



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA ESTADÍSTICA

PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN PARETO, BASADA EN LA INFORMACIÓN DE KULLBACK-LEIBLER

ANA CELIA LUQUE GUERRERO

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

DOCTOR EN CIENCIAS

MONTECILLO, TEXCOCO, ESTADO DE MÉXICO

2007

RESUMEN

PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN PARETO, BASADA EN LA INFORMACIÓN DE KULLBACK-LEIBLER

Ana Celia Luque Guerrero, Dr.

Colegio de Postgraduados, 2007

En este trabajo se presenta una prueba de bondad de ajuste para la distribución Pareto basada en la información de discriminación de Kullback-Leibler (1951) propuesta por Sheng Song (2002). Considerando que la distribución Pareto, no presenta el parámetro de escala, se aplicó la transformación logaritmo a esta distribución obteniendo como resultado la distribución Exponencial de dos parámetros, la cual es una distribución de localización y escala (Lehman y Casella, 1998). Se comprobó que la estadística de prueba de Kullback-Leibler es invariante. Se aplicó la metodología que propone Song, la cual se basa en los espacios m 'ésimos entre estadísticas de orden validados por la log verosimilitud. El cálculo de m se obtuvo a través de Simulación Monte Carlo . También mediante un experimento de Simulación Monte Carlo se calculó el poder de la prueba de Kullback-Leibler propuesta, la cual se comparó con el poder de la prueba de bondad de ajuste de Kolmogorov-Smirnov considerando las distribuciones, Lognormal, Weibull y Gamma como alternativas a la distribución Pareto. Con respecto al tamaño de la prueba, los resultados obtenidos por medio de simulación son muy parecidos a los valores α que se consideraron, confirmando que la prueba trabaja adecuadamente. Como conclusión más importante, se observó que la prueba de Kullback-Leibler resultó ser más poderosa que la prueba de Kolmogorov-Smirnov. Todos los cálculos se hicieron en lenguaje R.

Palabras Clave: Kullback-Leibler, Distribución Pareto, Simulación Monte Carlo

ABSTRACT

GOODNESS-OF-FIT TEST FOR PARETO DISTRIBUTION, BASED ON KULLBACK-LEIBLER INFORMATION

Ana Celia Luque Guerrero, Dr.

Colegio de Postgraduados, 2007

This paper presents a goodness-of-fit test for Pareto distribution based on the Kullback-Leibler (1951) discrimination information, as proposed by Sheng Song (2002). By considering the fact that Pareto distribution does not present the scale parameter, the logarithm transformation was used. As a result, an exponential distribution with location parameter was obtained, which is a localization and scale distribution (Lehman and Casella, 1998). It was proven that Kullback-Leibler test statistic is invariant. The methodology proposed by Song was used, which is based on m th-order spaces among order statistics validated by loglikelihood. The calculus of m was obtained through Monte Carlo Simulation. Additionally, the power of the proposed Kullback-Leibler test was also estimated. This was then compared to the Kolmogorov-Smirnov goodness-of-fit test by considering the Lognormal, Weibull and Gamma distributions as alternatives to Pareto distribution. With regard to the size of the test, the results obtained are quite similar to the α values considered. This confirmed the adequate function of the test. It was observed, as the most remarkable conclusion, that the Kullback-Leibler test was more powerful than the Kolmogorov-Smirnov one.

Key Words: Kullback-Leibler, Pareto Distribution, Monte Carlo Simulation

DEDICATORIA

A MIS PADRES

BLAS LUQUE Y CELIA GUERRERO, quienes con su gran ejemplo y apoyo me impulsaron constantemente con mucho amor, motivando mi superación.

Papi, sintiendo siempre tu presencia y tu protección desde cualquier lugar que te encuentres. Con mis mejores recuerdos para ti...- Siempre conmigo, siempre contigo-

A MI ESPOSO

LOMBARDO, mi compañero y amigo. Afortunadamente tuve y tengo la dicha de recorrer este camino y lo hago contigo, compartiendo momentos difíciles y muchos otros más, felices y gratificantes.

A MIS HIJAS

ARIADNA Y ANETT, mis pequeñas que siempre han sido un ejemplo de inteligencia y decisión. Gracias por acompañarme nuevamente en esta experiencia tan intensa que decidí vivir con el apoyo y comprensión de ustedes. Mi gran amor y admiración para ustedes.

AGRADECIMIENTOS

Alcanzar el objetivo de obtener un Doctorado en Ciencias es el mejor momento para dar las Gracias a todas las personas que me apoyaron y me ayudaron a través de estos años satisfactorios, pero también difíciles.

En primer lugar, expreso mi agradecimiento y reconocimiento al **Dr. Humberto Vaquera Huerta**, por su claridad, interés y entusiasmo, que orientó y motivó en todo momento mi proceso formativo; desde mi ingreso al doctorado hasta la dirección de esta tesis; investigación que presento con gran satisfacción y que complementa de manera importante mi formación como Estadística.

Agradezco de manera sensible a uno de los profesores que más he admirado desde que conocí el mundo de la Estadística, al **Dr. Ignacio Méndez Ramírez**, quien junto con su esposa **Lupita** me han brindado el privilegio de su amistad, e innumerables atenciones que motivaron mi pasión por el conocimiento estadístico.

A mis asesores **Dr. Gilberto Rendón Sánchez, Dr. Vicente González Romero y Dr. Enrique Villa Diharce**, por compartir conmigo su experiencia, sabiduría y fraterna compañía.

A mis sinodales **Dr. Sergio Pérez Elizalde y Dr. David del Valle Paniagua**, de quienes siempre recibí respaldo y atenciones como estudiante, y en la última parte de mi tesis, por sus sugerencias, de gran valor.

A todos mis profesores del Instituto de Socioeconomía, Estadística e Informática Estadística del Colegio de Postgraduados, por su gran apoyo y enseñanza comprometida.

A todo el personal del Instituto de Socioeconomía, Estadística e Informática Estadística del Colegio de Postgraduados por su gran comprensión y sincera amistad.

A las autoridades del Instituto Politécnico Nacional en especial al **Dr. Enrique Villa Rivera**, Director General del I.P.N., visionario que en todo momento impulsa y alienta la superación académica de todo el personal del Instituto.

Al **Dr. Luis Humberto Fabila Castillo**, Secretario de Investigación y Postgrado del I.P.N, por su respaldo y motivación para seguir adelante, hasta culminar exitosamente esta etapa.

Al **Dr. Adrián Luis García García**, ex Director del CICATA, Unidad Querétaro, quien alentó y apoyo en todo momento mi proceso formativo.

Al **Dr. Joaquín Salas Rodríguez**, Director del CICATA, Unidad Querétaro quien refrendó de manera comprometida y comprensiva, el interés institucional, para que pudiese alcanzar mi gran objetivo.

A mis compañeros del CICATA Unidad Querétaro, por su amistad y apoyo solidario.

Al Comité Técnico de Prestaciones a Becarios del IPN (**COTEPABE**), a la Comisión de Operación y Fomento de Actividades Académicas del IPN (**COFAA**), y al **CONACyT**, por el importante apoyo complementario que recibí, a través de permisos, beca y reconocimiento de mis estudios.

A mis compañeros de estudios tanto de Maestría como de Doctorado de Estadística, con quienes compartí muchos momentos especiales y emotivos. Mi especial agradecimiento a **Paulino Pérez Rodríguez** y a **Hortensia Reyes Cervantes**.

A todos los amigos con quienes tuve oportunidad de estudiar y convivir durante este tiempo. Gracias

A mis hermanos, cuñados, sobrinos y a Jair mi futuro yerno, quienes desde el inicio de mis estudios me alentaron e impulsaron hasta lograr mi objetivo.

CONTENIDO

1. INTRODUCCIÓN	1
2. OBJETIVOS	4
3. MARCO TEÓRICO	5
3.1 DISTRIBUCIÓN PARETO	5
3.1.1 REGLA 80/20	5
3.1.2 DEFINICIÓN DE LA DISTRIBUCIÓN PARETO	6
3.1.3 ESTIMADORES DE MÁXIMA VEROSIMILITUD DE LA DISTRIBUCIÓN PARETO	8
3.1.4 OTRAS APLICACIONES DE LA DISTRIBUCIÓN PARETO	9
3.2 CONCEPTOS DE LA TEORÍA DE LA INFORMACIÓN	9
3.2.1 ENTROPÍA.....	10
3.2.2 MEDIDAS DE ENTROPÍA.....	10
3.2.3 UNIDADES DE ENTROPÍA.....	11
3.2.4 ENTROPÍA DE SHANNON PARA UNA VARIABLE CONTINUA.....	12
3.2.5 ESTIMADOR NO PARAMÉTRICO DE ENTROPÍA	12
3.3 DISCREPANCIA INTRÍNSECA.....	13
3.3.1 MEDIDA DE DIVERGENCIA. DIVERGENCIA DE KULLBACK-LEIBLER.....	13
3.3.2 LA ENTROPÍA Y LA DIVERGENCIA DE KULLBACK-LEIBLER	14
3.4 PODER O POTENCIA DE UNA PRUEBA.....	14
3.5 PRUEBA DE HIPÓTESIS Y TAMAÑO DE LA PRUEBA.....	15
4. PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN PARETO, BASADA EN LA INFORMACIÓN DE KULLBACK-LEIBLER.....	17
4.1 PLANTEAMIENTO DEL PROBLEMA	17
4.2 TRANSFORMACIÓN DE LA DISTRIBUCIÓN PARETO	17
4.3 CONSTRUCCIÓN DE LA PRUEBA	18
4.4 IMPLEMENTACIÓN DE LA PRUEBA	19
5. ESTIMACIÓN DEL PODER DE LA PRUEBA.....	21
5.1 DISTRIBUCIÓN EMPÍRICA DEL ESTADÍSTICO DE PRUEBA DE KULLBACK-LEIBLER	21
5.2 ESTIMACIÓN DEL PODER DE LA PRUEBA.	22
5.3 PRUEBA DE KULLBACK-LEIBLER CON LAS DISTRIBUCIONES ALTERNATIVAS	22
5.3.1 VALORES CRÍTICOS	23
5.4 PRUEBA DE KOLMOGOROV-SMIRNOV.....	25
6. EJEMPLOS DE APLICACIÓN DE LA PRUEBA DE KULLBACK-LEIBLER	27
7. DISCUSIÓN DE LOS RESULTADOS.....	28
8. CONCLUSIONES	30
9. LITERATURA CITADA	31
ANEXOS	34
ANEXO A PROGRAMA EN R PARA REALIZAR LA PRUEBA DE BONDAD DE AJUSTE	34
ANEXO B TABLAS VALORES CRÍTICOS $c_{MN, \alpha}$ DE LA ESTADÍSTICA KL_{MN}	44
ANEXO C TABLAS DE LA ESTIMACIÓN DE LA POTENCIA.....	48
ANEXO D TABLAS DE LA ESTIMACIÓN DEL TAMAÑO DE LA PRUEBA	57

