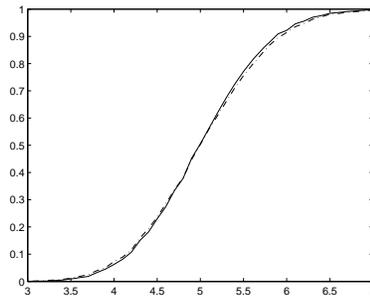
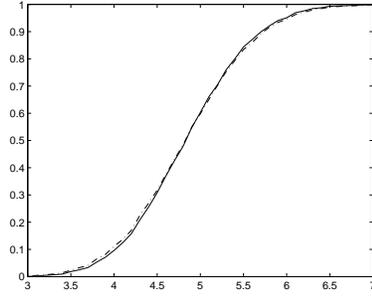


FIGURA 3. Gráficas de $P_\alpha(C(\alpha_0; T) \leq u)$ y $P_\alpha(C(\alpha_0; S) \leq u)$, $u = 0.5$.

Del algoritmo 1 se tiene el siguiente algoritmo para obtener una estimación de la distribución de confianza de α .

FIGURA 4. Gráficas de $P_\alpha(C(\alpha_0; T) \leq u)$ y $P_\alpha(C(\alpha_0; S) \leq u)$, $u = 0.6$.**Algoritmo 2**

1. Dadas las observaciones x_1, \dots, x_n de la muestra aleatoria X_1, \dots, X_n , sea

$$t_0 = \prod_{i=1}^n x_i.$$

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, por ejemplo, α_{inf} es el alfa tal que el área bajo la densidad de confianza a la izquierda de α_{inf} es un valor predeterminado, por ejemplo 0.01, y α_{sup} es el alfa tal que el área bajo la densidad de confianza a la derecha de α_{sup} es un valor predeterminado, por ejemplo 0.01.
3. Obtener estimaciones de $C(\alpha_i; t_0)$, $i = 0, \dots, r$.
4. Obtener una estimación de la densidad de confianza $c(\alpha; t_0)$.

En el siguiente ejemplo ilustramos numéricamente el lema de Neyman-Pearson para distribuciones de confianza (teorema 3).

Ejemplo 6. Sea la estadística $S = X_1 + \dots + X_n$. Por el ejemplo 1, sabemos que la densidad de confianza de α basada en S , una vez que se observó s , es

$$c(\alpha; s) = -(d/d\alpha)F(s; n\alpha) = n \int_0^s (\psi(n\alpha) - \log(t/\beta))f(t; n\alpha)dt. \quad (6)$$

Aunque la integral en (6) no la podemos resolver en forma cerrada, la podemos aproximar numéricamente. Para $n = 5, 10, 15$ y 30 se simularon muestras x_1, \dots, x_n de la muestra aleatoria X_1, \dots, X_n de la distribución $\text{gama}(\alpha = 3, \beta = 5)$. Para cada tamaño de muestra por medio de (6) se estimaron (asumiendo $\beta = 5$) las densidades de confianza de α basadas en el estadístico S ; la gráfica de cada una de ellas es la punteada en las figuras 5, 6, 7 y 8, respectivamente. Asimismo, por medio del algoritmo 2 se estimaron (también suponiendo $\beta = 5$) las densidades de confianza de α basadas en el estadístico suficiente T ; la gráfica de cada una de ellas es la continua en las mismas figuras. En éstas observamos que las gráficas de las densidades de confianza de α basadas en la estadística S , que en este caso no es suficiente, son más dispersas que las gráficas de las

densidades de confianza de α basadas en el estadístico suficiente T , de acuerdo a como lo establece el teorema de Neyman–Pearson para distribuciones de confianza.

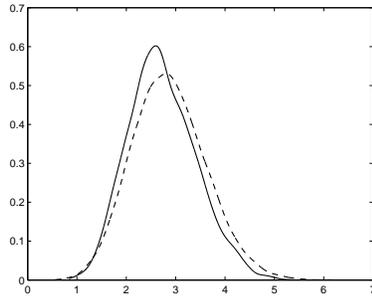


FIGURA 5. Gráficas de $c(\alpha; s)$ y $c(\alpha; t)$: $\alpha = 3$, $\beta = 5$, $n = 5$