LISTA DE FIGURAS

FIGURA 1. FORMA DE LA DISTRIBUCIÓN DE LA ESTADÍSTICA KL_{mn}	21
--	----

LISTA DE TABLAS

TABLA 1. VALORES CRÍTICOS $C_{mn, \alpha}$ DE LA ESTADÍSTICA KL_{mn} . TABLA RESUMIDA	24
TABLA 2. VALORES CRÍTICOS D DE LA ESTADÍSTICA DE KOLMOGOROV-SMIRNOV	26
TABLA 3. TIEMPOS DE FALLA.....	27
TABLA 4. VALORES CRÍTICOS $C_{mn, \alpha}$ DE LA ESTADÍSTICA KL_{mn} , $\alpha = 0.01$	43
TABLA 5. VALORES CRÍTICOS $C_{mn, \alpha}$ DE LA ESTADÍSTICA KL_{mn} , $\alpha = 0.025$	44
TABLA 6. VALORES CRÍTICOS $C_{mn, \alpha}$ DE LA ESTADÍSTICA KL_{mn} , $\alpha = 0.05$	45
TABLA 7. VALORES CRÍTICOS $C_{mn, \alpha}$ DE LA ESTADÍSTICA KL_{mn} , $\alpha = 0.10$	46
TABLA 8. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=10, $\alpha = 0.10$	47
TABLA 9. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=10, $\alpha = 0.05$	47
TABLA 10. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=10, $\alpha = 0.01$	48
TABLA 11. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=30, $\alpha = 0.10$	48
TABLA 12. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=30, $\alpha = 0.05$	49
TABLA 13. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=30, $\alpha = 0.01$	49
TABLA 14. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=50, $\alpha = 0.10$	50
TABLA 15. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=50, $\alpha = 0.05$	50
TABLA 16. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=50, $\alpha = 0.01$	51
TABLA 17. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=100, $\alpha = 0.10$	51
TABLA 18. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=100, $\alpha = 0.05$	52
TABLA 19. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=100, $\alpha = 0.01$	52
TABLA 20. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=200, $\alpha = 0.10$	53
TABLA 21. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=200, $\alpha = 0.05$	53
TABLA 22. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=200, $\alpha = 0.01$	54
TABLA 23. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=225, $\alpha = 0.10$	54
TABLA 24. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=225, $\alpha = 0.05$	55
TABLA 25. POTENCIA DE KOLMOGOROV-SMIRNOV Y KULLBACK-LEIBLER PARA N=225, $\alpha = 0.01$	55
TABLA 26. ESTIMACIÓN DEL TAMAÑO DE LA PRUEBA DE BONDAD DE AJUSTE DE KULLBACK-LEIBLER	56
TABLA 27. ESTIMACIÓN DEL TAMAÑO DE LA PRUEBA DE BONDAD DE AJUSTE DE KULLBACK-LEIBLER	56

1. INTRODUCCIÓN

Desde el punto de vista estadístico, probar si las especificaciones de un modelo son las adecuadas es extremadamente importante. La validez de los procedimientos inferenciales dependen en gran parte de la suposición de que una distribución de datos en particular, sea la correcta.

Se han desarrollado muchos procedimientos para determinar si una variable aleatoria pertenece a una distribución específica. Todos estos procedimientos se conocen como “pruebas de bondad de ajuste”.

En este trabajo se presenta el caso de una prueba de bondad de ajuste para la distribución Pareto, basada en la información de discriminación de Kullback-Leibler (1951), propuesta por Song (2002).

Para la distribución Pareto se han publicado varios estudios de bondad de ajuste, entre ellos se tienen los de Porter III y cols. (1992), quienes utilizaron pruebas modificadas de Kolmogorov-Smirnov, Anderson Darling y Cramer-von Mises. También Beirlant, de Wet y Goegebeur (2005, 2006), aplicaron pruebas de bondad de ajuste para la distribución Pareto, modificando la estadística de Jackson originalmente propuesta para probar exponencialidad.

Estudios relacionados con el uso de la información de Kullback-Leibler en pruebas de bondad de ajuste, los realizaron Arizono y Ohta (1989), aplicándola s en una prueba de normalidad, Ebrahimi *et al* (1992,) y Ebrahimi (1998) publicaron resultados sobre pruebas de exponencialidad de vida residual, también basándose en esta información de discriminación de Kullback-Leibler, Senoglu y Sürücü (2004), hicieron pruebas de bondad de ajuste para probar normalidad, exponencialidad y uniformidad. Utilizando la misma prueba, Pérez, Vaquera y Villaseñor (2005), la aplican con la distribución Gumbel y Park (2005), aplicó la prueba de Kullback-Leibler para exponencialidad considerando datos censurados tipo II.

Las áreas de aplicación de la distribución Pareto, incluyen la Economía, en donde se utiliza para calcular riesgos y obtener información de valores extremos mínimos en finanzas y seguros, en esta especialidad se le conoce con el nombre de Distribución Bradford, se aplica también en investigación de operaciones (Porter III and et al, 1992)

Lawless (1982), Meeker & Escobar (1998), Wu & Chang (2003) y Balakrishnan y cols. (2004) la han considerado en estudios de tiempos de vida en experimentos con censura Tipo II. Este tipo de censura, se refiere al caso en el que de un total de n unidades que se ponen a prueba, ésta finaliza al tiempo en el que la m -ésima unidad falle ($1 \leq m \leq n$). De esta forma, la muestra censurada Tipo II, tiene solo las m observaciones más pequeñas en una muestra aleatoria de n unidades.

Si la distribución Pareto se utiliza en términos de función de sobrevivencia ($1-F$), se aplica de esta forma en disciplinas como confiabilidad, telecomunicaciones, finanzas, seguros, estudios ambientales, geológicos y climatológicos.

Se tienen algunas referencias relacionadas con la distribución Pareto en estudios en confiabilidad, como el de Ouyang y Wu (1994) quienes publicaron resultados sobre intervalos predicción para observaciones ordenadas en estudios de pruebas de vida; también el trabajo de Wu y Chang, (2003) trata sobre inferencia de la distribución Pareto, basada en censura progresiva tipo II con retiros aleatorios. D. Hanagal (1996a, 1996b) publicó diferentes estudios de estimación en sistemas de confiabilidad bajo la distribución Pareto bivariada y otros resultados con la distribución Pareto multivariada.

En el tema de estimación, se tiene la publicación de K. Vannman (1976), en donde presenta estimadores basados en las estadísticas de orden de una distribución Pareto. B. Arnold (1983,1989) también ha estudiado desde el punto de vista bayesiano esta distribución, realizando estudios de estimación y predicción bayesiana, así como inferencia para poblaciones con datos Pareto.

En la sección 4 se presenta el procedimiento de prueba de bondad de ajuste de Kullback-Leibler, sin embargo, considerando que la distribución Pareto, no presenta el parámetro de escala, se requirió aplicar la transformación logaritmo a la distribución Pareto, obteniendo como resultado la distribución exponencial recorrida, la cual es una distribución de localización y escala.

Con esta transformación, se aplicó la metodología metodología propuesta por Song (2002), en la que desarrolla una prueba de bondad de ajuste con la información de discriminación de Kullback-Leibler. Se calcularon los valores críticos de $C_{mn}(\alpha)$ de la estadística KL_{mn} y considerando que su distribución no depende de los parámetros de localidad y escala, se aplicó la simulación Monte Carlo. Se consideraron los niveles de significancia $\alpha = 0.1, 0.025, 0.05$ y 0.01 , tamaños de muestra de $n = 20$ a 200 con separación de diez unidades entre un cálculo y otro y se generaron $B = 5,000$ muestras aleatorias de tamaño n , posteriormente, se calculó la estadística de prueba KL_{mn} para cada $m < n/2$. El valor de $C_{mn}(\alpha)$ para cada m y n se determinó con el cuantil $(1-\alpha) \times 100$ de la distribución empírica de KL_{mn} .

En la sección 5 mediante simulación Monte Carlo se comparó el poder de las pruebas de Kullback-Leibler propuesta y de Kolmogorov-Smirnov, se utilizaron las distribuciones Gamma, Weibull y lognormal, como alternativas, y las tres pruebas de bondad de ajuste se realizaron con los niveles de significancia de $\alpha = 0.1, 0.05$ y 0.01 , tamaños de muestra de $n = 10, 30, 50, 100, 200$ y 225 y se generaron $B = 5,000$ muestras aleatorias para cada combinación de α y n , posteriormente, se calculó el valor de la estadística correspondiente para cada una de las pruebas.

El tamaño de prueba, resultó conveniente para diferentes tamaños de muestra, a diferentes niveles de confianza.

2. OBJETIVOS

2.1 Objetivo General

- ∞ Obtener una prueba de bondad de ajuste para la distribución Pareto, utilizando la distancia de Kullback-Leibler.

2.2 Objetivos Particulares

- ∞ Obtener los valores críticos de bondad de ajuste para la distribución Pareto considerando diferentes niveles de significancia y diferentes tamaños de muestra obteniendo la tabla de valores críticos.
- ∞ Conocer el comportamiento de la prueba estudiada con respecto a su tamaño.
- ∞ Hacer la comparación de potencia y tamaño de la prueba de bondad de ajuste utilizando la distancia de Kullback-Leibler con la de Kolmogorov Smirnov, aplicándolas con las distribuciones alternativas, Weibull, Gamma, lognormal.



VILFREDO PARETO
July 15, 1848-August 19, 1923

3. MARCO TEÓRICO

3.1 Distribución PARETO

El nombre de ésta distribución es dado en honor al economista y sociólogo italiano, Vilfredo Pareto (1848-1923), quien fue uno de los líderes de la llamada Escuela de Lausanne, fundada por León Walras en el siglo XIX. Antes de que Pareto cambiase su dirección política (hacia ideas del tipo de las de Sorel y Mussolini), en su "Cours d'Économie Politique" publicado en 1896 y 1897, presentó una exposición de la llamada Ley de Pareto de la distribución del ingreso, (Bouchaud, 2000)

Pareto, argumentó que en todos los países y épocas, la distribución de la riqueza sigue un patrón regular logarítmico capturado en imágenes como una "cola" - en cuyo extremo derecho, una pequeña fracción de la población es dueña de la mayoría de la riqueza. Por ejemplo, en los Estados Unidos, 300 mil personas (menos del 1,5%) posee el 10% de la riqueza. Bouchaud, (2000) indica que, la ley de Pareto, sólo describe la cola de la distribución correspondiente a grandes ingresos/riquezas como las que se encuentran empíricamente. Sin embargo este concepto no se puede aplicar sobre la totalidad de la distribución.

3.1.1 Regla 80/20

La regla mide el porcentaje de desequilibrio entre la entrada y la salida. El principio de Pareto no es una ley, sino una regla genérica que puede aplicarse

para muchos aspectos de la vida. Uno de los ejemplos en negocios más usado comúnmente es cuando el 80 por 100 de beneficio de una compañía es generado por el 20% de sus clientes.

Johnson & Kotz (1970), indican que la Ley de Pareto, se formuló para estudiar la distribución del ingreso en una población utilizando la siguiente expresión (conservando la notación original):

$$N = Ax^{-\alpha} \tag{3.1}$$

donde N es el número de personas que tienen un ingreso $\geq x$ y A, α son parámetros (α es el parámetro de forma y también es conocida como la constante de Pareto).

3.1.2 Definición de la distribución Pareto

Johnson & Kotz (1970), indican que la forma de la distribución Pareto es la siguiente:

$$P(x) = \Pr[X \geq x] = \left(\frac{k}{x}\right)^{-\eta} \quad k > 0, \quad \eta > 0, \quad x > k \tag{3.2}$$

Donde $P(x)$ es la probabilidad de que el ingreso sea igual o mayor a x y k representa el ingreso mínimo. Como una consecuencia de (3.1), la función de distribución acumulada de X , representando al “ingreso” puede escribirse como

$$F_x(x) = 1 - \left(\frac{k}{x}\right)^{-\eta} \quad k > 0, \quad \eta > 0, \quad x > k \tag{3.3}$$

Esta distribución es una forma especial Pearson Tipo VI

La relación dada por (3.2) es ahora mejor conocida como “Distribución Pareto de primera clase”.

Pareto propuso otras dos formas de la distribución. Una se refiere a la “Distribución Pareto de segunda clase” (en ocasiones llamada distribución *Lomax*) y está dada por

$$F_x(x) = 1 - \frac{K_1}{(x + C)^\eta} \tag{3.4}$$

(esta distribución también se conoce como de Pearson Tipo V)

La otra distribución propuesta por Pareto es la “Distribución Pareto de tercera clase” la cual tiene la siguiente función de distribución acumulada

$$F_x(x) = 1 - \frac{k_2 e^{-bx}}{(x+C)^\theta} \quad (3.5)$$

Considerando la función de distribución en (4.2), se tiene la correspondiente función de densidad Pareto

$$p_x(x) = \frac{k\eta^k}{x^{k+1}}; \quad k > 0, \quad x \geq \eta \geq 0, \quad (3.6)$$

donde x es cualquier número más grande que η , el cual es (necesariamente positivo), y k es un parámetro positivo.

La distribución Pareto también se le conoce como distribución Zeta o Ley de Zipf, como la contraparte discreta de la distribución Pareto.

El valor esperado de una variable aleatoria que sigue una distribución Pareto es

$$E(X) = \frac{k\eta}{k-1} \quad k > 1 \quad (3.7)$$

(como k es un parámetro positivo, si se tuviera $k \leq 1$, el valor esperado sería infinito).

Su desviación estándar es

$$\sigma = \frac{\eta}{k-1} \sqrt{\frac{k}{k-2}} \quad k > 2 \quad (3.8)$$

(siguiendo la misma condición del comentario anterior si $k \leq 2$, la desviación estándar sería infinita).

Los momentos originales se pueden calcular mediante la siguiente expresión

$$\mu'_n = \frac{k\eta^n}{k-n} \quad (3.9)$$

pero los momentos están definidos solo para $k > n$. Esto significa que la función generadora de momentos, la cual es una serie de Taylor, en x con $\frac{x^k}{k!}$ como coeficientes, no está definida.

La función característica está dada por

$$\varphi(t/k, \eta) = k(-i\eta t)^k \Gamma(-k, -i\eta t)$$

donde $\Gamma(a, x)$ es la función gama incompleta.

Otras características matemáticas de la distribución Pareto son:

Esta distribución Pareto está relacionada con la distribución exponencial de la forma siguiente

$$p(x/k, \eta) = f\left[\frac{\ln(x/\eta)}{k}\right]$$

Además también la función delta Dirac es un caso límite de la distribución Pareto

$$\lim_{k \rightarrow \infty} p(x/k, \eta) = \delta(x - \eta)$$

3.1.3 Estimadores de Máxima Verosimilitud de la distribución Pareto

Petersen (2000), indica que la función de verosimilitud, L , para la distribución Pareto con parámetros η y k , dada una muestra $x = (x_1, x_2, \dots, x_n)$ es

$$L(\theta, k/x) = \prod_{i=1}^n \frac{k\eta^k}{x_i^{k+1}}; \quad 0 < \theta \leq \min\{x_i\}, \quad k > 0$$

Obteniendo la función logarítmica de verosimilitud

$$\ell(k, \theta) = n \ln k + nk \ln \eta - (k+1) \sum_{i=1}^n \ln x_i$$

se puede ver que $\ell(k, \eta)$ aumenta monótonamente, con η , esto es, el valor más grande del valor de η , es el valor más grande de la función de verosimilitud, así $x \geq \eta$, por lo que

$$\hat{\eta} = \min_i x_i$$

Para encontrar el estimador para k , hacemos la correspondiente derivada parcial y se determina en donde es igual a cero

$$\frac{\partial \ell}{\partial k} = \frac{n}{k} + n \ln \eta - \sum_{i=1}^n \frac{1}{x_i} = 0$$

Así los estimadores de máxima verosimilitud para k y η son

$$\hat{r} = \min_i x_i, \quad \hat{k} = \frac{n}{\sum_i (\ln x_i - \ln \eta)}$$

3.1.4 Otras aplicaciones de la distribución Pareto

Se tienen algunas referencias relacionadas con la distribución Pareto en estudios en confiabilidad, como el de Ouyang y Wu (1994) quienes publicaron resultados sobre intervalos de predicción para observaciones ordenadas en estudios de pruebas de vida; también el trabajo de Wu y Chang en 2003, trata sobre inferencia de la distribución Pareto, basada en censura progresiva tipo II con retiros aleatorios. D. Hanagal (1996a, 1996b) publicó diferentes estudios de estimación en sistemas de confiabilidad bajo la distribución Pareto bivariada y otros resultados con la distribución Pareto multivariada.

En el tema de estimación, se tiene la publicación de K. Vannman (1976), en donde presenta estimadores basados en las estadísticas de orden de una distribución Pareto. B. Arnold (1983, 1989) también ha estudiado desde el punto de vista bayesiano esta distribución, realizando estudios de estimación y predicción bayesiana, así como inferencia para poblaciones con datos Pareto.

El uso de la distribución Pareto en estudios de confiabilidad es muy útil para modelar unidades de una población que tengan distribución exponencial. Estas unidades presentan una tasa de falla que varía de unidad en unidad de acuerdo a una distribución gama (θ, κ) , resultando que el tiempo de falla incondicional de una unidad seleccionada al azar de la población, tiene una distribución Pareto de la forma en (3.3) y con función de densidad dada en (3.6) Meeker & Escobar (1998).

3.2 Conceptos de la Teoría de la Información

Estos conceptos se utilizan en investigaciones de modelos neuronales y son muy útiles en el ámbito de la ingeniería, especialmente en el entorno de la teoría de comunicaciones, donde una de sus aplicaciones más comunes es la *compresión* de los datos. Los conceptos a los que se refiere esta teoría son, entre otros, *entropía*, *información mutua*, *entropía de Kullback-Leibler*, etc,

3.2.1 Entropía

El concepto básico es la entropía, definida como una medida de incertidumbre acerca del valor de una variable aleatoria x . Especificada por Shannon (1948) como $H(X)$, tiene la siguiente expresión formal:

$$\begin{aligned} H(X) &= -\sum_{x=1}^n p(x) \log[p(x)] \\ &= E \left[\log \left(\frac{1}{p(x)} \right) \right] \end{aligned}$$

donde, $E(\cdot)$ es el operador esperanza y $p(x)$ la distribución de probabilidad de la variable aleatoria. Si consideramos más de una variable, la entropía tendrá un carácter *conjunto* y si se condiciona su comportamiento, la entropía será de naturaleza *condicional*, ambos conceptos se sitúan en el ámbito bidimensional. Así, sean “ x ” e “ y ”, dos variables aleatorias de tipo discreto, donde $p(x, y)$ es la probabilidad conjunta y $p(y|x)$ es la probabilidad condicionada, y $H(x, y)$ se define como la entropía *conjunta*:

$$\begin{aligned} H(X, Y) &= -\sum_{x=1}^n \sum_{y=1}^m p(x, y) \log[p(x, y)] \\ &= E \left[\log \left(\frac{1}{p(x, y)} \right) \right] \end{aligned}$$

la cual es una medida de incertidumbre conjunta entre las dos variables aleatorias.

Y la entropía *condicional* $H(y|x)$ se define como

$$H(Y|X) = -\sum_{x=1}^n \sum_{y=1}^m p(x, y) \log[p(y|x)]$$

la cual es una medida del grado de incertidumbre de “ y ”, una vez que se conocen los valores concretos de “ x ”.

3.2.2 Medidas de Entropía

En el caso de variables aleatorias continuas se definen las diferentes entropías en una forma análoga.

La entropía fue introducida como medida cuantitativa de la información o, como medida de la incertidumbre sobre una experiencia cualquiera con un número finito de resultados posibles.

Si se considera un experimento aleatorio cuyos posibles resultados tienen sus probabilidades respectivas de ocurrir, se sabe que existe una “cierta” incertidumbre acerca del resultado particular que se presentará en caso de realizar el experimento.

Es necesario cuantificar la incertidumbre asociada a un experimento aleatorio. Para este fin, existen muchas medidas de incertidumbre, entre las más conocidas se tiene la entropía de Shannon. Es importante ver que la entropía de Shannon depende del número de resultados y de la probabilidad de ocurrencia de los mismos.

Otras medidas como la de Hartley, únicamente depende del número de resultados y no de la probabilidad de ocurrencia de los mismos, y se define como el logaritmo del número de resultados posibles del experimento.

El objetivo de las medidas de entropía es cuantificar la incertidumbre asociada a una variable aleatoria. (Pardo, 1997)

3.2.3 Unidades de Entropía

Al requerir medir la entropía, es necesario considerar en qué unidades se tiene que medir. Se considera que las unidades de entropía se dá en logaritmos, los cuales se pueden tomar con respecto a cualquier base que sea mayor que la unidad. Si se toma en base dos, la unidad correspondiente se denomina BIT (Binary digit) y puede definirse como la entropía correspondiente a una variable aleatoria con dos resultados equiprobables.

Si se toma en base 10, la unidad correspondiente se denomina DIT (Decimal digit o unidad de Hartley) y se define como la entropía de una variable aleatoria con diez resultados equiprobables.

Si se toman en base e , la unidad correspondiente se denomina NAT y se aplica en el concepto de entropía para variables aleatorias continuas. Cuando se toman los logaritmos en base natural se tiene la ventaja de que se simplifican los cálculos. En este caso se define el NAT, unidad de entropía en el caso continuo,

como la entropía correspondiente a una variable aleatoria con distribución uniforme en el intervalo $(0, e)$ (Pardo, 1997)

3.2.4 Entropía de Shannon para una variable aleatoria continua

Definición. Sea $X \equiv (X_1, \dots, X_n)$ una variable aleatoria n-dimensional continua con función de densidad $f(x)$, se denomina entropía de X a la expresión

$$h(X) = - \int_{R^n} f(x) \log f(x) dx$$

Supuesto que la integral existe.

La entropía de Shannon en el caso continuo verifica que la entropía $h(X)$ puede ser negativa.

3.2.5 Estimador no paramétrico de entropía

Para conocer la entropía de una función de densidad de probabilidades su distribución tiene que ser completamente especificada, pero muchos investigadores han desarrollado métodos no paramétricos para estimarla, como los resultados que presenta Park (1995) considerando la propuesta de Vasicek (1976), en la que presenta el estimador de la entropía, que la expresa $H(f)$ como:

$$H(f) = \int_0^1 \log \left\{ \frac{d}{dp} F^{-1}(p) \right\} dp \quad (3.10)$$

usando el hecho de que la pendiente $\frac{d}{dp} F^{-1}(p)$ se puede expresar en la forma:

$$\frac{d}{dp} F^{-1}(p) = \frac{1}{f(F^{-1}(p))} \quad (3.11)$$

De este modo para tener una idea del valor de la pendiente hay que estimar F con la función de distribución empírica F_n y reemplazar el operador de diferenciación por una diferencia entre dos cantidades. Esto conduce a un estimador muy simple de la pendiente que se obtiene como el producto de $\frac{n}{2m}$ y la diferencia entre dos cuantiles muestrales, con lo cual el estimador de la entropía está dado por:

$$H_{mn} = \frac{1}{n} \sum_{i=1}^n \log \left\{ \frac{n}{2m} (X_{(i+m)} - X_{(i-m)}) \right\} \quad (3.12)$$

Donde m es un entero positivo más pequeño que $n/2$, $X_{(j)} = X_{(1)}$

si $j < 1$, $X_{(j)} = X_{(n)}$ si $j > n$ y $X_{(1)} \leq \dots \leq X_{(n)}$ son las correspondientes estadísticas de orden, basadas en una muestra aleatoria de tamaño n .

3.3 Discrepancia intrínseca

Es importante considerar la definición original de Kullback (1959), relacionada con el concepto de discrepancia intrínseca la cual se presenta textualmente :

“Sea $S = \{p(x)\}$ el conjunto de todas las funciones de densidad regulares y absolutamente continuas entre sí en un espacio muestral χ con respecto a una medida dominante P .

$$K(q | p) = \int p(x) \log \frac{p(x)}{q(x)} dx$$

está definida, es no negativa y, para todo $p, q \in S$.

$K(q | p) = 0$, si y sólo si $p = q$, c.p.p.

En general, la divergencia de Kullback-Leibler no es simétrica, esto es

$$K(q | p) \neq K(p | q).$$

Además si los soportes de p y q , χ_p y χ_q , respectivamente, son tales que $\chi_q \subseteq \chi_p$, entonces $K(q | p)$ diverge”.

3.3.1 Medida de Divergencia. Divergencia de Kullback-Leibler

Una vez considerada la definición de 3.3, se puede definir que una medida de divergencia cuantifica la cantidad de información proporcionada por los datos, y mide el grado de discrepancia entre dos poblaciones caracterizadas por sus correspondientes distribuciones de probabilidad.

La divergencia de Kullback-Leibler o entropía relativa es una cantidad que mide la diferencia entre dos funciones de densidad de probabilidad.

La divergencia de Kullback-Leibler a favor de $g(x)$ contra $f(x)$ está definida por:

$$K(g; f) = \int_{-\infty}^{\infty} g(x) \log \frac{g(x)}{f(x)} dx \quad (3.13)$$

Como $K(g; f)$ tiene la propiedad de que $K(g; f) \geq 0$ y la igualdad se mantiene si $g = f$, el estimador de la información de Kullback-Leibler también ha sido considerada como una estadística de prueba de bondad de ajuste por Arizono (1989) y Ebrahimi (1992) entre otros.