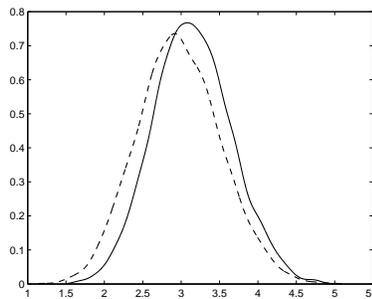


FIGURA 6. Gráficas de $c(\alpha; s)$ y $c(\alpha; t)$: $\alpha = 3$, $\beta = 5$, $n = 10$

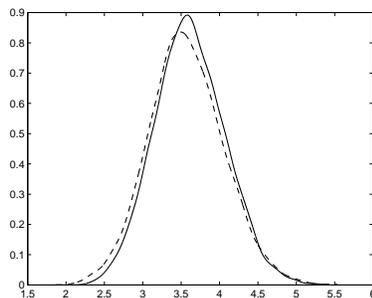
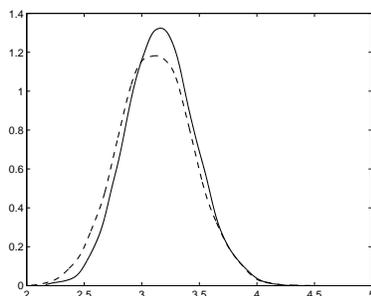


FIGURA 7. Gráficas de $c(\alpha; s)$ y $c(\alpha; t)$: $\alpha = 3$, $\beta = 5$, $n = 15$

FIGURA 8. Gráficas de $c(\alpha; s)$ y $c(\alpha; t)$: $\alpha = 3$, $\beta = 5$, $n = 30$

5. Conclusiones

Las distribuciones de confianza son muy importantes porque nos permiten ver de manera unificada y clara la estrecha relación que hay entre las pruebas de hipótesis y los intervalos de confianza, hecho que comúnmente queda poco claro cuando se utiliza el procedimiento frecuentista tradicional. Además, a través de ellas es directo y claro llevar a cabo pruebas de hipótesis unilaterales y obtener intervalos de confianza. En este contexto, las generalizaciones aquí consideradas del lema de Neyman-Pearson obtenidas por SCHWEDER y HJORT cobran relevancia, porque la primera nos dice cómo llevar a cabo las mejores pruebas de hipótesis unilaterales, y la segunda cómo obtener los mejores intervalos de confianza, tareas para las cuales proporcionamos un algoritmo. Incluimos además varios ejemplos numéricos que ilustran con claridad los resultados referidos.

Referencias

- [1] ASH, R. B. *Real analysis and probability*, Academic Press, New York, 2002.
- [2] CASELLA, G. & BERGER, R. L. *Statistical inference*, Duxbury, USA, 2002.
- [3] CLARKE, L. E. *Random variables*, Longman Group Limited, USA, 1975.
- [4] DUDLEY, R. M. *Real analysis and probability*, Wadsworth and Brooks/Cole, USA, 1989.
- [5] EFRON, B. *R. A. Fisher in the 21st century*, *Statistical Science*, **13** (1998), 95–122.
- [6] FISHER, R. A. *Statistical methods and scientific inference*, Halfner, London, 1973.
- [7] LEHMANN, E. L. *Testing statistical hypotheses*, Wiley, New York, 1959.
- [8] LEHMANN, E. L. *The Fisher, Neyman-Pearson theories of testing hypotheses: one theory of two?*, *Journal of the American Statistical Association*, **88** (1993), 1242–1249.
- [9] SCHWEDER, T. & HJORT, N. L. *Confidence and likelihood*, *Scandinavian Journal of Statistics*, **29** (2002), 309–332.

(Recibido en agosto de 2013. Aceptado para publicación en diciembre de 2013)

EDILBERTO NÁJERA RANGEL
DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS
UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

C.P. 86690, CUNDUACÁN, TABASCO, MÉXICO

e-mail: edilberto.najera@ujat.mx

BRALY GUADALUPE PERALTA REYES

INSTITUTO TECNOLÓGICO SUPERIOR DE COMALCALCO

C.P. 86650, COMALCALCO, TABASCO, MÉXICO

e-mail: braly.peralta@campus.itsc.edu.mx

AROLDO PÉREZ PÉREZ

DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS

UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

C.P. 86690, CUNDUACÁN, TABASCO, MÉXICO

e-mail: aroldo.perez@ujat.mx