3.3.2 La entropía y la divergencia de Kullback-Leibler en las pruebas de bondad de ajuste

Las pruebas de bondad de ajuste basadas en la entropía muestral o bien en divergencia de Kullback-Leibler son consistentes y presentan un buen desempeño. Vasicek (1976) propuso una prueba de normalidad, basada en el hecho de que la entropía de esta distribución excede a la de cualquier otra distribución que tenga la misma varianza. Arizono y Otha (1976) desarrollaron otra prueba para normalidad, pero basada en la divergencia de Kullback-Leibler, que puede ser aplicada para hipótesis simples y compuestas.

Ebrahimi *et al.* (1992) desarrollaron una prueba de exponencialidad basada en la divergencia de Kullback-Leibler. Song (2002) presentó una metodología general para desarrollar pruebas de bondad de ajuste con *distribución libre asintótica* basada en la divergencia de Kullback-Leibler. El procedimiento puede ser aplicado a una gama amplia de distribuciones, entre las que destacan las distribuciones siguientes: Normal, log-normal, logística, Gumbel, exponencial,

3.4 Poder o Potencia de la prueba

La potencia de una prueba es la probabilidad de rechazar la hipótesis nula H_0 , cuando esta es verdadera. La definición de potencia se usa principalmente para una hipótesis alternativa simple. Si la hipótesis alternativa es compuesta se emplea el término de Función Potencia, la cual se define como la función a la que un valor θ del parámetro se le asocia la probabilidad de rechazar la H_0 si este valor es el verdadero y su expresión es

$$\beta(\theta) = P_{\theta}[\text{Rechazar } H_0]$$

Si la hipótesis H_0 es simple del tipo $\theta = \theta_0$ entonces el valor de la función potencia para θ_0 es el *Error Tipo I* :

$$\beta(\theta_0) = \alpha$$

3.5 Prueba de hipótesis y Tamaño de la prueba

Cuando se usa una prueba φ , estamos sujetos a cometer cualquiera de los dos tipos de errores posibles

$$\varphi(x) = 1 \quad \text{Si } x > 0$$

$$\varphi(x) = 0 \quad \text{Si } x \leq 0$$

El criterio para encontrar la prueba óptima, es escoger la prueba φ^* tal que las probabilidades de ambos errores sean mínimas.

En general, cuando se minimiza la probabilidad del *Error Tipo I* se aumenta la probabilidad del *Error Tipo II* y viceversa.

Por esta razón fijamos un nivel de probabilidad α para la probabilidad del *Error Tipo I* y se trata de obtener la prueba que hace mínima la probabilidad del *Error Tipo II*. Es decir, entre todas las pruebas φ que satisfacen:

$$P(\text{Error Tipo I usando } \varphi) = \alpha < \alpha_1$$

$$0 < \alpha < 1 = \alpha \in (0,1)$$

Note que si φ es de tamaño α y $\alpha < \alpha_1$ entonces φ también es de tamaño α_1 .

Definición: Una prueba φ que satisface * es llamada una prueba de tamaño α .

El interés es que las pruebas, sean del tamaño lo más pequeño posible.

a) Probabilidad de rechazar H_0 cuando es verdadera

$$P(\text{Error Tipo I usando } \varphi) = P(\text{Rechazar } H_0 \text{ usando } \varphi \mid \theta \in w)$$

$$= P(\varphi(\underline{x}) = 1 \mid \theta \in w) \quad \text{De donde } \varphi \text{ es de tamaño } \alpha \text{ si}$$

$$\max_{\theta \in w} P(\varphi(\underline{x}) = 1 \mid \theta) \leq \alpha \quad **$$

b) Probabilidad de No rechazar H_0 cuando es falsa.

$$P(\text{Error Tipo II usando } \varphi) = P(\varphi(\underline{x}) = 0 \mid \theta \in \Omega - w) \quad ***$$

$$= 1 - P(\varphi(x) = 1 | \theta \in \Omega - w)$$

Entonces se desea encontrar la φ que satisface

$$\max_{\theta \in w} P(\varphi(x) = 1 | \theta) \leq \alpha \quad ** \quad \text{y minimiza}$$

$$P(\varphi(x) = 0 | \theta \in \Omega - w) \quad \text{o}$$

$$= 1 - P(\varphi(x) = 1 | \theta \in \Omega - w)$$

4. PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN PARETO, BASADA EN LA INFORMACIÓN DE KULLBACK-LEIBLER

4.1 Planteamiento del problema

Una variable aleatoria X , tiene función de distribución Pareto, $X \sim P(k, \eta)$ si su función de densidad es de la forma

$$f(x; \cdot) = f_0(x, k, \eta) = \frac{k\eta^k}{x^{k+1}} \quad k > 0, \quad x \geq \eta \geq 0 \quad (4.1)$$

donde η es el parámetro de localización y k es el parámetro de forma. Y \hat{k} y $\hat{\eta}$ son respectivamente los estimadores de máxima verosimilitud:

$$\hat{k} = \frac{n}{\sum_{i=1}^n (\ln y_i - n \ln \hat{\eta})} \quad \hat{\eta} = \min_{1 \leq i \leq n} \{y_i\}$$

Sean $\{X_1, \dots, X_n\}$ observaciones independientes de una distribución F , con función de densidad de probabilidad $f(x; \cdot)$, $x \in \mathfrak{X}$.

Se desea probar el siguiente juego de hipótesis

$$H_0: f(x; \cdot) = f_0(x, k, \eta) \quad (4.2)$$

contra la hipótesis alternativa:

$$H_1: f(x; \cdot) \neq f_0(x, k, \eta) \quad (4.3)$$

El objetivo de esta hipótesis es conocer si una muestra aleatoria proviene de la distribución Pareto.

4.2 Transformación de la distribución Pareto

Sin embargo, antes de probar la hipótesis de interés, se requiere aplicar una transformación logaritmo a la distribución Pareto para obtener una distribución Exponencial de dos parámetros (Lehman y Casella, (1998) pág. 486).

Considerando que Y se distribuye Exponencial de dos parámetros

$Y \sim \text{Exp}(\xi, b)$ con función de densidad

$$f_{(Y)}(y) = \frac{1}{b} \exp\left\{-\frac{(y-\xi)}{b}\right\} I_{(\xi, \infty)}(y) \quad (4.4)$$

donde $\xi = \ln \eta$ y $b = 1/k$

4.3 Construcción de la prueba

Ahora se va a considerar la información de discriminación de Kullback-Leibler entre dos funciones de distribución para probar la hipótesis de interés.

$$KL(F, F_0; \theta) = \int_{-\infty}^{\infty} f(y) \ln[f(y)/f_0(y, \theta)] dy \quad (4.5)$$

donde θ es un vector de parámetros que en este caso, contiene a η, k .

La evaluación de $KL(F, F_0; \theta)$ requiere conocer a F y F_0 , por lo que ahora se debe obtener un estimador muestral de $KL(F, F_0; \theta)$, considerando la hipótesis que se quiere probar. Para este fin se aplicará la propuesta hecha por Song (2002). Considerando la información de Kullback-Leibler dada en (4.8) y por propiedades de los logaritmos se tiene:

$$KL(F, F_0; \theta) = \int_{-\infty}^{\infty} f(y) \ln f(y) dy - \int_{-\infty}^{\infty} f(y) \ln f_0(y, \theta) dy \quad (4.6)$$

donde: $\int_{-\infty}^{\infty} f(y) \ln f(y) dy = -H(F)$ es la entropía de F .

Para obtener la estimación de la entropía, $H(F)$, se puede considerar el estimador propuesto por Vasicek (1976), dado por:

$$H_{mm} = \frac{1}{n} \sum_{i=1}^n \ln \left\{ \frac{n}{2m} (Y_{(i+m)} - Y_{(i-m)}) \right\} \quad (4.7)$$

donde m es un entero positivo menor que $n/2$, $\{Y_{(1)}, \dots, Y_{(n)}\}$, son las estadísticas de orden de la muestra $\{Y_1, \dots, Y_n\}$, $Y_j = Y_{(1)}$, si $j < 1$ y $Y_j = Y_{(n)}$ si $j > n$.

Para estimar de (4.5) la función $\int_{-\infty}^{\infty} f(y) \ln f_0(y, \theta) dy$ se utilizará la expresión propuesta por Song (2002) dada por:

$$n^{-1} \sum_{i=1}^n \ln f_0(Y_i, \hat{k}, \hat{\eta}) \quad (4.8)$$

sustituyendo (4.7) y (4.8) en (4.5) se tiene la ecuación completa para el estimador

KL_{mn} :

$$KL_{mn} = \frac{1}{n} \sum_{i=1}^n \ln \left\{ \frac{n}{2m} Y_{(i+m)} - Y_{(i-m)} \right\} - n^{-1} \sum_{i=1}^n \ln f_0(Y_i, \hat{k}, \hat{\eta})$$

teniendo como resultado la estadística de prueba para la distribución Exponencial de dos parámetros.

A partir de este resultado se generaliza a una familia de localidad y escala, la cual se consideró en este estudio para aplicarla en la prueba de bondad de ajuste de Kullback-Leibler.

4.4 Implementación de la prueba

Una vez que se aplica la prueba si se obtienen valores grandes del estimador KL_{mn} se rechaza H_0 a favor de la hipótesis alternativa, es decir; se rechaza la hipótesis nula si $KL_{mn} \geq C_{mn}(\alpha)$, así el valor de la constante crítica $C_{mn}(\alpha)$ se determina por el cuantil $(1 - \alpha)100$ de la distribución de KL_{mn} bajo la hipótesis nula.

Considerando que se tiene definido el tamaño de muestra n , ahora se tiene que especificar el parámetro m . Song (2002) sugiere que de acuerdo a su teoría, m debería escogerse de acuerdo al tamaño de muestra finito. Sin embargo en la práctica escoger el valor de m óptimo es problemático; ya que no sólo depende del tamaño de muestra, sino también de la alternativa en particular que se considere. Por lo que Song (2002), probó hipótesis compuestas de normalidad con parámetros no especificados, contra siete opciones de hipótesis alternativas obteniendo el poder de la prueba por medio de simulación MonteCarlo.

Estas simulaciones mostraron que aún para n fija, no existe una m que sea óptima para todas las alternativas consideradas.

O sea, si se tiene una distribución alternativa de especial interés, entonces la mejor forma de escoger m , podría ser la m que proporcione el poder más alto en la dirección de esta alternativa para el tamaño de muestra n dado y el nivel α requerido.

Para este fin, se tiene el hecho de que la información de Kullback-Leibler $KL((F, F_0; \theta)) \geq 0$ para todo $\theta \in \Theta$ en donde la igualdad se mantiene para algún $\theta \in \Theta$ si y solo si $f(x) = f_0(x, \theta)$. En otras palabras, bajo $KL((F, F_0; \theta)) = 0$, los valores grandes de $KL((F, F_0; \theta))$ favorecen la hipótesis alternativa sobre H_0 .

Sobre esta propiedad se basa el método para seleccionar m . Dadas las observaciones $\{X_i\}_{i=1}^n$, se estima $KL((F, F_0; \theta))$, con su estimador muestral, KL_{mn} . La idea básica es escoger m que minimice KL_{mn} .

La siguiente expresión nos presenta la forma de obtener m : $\hat{m} = \min \left\{ m^* : m^* = \arg \max \left\{ H_{mn} : H_{mn} \leq -\frac{1}{n} \sum_{i=1}^n \ln f_0(X_{(i)}, \hat{k}, \hat{\eta}) \right\} \right\}$ (4.10)

es decir, \hat{m} debe de ser el valor mas pequeño de \hat{m}^* , que maximice la entropía muestral H_{mn} restringida por la log verosimilitud observada.

5. ESTIMACIÓN DEL PODER DE LA PRUEBA

5.1 Distribución empírica del estadístico de prueba KL_{mn}

Aplicando simulación Monte Carlo se probó que para este caso particular la distribución empírica del estadístico de prueba KL_{mn} para la distribución Pareto, bajo H_0 , es prácticamente idéntica aún para tamaños de muestra pequeños y no depende de los parámetros k y η . Para obtener la distribución empírica de KL_{mn} se mantuvieron fijos n, m, \hat{k} y $\hat{\eta}$ y se generaron B muestras aleatorias de tamaño n para la distribución Pareto con parámetros k y η . Para calcular el valor de KL_{mn} se hicieron B realizaciones de la estadística de prueba y se generó la forma aproximada de la distribución, obteniendo el histograma de las diferentes gráficas

En la figura 1 se presenta la forma de la distribución de la estadística de prueba KL_{mn} con $n=10$, $m=4$ y $B=10,000$ para cada combinación de los diferentes valores de parámetros especificados en la gráfica. Los cálculos se hicieron con el programa en lenguaje R.

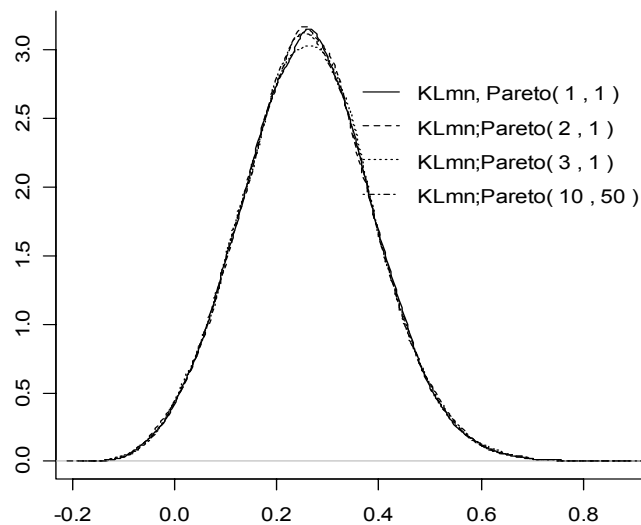


FIGURA 1 Distribución de la estadística de prueba de Kullback-Leibler

5.2 Estimación del poder de la prueba

En esta sección se presenta la obtención del poder de la prueba de manera comparativa para las pruebas de Kullback-Leibler y Kolmogorov Smirnov, considerando las distribuciones, Lognormal, Weibull y Gamma como alternativas a la distribución Pareto.

Se utilizaron para las tres pruebas de bondad de ajuste los niveles de significancia de $\alpha = 0.1, 0.05$ y 0.01 , tamaños de muestra de $n = 10, 30, 50, 100, 200$ y 225 y se generaron $B = 5,000$ muestras aleatorias para cada combinación de α y n , posteriormente, se calculó el valor de la estadística correspondiente para cada una de las pruebas.

5.3 Prueba de Kullback-Leibler con las distribuciones alternativas

Para obtener la estadística de prueba KL_{mn} de Kullback-Leibler, se calculó el estimador de entropía (H_{mn}). Este estimador considera las estadísticas de orden, el valor de n y el valor de m .

Se calculó el poder de la prueba para las 16 distribuciones alternativas con diferentes combinaciones de los niveles de significancia y tamaños de muestra antes mencionados. En cada caso las pruebas se realizaron con los parámetros que se presentan a continuación.

Distribución Weibull: (1,1), (1,3), (2,3), (2,1), (3,1) y, (3.439,1)

Distribución Gamma: (1,1), (1,3), (2,3), (2,1), (3,1) y, (3.439,1)

Distribución Log-normal: (1,1), (1,0.4), (5,3), (10,0.2)

Para la distribución Weibull, se realizaron las pruebas con diferentes valores de parámetros, estableciendo las siguientes correspondencias con otras distribuciones de acuerdo a Makino(1984),

Weibull(1,3) = Exp(3)

Weibull(2,3) = Raleigh(;)

Weibull(3.439,1) = Normal(;)

La distribución Weibull que se consideró, tiene parámetro de forma α y parámetro de escala β con función de densidad

$$f(x) = (\alpha / \beta)(x / \beta)^{(\alpha-1)\exp(-(x / \beta)^\alpha)} \quad \text{para } x > 0$$

La distribución Gamma que se utilizó tiene parámetro de forma θ y parámetro de escala β con función de densidad

$$f(x) = \frac{1}{\beta^\theta \Gamma(\theta)} x^{(\theta-1)} \exp^{-(x/\beta)} \quad \text{para } x > 0, \theta > 0 \text{ y } \beta > 0$$

Para la distribución Gamma; solamente por comparación, se aplicaron los mismos valores de los parámetros que para la distribución Weibull.

Se trabajó con la distribución Lognormal, con parámetro de localización μ y parámetro de forma σ^2 con función de densidad:

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \quad \forall x > 0$$

Tamborero del Pino y Cejalvo (1996) indican las siguientes características para la distribución Lognormal de acuerdo a los valores del parámetro de escala:

Lognormal (5,3) = Asimétrica, para valores medios y altos de σ

Lognormal (1,0.4) = Conforme σ decrece, la distribución es más simétrica

Lognormal (1,1) = Exponencial negativa. Si σ se acerca a la unidad

Lognormal (10,0.2) = Normal. Para valores de $\sigma < 0.2$

Para esta distribución se trabajó con los parámetros antes indicados.

Las potencias de las pruebas de Kullback-Leibler se realizaron con la estadística de prueba KL_{mn} , obtenidas por medio de simulación Monte Carlo utilizando el programa en lenguaje R :

5.3.1 Valores críticos

Para determinar los valores críticos de $C_{mn}(\alpha)$ y KL_{mn} , y considerando que su distribución no depende de los parámetros de localidad y escala, se aplicó la

simulación Monte Carlo. Se consideraron los niveles de significancia $\alpha = 0.1, 0.025, 0.05$ y 0.01 , tamaños de muestra de $n = 20$ a 200 con separación de diez unidades entre un cálculo y otro y se generaron $B = 5,000$ muestras aleatorias de tamaño n , posteriormente, se calculó la estadística de prueba KL_{mn} para cada $m < n/2$. El valor de $C_{mn}(\alpha)$ para cada m y n se determinó con el cuantil $(1-\alpha) \times 100$ de la distribución empírica de KL_{mn} .

En la tabla 1 se muestran de manera resumida los valores críticos $C_{mn}\alpha$ de la estadística KL_{mn} para la distribución Pareto, y en las tablas 2, 3, 4 y 5 (anexas) se presentan los valores críticos para la estadística KL_{mn} para los niveles de significancia, de $0.01, 0.025, 0.05$ y 0.1 respectivamente.

Nivel de significancia (α)								
n	0.01		0.025		0.05		0.1	
	Cmn	m	Cmn	m	Cmn	m	Cmn	m
20	0.3889	4	0.3440	4	0.3058	4	0.2690	4
30	0.2910	5	0.2573	5	0.2329	5	0.2079	5
40	0.2366	6	0.2127	6	0.1931	6	0.1707	6
50	0.2050	6	0.1851	6	0.1669	6	0.1492	6
60	0.1792	8	0.1603	7	0.1470	6	0.1324	7
70	0.1598	7	0.1443	7	0.1320	7	0.1190	7
80	0.1456	7	0.1322	8	0.1206	8	0.1091	8
90	0.1363	8	0.1226	9	0.1131	10	0.1016	10
100	0.1248	9	0.1142	9	0.1050	9	0.0950	8
110	0.1187	10	0.1067	10	0.0981	10	0.0887	10
120	0.1111	8	0.1010	11	0.0927	11	0.0834	11
130	0.1059	11	0.0952	9	0.0881	10	0.0796	11
140	0.1005	13	0.0921	9	0.0843	11	0.0760	12
150	0.0939	11	0.0864	11	0.0795	12	0.0719	12
160	0.0913	11	0.0834	13	0.0768	11	0.0696	13
170	0.0880	11	0.0800	12	0.0735	12	0.0666	12
180	0.0845	12	0.0772	12	0.0713	12	0.0641	14
190	0.0817	13	0.0748	15	0.0687	15	0.0618	15
200	0.0787	12	0.0714	12	0.0660	14	0.0594	14

Tabla 1. Valores críticos $C_{mn}\alpha$ de la estadística KL_{mn} para la distribución Pareto (Forma resumida)

5.4 Prueba de Kolmogorov-Smirnov

La prueba de Kolmogorov-Smirnov indica si un conjunto de observaciones provienen de alguna distribución continua, completamente especificada. Sin embargo una de las limitaciones para aplicar esta prueba, es cuando uno o más parámetros de la distribución deben estimarse a partir de la muestra, por lo que en estos casos no es muy conveniente aplicar la prueba, ya que no se pueden utilizar los puntos críticos tabulados comúnmente, (Lilliefors, 1969).

Considerando que la distribución Pareto, no presenta el parámetro de escala, por lo que no cumple con esta condición para aplicar la prueba de Kolmogorov-Smirnov. Sin embargo, este problema se resolvió aplicando la transformación logaritmo a la distribución Pareto, obteniendo como resultado la distribución exponencial recorrida, la cual es una distribución de localización y escala (Lehman y Casella, 1998) (explicada en el capítulo 2).

Lilliefors (1969), presenta tablas de la estadística de Kolmogorov-Smirnov, para probar si un conjunto de observaciones provienen de una población exponencial cuando la media no es especificada, y se estima a partir de la muestra.

En su trabajo, Lilliefors (1969), menciona que David y Johnson (1948), indican que si los parámetros estimados son parámetros de escala o localización y los estimadores satisfacen ciertas condiciones generales, entonces se aplica la transformación integral de probabilidad, y entonces, la distribución conjunta de las variables transformadas no dependerán del valor verdadero del parámetro. La distribución dependerá de la forma funcional de la distribución de las variables originales. Así se pueden construir las tablas con la estadística de Kolmogorov-Smirnov para esa distribución en particular.

El procedimiento que presenta Lilliefors (1969) es:

dada una muestra de n observaciones se determina

$$D = \text{Máximo} |F^*(X) - S_n(X)|$$

donde, $S_n(X)$ es la función de distribución acumulada de la muestra y $F^*(X)$ es la función de distribución acumulada exponencial con $1/\lambda = \bar{X}$ que es la media

muestral. Si el valor de D excede el valor crítico en la tabla, entonces se rechaza la hipótesis de que las observaciones vienen de una población exponencial.

La tabla 2 muestra los valores críticos de D

<i>Nivel de significancia para $D = \text{Máximo} F^*(X) - S_N(X)$</i>					
n	0.20	0.15	0.1	0.05	0.01
3	0.451	0.479	0.511	0.551	0.600
4	0.396	0.422	0.449	0.487	0.548
5	0.359	0.382	0.406	0.442	0.504
6	0.331	0.351	0.375	0.408	0.470
7	0.309	0.327	0.350	0.382	0.442
8	0.291	0.308	0.329	0.36	0.419
9	0.277	0.291	0.311	0.341	0.399
10	0.263	0.277	0.295	0.325	0.380
11	0.251	0.264	0.283	0.311	0.365
12	0.241	0.254	0.271	0.298	0.351
13	0.232	0.245	0.261	0.287	0.338
14	0.224	0.237	0.252	0.277	0.326
15	0.217	0.229	0.244	0.269	0.315
16	0.211	0.222	0.236	0.261	0.306
17	0.204	0.215	0.229	0.253	0.297
18	0.199	0.210	0.223	0.246	0.289
19	0.193	0.204	0.218	0.239	0.283
20	0.188	0.199	0.212	0.234	0.278
25	0.170	0.180	0.191	0.210	0.247
30	0.155	0.164	0.174	0.192	0.226
Más de 30	.86/ \sqrt{N}	.91/ \sqrt{N}	.96/ \sqrt{N}	1.06/ \sqrt{N}	1.25/ \sqrt{N}

Tabla 2. Valores críticos de D . Lilliefors (1969)

6. APLICACIÓN DE LA PRUEBA DE KULLBACK-LEIBLER EN UN EJEMPLO DE OBSERVACIONES DE TIEMPOS DE VIDA

Considerando los resultados de 20 tiempos de falla en un estudio de sobrevivencia, reportados por Ouyang y Wu (1994) en su trabajo sobre intervalos de predicción para observaciones ordenadas con distribución Pareto, se aplicaron los resultados obtenidos en este trabajo. El programa en lenguaje R para este ejemplo se anexa al final.

30.101	30.150	30.374	30.581	30.871	31.086	31.398	31.752	31.792	31.960
32.260	32.517	32.636	33.002	33.552	33.721	34.002	34.023	34.150	35.274

Tabla 3. Tiempos de falla

Se aplicó la prueba de Kullback-Leibler, para probar si los datos de tiempo de falla tienen distribución Pareto, contra la alternativa de que no se distribuyen Pareto.

En la prueba se consideró un tamaño de muestra $n = 20$, un valor de $m = 4$ y un nivel de significancia de 0.05

Los resultados obtenidos con el programa de cómputo R se anexan al final del trabajo.

Se considera la siguiente regla de decisión:

Se rechaza la hipótesis nula si, $KL_{4,20} \geq C_{4,20}(.05)$

El valor de $KL_{4,20} = 0.3514342$

Valor crítico $C_{4,20}(0.05) = 0.3572$

Como $KL_{4,20} < C_{4,20}(.05)$ entonces No se rechaza la hipótesis nula

El resultado de la prueba indica que los datos de tiempo de falla que se consideraron en el ejemplo tienen distribución Pareto con un nivel de significancia de 0.05

7. DISCUSIÓN DE LOS RESULTADOS

- Se aplicó la prueba de bondad de ajuste de Kullback-Leibler propuesta por Song (2000), para la distribución Pareto. Debido a que la distribución Pareto no es invariante, se le aplicó la transformación logaritmo obteniendo como resultado la distribución Exponencial de dos parámetros, la cual es una distribución de localización y escala (Lehman y Casella, 1998).
-
- Se obtuvo la tabla de valores críticos $C_{mn,\alpha}$, de la estadística KL_{mn} para la distribución Pareto, estos valores se obtuvieron por simulación Montecarlo a niveles de significancia, 0.01, 0.025, 0.05 y 0.1, tamaños de muestra n de 5 a 200 y para valores de m de 4 a 15. (Tabla 1)
- La prueba de bondad de ajuste de Kullback-Leibler, para la distribución Pareto, considerando como distribuciones alternativas la Weibull, Gamma, y Lognormal, observando en general que para las tres distribuciones, la potencia de la prueba, aumenta, conforme se incrementa el tamaño de la muestra.
- Para el caso de la distribución alternativa Weibull, se consideraron 5 combinaciones de parámetros. Para el caso de tamaño de muestra $n = 10$ a los diferentes niveles de significancia considerados (0.10, 0.05 y 0.1), el poder de la prueba de Kullback-Leibler, presentó valores más altos en la mayoría de los casos, que para la prueba de Kolmogorov-Smirnov. (Tablas 4, 5 y 6).
- Para esta misma distribución, la potencia de Kullback-Leibler, para $n=30$ fue mayor en todos los casos y para el tamaño de muestra $n=50$, $n=100$, $n=200$ y $n=225$, con α diferentes, la potencia fue igual a uno, viendo que para estos tres últimos tamaños de muestra, resultó idéntica a la potencia de la prueba de Kolmogorov-Smirnov. (Tablas de la 7 a la 21)

- Para la distribución Gamma, con tamaño de muestra igual a 10, el resultado del poder de la prueba con los diferentes niveles de significancia considerados (0.01, 0.05 y 0.1), se presentó de manera muy semejante a la distribución Weibull, siendo mayor en prácticamente todos los casos que para la prueba de Kolmogorov- Smirnov. (Ver Tablas 4 , 5 y 6)
- Para esta misma distribución, la potencia de Kullback-Leibler, para $n=30$ fue mayor en los 16 casos. Para los tamaños de muestra; $n=50$, $n=100$, $n=200$ y $n=225$, la potencia fue igual a uno, viendo que para los tres últimos tamaños de muestra, resultó idéntica a la potencia de la prueba de Kolmogorov- Smirnov. (Tablas de la 7 a la 21)
- Tamborero del Pino y Cejalvo (1996) indican que para la distribución Lognormal, de acuerdo a los valores del parámetro de escala σ , si este tiene valores medios y altos entonces la distribución es asimétrica, en este trabajo se probó la hipótesis alterna de Lognormal (5,3) y en las tablas 8, 9 y 10 se observa que para esta combinación de parámetros al tamaño de muestra más bajo ($n=10$) y para los tres niveles de significancia (0.10, 0.05 y 0.01), la prueba de Kullback-Leibler no reportó valor para la potencia, por lo que la prueba no converge para estos casos.
- Con tamaños de muestra mayores de 10, la distribución Lognormal, presentó resultados del potencia de la prueba muy parecidos a los que se obtuvieron con las distribuciones Weibull y Gamma. (Tablas de la 4 a la 21)
- Con respecto al tamaño de la prueba, los resultados obtenidos por medio de simulación son muy parecidos a los valores α de que se consideraron. Se obtuvo el tamaño de la prueba con diferentes parámetros a diferentes combinaciones de niveles de significancia y tamaños de muestra; confirmando que la prueba trabaja adecuadamente. (Tablas de la 22 y 23)

8. CONCLUSIONES

- Se derivó una prueba de bondad de ajuste de Kullback-Leibler, para la distribución Pareto observando que la potencia de la prueba de Kullback-Leibler, aumenta conforme se incrementa el tamaño de la muestra, lo que demuestra que es una PRUEBA CONSISTENTE
- Con respecto al tamaño de la prueba, los resultados obtenidos por medio de simulación son muy parecidos a los valores de α que se consideraron.
- Como conclusión más importante, se observó que la prueba de Kullback-Leibler resultó ser más poderosa que la prueba de bondad de ajuste de Kolmogorov-Smirnov.

9. LITERATURA CITADA

- Arizono, I. and H. Ohta. 1989. A test for normality based on Kullback-Leibler Information. *American Statistician* 34, 20-23.
- Arnold B., C. y J. Press. 1983. Bayesian inference for Pareto populations. *Journal of Econometrics* 21, 287-306.
- Arnold B., C. y J. Press. 1976. Bayesian estimation and prediction for Pareto data. *Journal of the American Statistical Association* 84, (408), 1079-1084.
- Balakrishnan N., H. K.T-Ng, and N. Kannan. 2004. Goodness-of-fit based on spacings for progressively Type-II censored data from a general location-scale distribution. *IEEE Transactions on Reliability* 53 (3), 349- 356.
- Beirlant J., T. de Wet and Y. Goegebeur. 2006. A goodness, of-fit statistic for Pareto-type behaviour. *Journal of Computational and Applied Mathematics*. 186, 99-116.
- Bouchaud, J. 2000. *Theory of Financial Risks*. Cambridge University Press
- Bowman, A. W. 1992. Density based tests for goodness-of-fit. *Journal Statistical Computation and Simulation*, 40, 1-13 Cit on. Park, S.1995.
- David F., N. and N. L Johnson.1948. The probability integral transformation when parameters are estimated from the sample. *Biometrika*, 35, 182-190. Cit on Lilliefors 1969.
- Ebrahimi, N and M. Habibullah and E. S. Soofi. 1992. Testing exponentiality based on Kullback-Leibler information. *Journal of the Royal Statistical Society ser. B* 54 (3), 739-748.
- Ebrahimi, N. 1998. Testing exponentiality of the residual life, based on Kullback-Leibler information. *IEEE Transactions on Reliability* 47, 97-201.
- Hanagal, D. D. 1996a. Estimation of system reliability under bivariate Pareto distribution. *Parishankyan Samikkha*, 3, 3-18.
- Hanagal, D. D., 1996b. Estimation of system reliability in two component stress streng models. *Economic Quality Control* 11, 145-154.
- Johnson, N. and S. Kotz. 1970. *Distributions in Statistics. Continuous Univariate Distributions*. 1. Houghton Mifflin Company. Boston. Chapter 19.
- Kullback, S. and Leibler, R. A., 1951. On information and sufficiency, *Annals of Mathematical Statistics* 4, 49-70.

- Lehman, E. L. and G. Casella, 1998. Theory of point estimation. Springer. Verlag. New York, INC. p. 486
- Lilliefors, H. W. 1969. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. Journal of the American Statistical Association. 64 (325), 387-389.
- Makino, T. 1984. Mean hazard rate and its application to the normal approximation of the Weibull distribution. Naval Research Logistics Quarterly, 31 (1), 1-8, ISSN 0028-1441. Cit on S. Jayatilleka and G. Okogbaa. 2003. Use of accelerated life tests on transmission belts for predicting product live, identifying better designs, materials and suppliers. IEEE Procs. Annual Reliability and Maintainability Symposium. 101-105.
- Meeker, W. Q, and L. A. Escobar. 1998. Statistical methods for reliability data. Wiley Series in Probability and Statistics.
- Michael, J.R. 1983. The stabilized probability plot. Biometrika. 70, 11-17.
- Ouyang, L. - Y. and S. - J, Wu. 1994. Prediction intervals for an ordered observation From a Pareto distribution. IEEE Transactions on Reliability, 43 (2), 264-269.
- Pardo, L. 1997. Teoría de la Información Estadística. (Análisis de datos Categorizables). Editorial Hespérides, Salamanca, España.
- Park, S. 1995. The entropy of consecutive order statistics, IEEE Transactions Information Theory. 41. 2003-2007.
- Park, S. 2005. Testing exponentiality based on the Kullback-Leibler information with the type II censored data. IEEE Transactions on reliability, 54(1), 22-26.
- Petersen, J. L. 2000. Estimating the Parameters of a Pareto Distribution. Technical Report. University of Montana.
- Porter III, J. E., J. W. Coleman and A. H. Moore. 1992. Modified KS, AD, and C-vM tests for the Pareto distribution with unknown location & scale parameters. IEEE Transactions on Reliability, 41(1), 112-117.
- R Development CoreTeam. R: A language and environment for statistical Computing. The R Foundation for Statistical Computing, Viena, Austria. Version 2.3.1 (2006-06-01) ISBN 3-900051-07-0, URL . <http://www.R-project.org>- Copyright 2006.
- Pérez, R. P., H. Vaquera y J. Villaseñor. 2006. Prueba de bondad de ajuste para la distribución Gumbel, basada en la divergencia de Kullback-Leibler. Colegio de Postgraduados, México. (En prensa).

- Senoglu, B. and B. Sürücü. 2004. Goodness-of-fit based on Kullback-Leibler information. IEEE Transactions on Reliability, 53, 357-361.
- Shannon, C. E., 1948. A mathematical theory of communications. Bell System Tech. J. 27, 379-423. Cit on Park, S. 1995.
- Song, K. S. 2002. Goodness-of-fit based on Kullback-Leibler discrimination information. IEEE Transactions on Information Theory. 48, 1103-1117.
- Teitler et al. 1986. Maximum entropy, reliability distributions. IEEE Transactions on Reliability, 35, 391-395. Cit on Park, S. 2005.
- Vännman, K. 1976. Estimators based on order statistics from a Pareto distribution. Journal of the American Statistical Association, 71 (355), 704-708.
- Vasicek, O. 1976. A test for normality based on sample entropy. Journal of the Royal Statistical Society Ser. B, (38), 54-59.

ANEXOS

ANEXO A

#PROGRAMA PARA REALIZAR LA PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN PARETO.

#SE HACE LA TRANSFORMACIÓN DE LA FAMILIA ORIGINAL DE DISTRIBUCIÓN A UNA FAMILIA DE LOCALIDAD-ESCALA RESULTANDO LA DISTRIBUCIÓN EXPONENCIAL DE 2 PARÁMETROS .
#Tomado de Theory Puntual Estimation (TPE) (Lehmann, pág. 486).

1) PROGRAMA PARA GENERAR NÚMEROS ALEATORIOS DE LA DISTRIBUCIÓN PARETO

#n es el tamaño de muestra
#a
#b

```
rpareto<-function(n,a,b)
{
  u<-runif(n)
  return(b*(1/(1-u))^(1/a))
}
```

2) PROGRAMA PARA EVALUAR LA FUNCIÓN DE DENSIDAD DE LA DISTRIBUCIÓN EXPONENCIAL DOS PARÁMETROS

#Se corre primero el programa 1)
#a parámetro de escala, d parametro de localidad
fexpdosparámetros <-function(x,a,d)
{
 a*exp(-a*(x-d))
}

3) PROGRAMA PARA OBTIENE LOS ESTIMADORES DE MÁXIMA VEROSIMILITUD (EMV) DE LA DISTRIBUCIÓN PARETO

##Se corren primero los programas 1) & 2)
#ver página 486, TPE (Lehmann).
#datos es el vector de datos, corregir es un valor lógico, si es verdadero
#reemplaza el emv contenido en solucion[2] por el UMVUE del mismo.
emvpareto<-function(datos,corregir=TRUE)
{
 n<-length(datos)
 solucion<-rep(0,2)

```

solucion[2]<-min(datos)
solucion[1]<-n/(sum(log(datos))-n*log(solucion[2]))
if(corregir)
{
  solucion[2]<-solucion[2]*(1-1/((n-1)*solucion[1]))
}
return(solucion)
}

```

```

#Calcula Hmn
Hmn<-function(x,m)
{
  n<-length(x) #Obtiene el número de datos en la muestra
  x<-sort(x) #Ordena los datos
  suma<-0
  ximasm<-0
  ximenosm<-0
  for(i in 1:n)
  {
    if((i+m)>n)
    {
      ximasm<-x[n]
    }
    else
    {
      ximasm<-x[i+m]
    }

    if((i-m)<1)
    {
      ximenosm<-x[1]
    }
    else
    {
      ximenosm<-x[i-m]
    }
    suma<-suma+log(n/(2*m)*(ximasm-ximenosm))
  }
  return(suma/n)
}

```

#4) PROGRAMA PARA CALCULA EL ESTIMADOR DE KLMN(F:F0) DE LA INFORMACIÓN DE KULLBACK-LEIBLER

```

#Bajo la hipótesis nula si  $X \sim \text{Pareto}(a,c)$  luego  $Z = \log(X) \sim \text{Exp}$  de dos parámetros()
#ver página 486, TPE (Lehmann)
KLmn<-function(x,m)
{

```

```

estimadores<-emvpareto(x)
a<-estimadores[1]
d<-log(estimadores[2])
z<-log(x)
return(-Hmn(z,m)-mean(log(fexp dos parámetros (z,a,d))))
}

```

5) PROGRAMA PARA CALCULAR LA Distribución empírica de KLmn, para diversos tamanos de muestra(n)

#y diferentes m, el parámetro B indica el número de muestras a generar

```
DistrKLmn<-function(n=50,m=10,B=100,params)
```

```
{
  aleatorios<-rep(0,B)
  i<-1
  while(i<=B)
  {
    muestra<-rpareto(n,params[1],params[2])
    aleatorios[i]<-KLmn(muestra,m);
    i<-i+1;
  }
  return(aleatorios)
}
```

#6) Función para calcular los puntos críticos para diversos tamaños de muestra(n)

#y diferentes m, el parámetro B indica el número de muestras a generar, el parámetro

#alpha es el nivel de significancia de la prueba

#Aqui hemos considerado el caso de la "Pareto estándar", P(1,1)

#Para la distribución de KLmn

```
Cmalpha<-function(n=50,m=10,B=100, alpha=c(0.01,0.025,0.05,0.10))
```

```
{
  return(quantile(DistrKLmn(n,m,B,c(1,1)),1-alpha))
}
```

7) Función para calcular la tabla de valores críticos para la prueba

#tmuestras es un vector que indica el tamaño de las muestras del cual se calculan los

#puntos críticos, B es en número de muestras de tamaño n a generar para calcular el

#punto crítico para alpha dado

```
TPCKLmn<-function(tmuestras,B=100,alpha=c(0.01,0.025,0.05,0.10))
```

```
{
  nmax<-max(tmuestras) #Calcula el tamaño de muestra máximo
  mmax<-4*nmax^(1/3)
```

```
  #Creación de la tabla para contener los puntos críticos
```

```

Tabla<-matrix(nrow=length(alpha)*length(tmuestras),ncol=mmax)
cat("\nTrabajando...")
cat("\nAI 100% ")
for(i in 1:length(tmuestras))
{
  cat("#")
}
cat("\nActual ")

for(i in 1:length(tmuestras))
{
  if(tmuestras[i]%%2==0)
  {
    #i es par
    mmaxj<-(tmuestras[i]/2)-1
    if(mmaxj>mmax) mmaxj<-mmax
  }
  else
  {
    #i es impar
    mmaxj<-as.integer(tmuestras[i]/2.0)
    if(mmaxj>mmax) mmaxj<-mmax
  }
  for(j in 1:mmaxj)
  {
    Constantes<-Cmalpha(tmuestras[i],j,B,alpha)
    Tabla[i,j]<-Constantes[1]
    Tabla[length(tmuestras)+i,j]<-Constantes[2]
    Tabla[2*length(tmuestras)+i,j]<-Constantes[3]
    Tabla[3*length(tmuestras)+i,j]<-Constantes[4]
  }
  cat("#")
}
cat("\n")
dimnames(Tabla)<-list(rep(tmuestras,4),c(1:mmax))
return(Tabla)
}

```

8) Función para obtener los valores de m adecuados a cada tamaño de muestra
#Regresa una tabla apilada. la primera parte corresponde a los valores para el primer nivel
#de significancia, la segunda corresponde a los valores para el segundo nivel de significancia y así
#sucesivamente.
depura<-function(datos)
{

```

  filas<-nrow(datos)
  Criticos<-matrix(nrow=filas,ncol=4)

```

```

for(i in 1:filas)
{
  filai<-datos[i,] #Asigna la fila i-ésima de la matriz de datos
  minimo<-min(filai,na.rm=TRUE) #Obtiene el mínimo de la fila i
  maximo<-max(filai,na.rm=TRUE) #Obtiene el máximo de la fila i
  Criticos[i,1]<-minimo
  Criticos[i,2]<-match(minimo,filai) #Obtiene la posición del mínimo en
la tabla
  Criticos[i,3]<-maximo
  Criticos[i,4]<-match(maximo,filai) #Obtiene la posición del máximo
}
return(Criticos)
}

```

9) Función para obtener el punto critico de la prueba basada en K-L,
#n es el tamaño de muestra, y alpha es el nivel de significancia,
#si n o alpha no están en la tabla regresa NA NA

```

CriticoKL<-function(n,alpha)
{
  #Crea la tabla de valores críticos
  Tabla<-matrix(nrow=21,ncol=8)
  Tabla[1,]<-c(0.7284,3,0.6589,3,0.5901,3,0.5293,3)
  Tabla[2,]<-c(0.4312,4,0.3941,4,0.3572,4,0.3235,4)
  Tabla[3,]<-c(0.3218,5,0.2911,5,0.2669,5,0.2399,5)
  Tabla[4,]<-c(0.2643,6,0.2387,6,0.2176,6,0.1978,6)
  Tabla[5,]<-c(0.2254,7,0.2051,5,0.1879,7,0.1701,7)
  Tabla[6,]<-c(0.1987,6,0.1816,8,0.1649,8,0.1499,7)
  Tabla[7,]<-c(0.1774,7,0.1604,7,0.1475,7,0.1336,7)
  Tabla[8,]<-c(0.1609,9,0.1473,9,0.1348,8,0.1231,7)
  Tabla[9,]<-c(0.1474,10,0.1331,8,0.1239,8,0.1125,8)
  Tabla[10,]<-c(0.1381,9,0.1259,8,0.1158,9,0.1046,9)
  Tabla[11,]<-c(0.1261,11,0.1155,10,0.1071,9,0.0981,9)
  Tabla[12,]<-c(0.1194,8,0.1096,11,0.1014,10,0.0922,11)
  Tabla[13,]<-c(0.1125,9,0.1038,9,0.0954,12,0.0871,12)
  Tabla[14,]<-c(0.1049,11,0.0976,11,0.0903,11,0.0828,12)
  Tabla[15,]<-c(0.1014,12,0.0933,12,0.0866,12,0.0786,12)
  Tabla[16,]<-c(0.0969,10,0.0895,11,0.0833,11,0.0754,14)
  Tabla[17,]<-c(0.0935,11,0.0857,13,0.0787,13,0.0719,13)
  Tabla[18,]<-c(0.0891,11,0.0826,11,0.0766,13,0.0696,13)
  Tabla[19,]<-c(0.0857,11,0.0791,11,0.0733,11,0.0674,13)
  Tabla[20,]<-c(0.0836,12,0.0770,14,0.0711,14,0.0646,15)
  Tabla[21,]<-c(0.0763,15,0.0705,15,0.0650,15,0.0596,15)
  i<-
c(10,20,30,40,50,60,70,80,90,100,110,120,130,140,150,160,170,180,190,200,225
)
  j<-c(0.01,NA,0.025,NA,0.05,NA,0.10,NA)
  critico<-Tabla[match(n,i),match(alpha,j)]
  m<-Tabla[match(n,i),match(alpha,j)+1]

```

```

    return(c(critico,m))
}

#Función que indica si se rechaza o no se rechaza la Hipótesis nula
#Ho:F(x) = Pareto, Ha:F(x)<> Pareto
#Si se rechaza regresa un 1, caso contrario un 0,
#datos es el vector de datos,alpha es el nivel de significancia de la prueba
PhixKL<-function(datos,alpha)
{
  n<-length(datos)          #Tamaño de la muestra
  Constantes<-CriticoKL(n,alpha)  #Obtiene la constante crítica
  pCritico<-Constantes[1]      #Valor crítico
  m<-Constantes[2]           #"Window size" asociado
  if(KLmn(datos,m)>pCritico) return(1)
  else return(0)
}

# 10) PROGRAMA PARA OBTENER LA POTENCIA DE LA PRUEBA DE
KULLBACK-LEIBLER
#n es el tamaño de muestra, alpha es en nivel de significancia de la prueba
#B es el número de veces que se repite el proceso para estimar la potencia
#alternativa es el modelo distribucional que genera las muestras
#params, es un vector de parámetros de la distribución que genera las muestras
PotKL<-function(n=30,alpha=0.05,B=1000,alternativa="pareto",params=c(2,3))
{
  rechazos<-rep(0,B) #Crea un vector para contar el número de veces que se
rechaza Ho
  for(i in 1:B)
  {
    x<-switch(alternativa,weibull=rweibull(n,params[1],params[2]),
              gamma=rgamma(n,shape=params[1],scale=params[2]),
              normal=rnorm(n,params[1],params[2]),
              pareto=rpareto(n,params[1],params[2]),
              lognormal=rlnorm(n,params[1],params[2]))
    rechazos[i]<-PhixKL(x,alpha)
  }
  return(sum(rechazos)/B) #Regresa la potencia de la prueba calculada por
Monte-Carlo
}

# 11) Función para graficar la distribución empírica de Klmn para la distribución
Pareto
#t muestras es un vector con los tamaños de las muestras, m el "window size"
asociado
#B número de veces que se repite el proceso para obtener las distribuciones
empíricas
#params es un vector de parámetros para el modelo Pareto.

```

```

#grafica<-function(tmuestras,m=4,B=40000,params)
{
  cuantas<-length(tmuestras)

  plot(density(DistrKLMn(tmuestras[1],m,B,params)),main="",xlab="",ylab="",bty="l",l
ty=1)
  legpos<-locator(1)
  legend(legpos,paste("KLMn, Pareto(",params[1],",",params[2],")"),bty="n",lty=1)
  for(i in 2:cuantas)
  {
    lines(density(DistrKLMn(tmuestras[i],m,B,c(params[2*i-
1],params[2*i]))),bty="l",lty=i)
    legend(legpos$x,legpos$y-(i-1)*0.25,paste("KLMn;Pareto(",params[2*i-
1],",",params[2*i],")"),bty="n",lty=i)
  }
}

```

12) Función para calcular la estadística de Kolmogorov, ver Lilliefors, 1969.

#x es el vector de datos

Dstat<-function(x)

```

{
  #obtiene los estimadores máximo verosímiles para la distribución Pareto
  #transforma los datos a exponencial de dos parámetros y elimina luego el
  #parámetro de corrimiento, por lo que 'y' tendrá una muestra de variables
  #aleatorias #independientes idénticamente distribuidas exponencial(a)
  estimadores<-emvpareto(x)
  a<-estimadores[1]
  d<-log(estimadores[2])
  y<-log(x)-d
  y<-sort(y)
  z<-pexp(y,a)
  Dmas<-max(c(1:length(x))/length(x)-z) #Calcula D+
  Dmenos<-max(z-c(0:(length(x)-1))/length(x)) #Calcula D-
  return(max(Dmas,Dmenos))
}

```

13) PROGRAMA PARA OBTENER La tabla de puntos críticos de Lilliefors

CriticoKolmogorov<-function(n,alpha)

```

{
  #columna 1(alpha=0.10), columna 2(alpha=0.05), columna 3(alpha=0.01)
  Criticos<-matrix(nrow=4, ncol=3)
  Criticos[1,]<-c(0.295,0.325,0.380) #Tamaño de muestra 10
  Criticos[2,]<-c(0.212,0.234,0.278) # Tamaño de muestra 20
  Criticos[3,]<-c(0.174,0.192,0.226) # Tamaño de muestra 30
  Criticos[4,]<-c(0.960,1.060,1.250) # Tamaño de muestra infinito
  i<-c(10,20,30,NA) #índice de fila
  j<-c(0.10,0.05,0.01) #índice de columna
  if(n<=30)

```



```

    {
      return(Criticos[match(n,i),match(alpha,j)])
    }
    else
    {
      return((Criticos[4,match(alpha,j)]/sqrt(n))
    }
  }
}

```

14) Función que indica si se rechaza o no se rechaza la Hipótesis nula

#Ho:F(x) = Pareto, Ha:F(x)<> Pareto

#Si se rechaza regresa un 1, caso contrario un 0

#datos es el vector de datos, (la muestra que se presume es Pareto)

#alpha es el nivel de significancia de la prueba

PhixKolmogorov<-function(datos,alpha)

```

{
  n<-length(datos)
  pcritico<-CriticoKolmogorov(n,alpha)
  D<-Dstat(datos)
  if(D>pcritico) return(1)
  else return(0)
}

```

15) Obtiene la potencia de la prueba de Kolmogorov

#n es el tamaño de muestra, alpha es en nivel de significancia de la prueba

#B es el número de veces que se repite el proceso para estimar la potencia

#alternativa es el modelo distribucional que genera las muestras

#params, es un vector de parámetros de la distribución que genera las muestras

PotKolmogorov<-

function(n=30,alpha=0.05,B=1000,alternativa="pareto",params=c(2,3))

```

{
  rechazos<-rep(0,B) #Crea un vector para contar el número de veces que se
rechaza Ho
  for(i in 1:B)
  {
    x<-switch(alternativa,weibull=rweibull(n,params[1],params[2]),
              gamma=rgamma(n,shape=params[1],scale=params[2]),
              normal=rnorm(n,params[1],params[2]),
              pareto=rpareto(n,params[1],params[2]),
              lognormal=rlnorm(n,params[1],params[2]))
    rechazos[i]<-PhixKolmogorov(x,alpha)
  }
  return(sum(rechazos)/B) #Regresa la potencia de la prueba calculada por
Monte-Carlo
}

```

16) Muestra que la estadística de prueba es invariante

#grafica(rep(20,4),10,40000,c(1,1,2,1,3,1,10,50,10,100))

```

# 17) Calcula la tabla de valores críticos para la prueba para los tamaños de
muestra indicados y las alphas indicados
#y las almacena en archivos de texto separados por comas

#library(MASS)

#b<-
TPCKLmn(c(20,30,40,50,60,70,80,90,100,110,120,130,140,150,160,170,180,190,
200),B=5000,alpha=c(0.01,0.025,0.05,0.10))
#c<-depura(b)
#write.matrix(b,"criticos.csv",sep = ",")
#write.matrix(c,"depurados.csv",sep = ",")

#b<-TPCKLmn(c(10,225),B=5000,alpha=c(0.01,0.025,0.05,0.10))
#c<-depura(b)
#write.matrix(b,"criticos.csv",sep = ",")
#write.matrix(c,"depurados.csv",sep = ",")

#Potencias y tamaños, prueba de KL
#PotKL(10,0.05,10000,"pareto",c(3,1))

#Potencias y tamaños, prueba de Kolmogorov
#PotKolmogorov(100,0.01,10000,"pareto",c(3,1))

#EXAMPLE FAILURE
FAILURE<-
c(30.101,30.150,30.374,30.581,30.871,31.086,31.398,31.752,31.792,31.960,32.2
65,32.517,32.636,33.002,33.552,33.721,34.002,34.023,34.150,35.274)

alpha=0.05 #nivel de significancia
n=length(FAILURE) #Tamano de muestra
n #Muestra el tamano de muestra
Constantes<-CriticoKL(20,alpha) #Obtiene la constante critica de tablas
#Vector con dos componentes, primer componente
#constante critica, segundo componente
#m asociado
pCritico<-Constantes[1] #Valor critico
pCritico
m<-Constantes[2] #m
m
estimadores<-emvpareto(FAILURE) #EMV
estimadores
a<-estimadores[1] #Paremetro de escala de exp dos parámetros
a
d<-log(estimadores[2]) #Paremetro de loc de exp dos parámetros
d
z<-log(FAILURE)
z #Datos exp dos parámetros (escala, loc)

```

```
w<-z-d
w          #Datos exp(escala)
x11()
plot(ecdf(w),do.p=FALSE, main=" FAILURE_DISTRIBUCIONES_.Empírica vs
Teorica")
curve(pexp(x,a),add=T)

ifelse(KLmn(FAILURE,m)>pCritico,1,0)  #Si regresa 1 se rechaza H0
```

Tabla 4. Valores críticos $C_{mn}(\alpha)$ de la estadística KL_{mn} , para la distribución Pareto obtenidos mediante simulación Monte Carlo $\alpha = 0.01$

n	m																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
5	1.652575	1.240538	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	1.5068	1.097264	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	1.309291	0.97631	0.951256	0.86769	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	1.289051	0.890081	0.86769	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	1.191634	0.861083	0.793586	0.845133	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10	1.138812	0.784238	0.712432	0.766284	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
11	1.048764	0.750008	0.693334	0.688841	0.757042	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
12	0.998351	0.706855	0.647158	0.641061	0.685685	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	0.92245	0.684074	0.592221	0.62156	0.643406	0.679912	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
14	0.899345	0.626796	0.588691	0.562072	0.608238	0.656003	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
15	0.871696	0.602925	0.529638	0.559842	0.572381	0.589806	0.66197	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
16	0.838232	0.594934	0.524431	0.522919	0.52727	0.569106	0.61575	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
17	0.818025	0.566155	0.513141	0.486389	0.501519	0.522343	0.57637	0.60947	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
18	0.77829	0.555865	0.484688	0.471263	0.484486	0.503373	0.53678	0.57811	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
19	0.747918	0.543941	0.477607	0.453137	0.457767	0.485658	0.51016	0.54243	0.57843	NA	NA	NA	NA	NA	NA	NA	NA	NA
20	0.742135	0.498276	0.454094	0.444173	0.457995	0.458229	0.48816	0.52112	0.53857	NA	NA	NA	NA	NA	NA	NA	NA	NA
30	0.617426	0.404245	0.344535	0.331966	0.319449	0.329554	0.33724	0.34656	0.36708	0.38002	0.40164	0.4208345	0.4335695	0.44979	NA	NA	NA	NA
40	0.560986	0.347179	0.294022	0.270662	0.272473	0.263711	0.26845	0.27573	0.27805	0.28785	0.29363	0.3105175	0.3279144	0.33648	0.35519	0.35783	0.37143	0.387116
50	0.507832	0.318969	0.25657	0.2337	0.228573	0.219072	0.23238	0.22604	0.23578	0.23999	0.24631	0.2520889	0.2605182	0.27624	0.27015	0.28761	0.29448	0.30013
60	0.474198	0.295727	0.242172	0.214885	0.199143	0.197963	0.1952	0.19811	0.20323	0.20463	0.20846	0.2215975	0.2241966	0.22984	0.23146	0.23534	0.24993	0.252911
70	0.457967	0.28182	0.219361	0.195597	0.183755	0.173091	0.1775	0.17929	0.17576	0.1801	0.18457	0.1883123	0.1945716	0.19857	0.20152	0.20961	0.21126	0.219909
80	0.444125	0.266322	0.204097	0.182587	0.16926	0.16493	0.1602	0.15761	0.16039	0.16362	0.16563	0.1683491	0.1673798	0.17367	0.18158	0.18234	0.19078	0.188253
90	0.426612	0.251825	0.195864	0.174283	0.160468	0.150947	0.14776	0.1466	0.14898	0.1472	0.14907	0.1561679	0.1574815	0.15717	0.16159	0.17081	0.17008	0.172733
100	0.416666	0.24617	0.189325	0.165114	0.151381	0.140545	0.13939	0.1351	0.13712	0.13618	0.13735	0.1390839	0.1404432	0.14743	0.15274	0.15267	0.14959	0.156744

Tabla 5. Valores críticos $C_{mn}(\alpha)$ de la estadística KL_{mn} , para la distribución Pareto obtenidos mediante simulación Monte Carlo $\alpha=0.025$

n	m																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
5	1.46259	1.11293	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	1.32525	0.96578	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	1.18606	0.85993	0.86657	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	1.1174	0.79242	0.76868	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	1.04661	0.751	0.69133	0.75297	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10	0.97412	0.69628	0.64141	0.68917	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
11	0.93541	0.65035	0.61616	0.62862	0.66722	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
12	0.88911	0.61925	0.57827	0.57968	0.62483	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	0.84105	0.59809	0.54327	0.55501	0.57825	0.61544	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
14	0.7955	0.55513	0.51673	0.50128	0.54189	0.58849	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
15	0.77234	0.53792	0.48322	0.491	0.50804	0.53418	0.60136	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
16	0.74902	0.52226	0.47056	0.47787	0.47965	0.51123	0.55747	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
17	0.72491	0.51065	0.45613	0.44542	0.45093	0.476	0.52115	0.547	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
18	0.70834	0.48686	0.43839	0.42842	0.43791	0.46186	0.48305	0.517	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
19	0.67837	0.48024	0.41709	0.40987	0.40865	0.43622	0.4623	0.483	0.52391	NA	NA	NA	NA	NA	NA	NA	NA	NA
20	0.6632	0.45095	0.4092	0.39392	0.409	0.40925	0.43523	0.466	0.4864	NA	NA	NA	NA	NA	NA	NA	NA	NA
30	0.56019	0.36443	0.31314	0.29675	0.29098	0.29919	0.3044	0.315	0.33515	0.34574	0.36025	0.377	0.39265	0.409952	NA	NA	NA	NA
40	0.51166	0.31888	0.26511	0.24775	0.24582	0.23856	0.24774	0.248	0.25708	0.26227	0.27012	0.28006	0.29883	0.300105	0.32401	0.32828	0.33623	0.347737
50	0.47308	0.29044	0.23552	0.2137	0.20526	0.2018	0.21155	0.208	0.21452	0.21986	0.22625	0.23124	0.23711	0.247313	0.25017	0.264	0.26812	0.273579
60	0.4473	0.27036	0.21846	0.1947	0.18191	0.18107	0.17863	0.179	0.18734	0.18832	0.19122	0.20098	0.20243	0.206856	0.21145	0.21453	0.22488	0.232788
70	0.42487	0.26154	0.20274	0.17978	0.16703	0.15791	0.16138	0.163	0.16123	0.1647	0.16805	0.17199	0.17518	0.181479	0.18229	0.18825	0.19308	0.198231
80	0.41884	0.24328	0.18805	0.16879	0.15428	0.14855	0.14745	0.145	0.14507	0.14787	0.14984	0.15379	0.1527	0.15905	0.16524	0.1669	0.1736	0.173448
90	0.4079	0.23388	0.18136	0.16017	0.14565	0.14072	0.13474	0.133	0.1344	0.13491	0.13659	0.13905	0.14364	0.144665	0.1479	0.15309	0.15324	0.156645
100	0.39793	0.22845	0.17396	0.15194	0.13725	0.12809	0.1272	0.125	0.12564	0.12464	0.12567	0.12828	0.12937	0.134972	0.13748	0.14012	0.13726	0.144833

Tabla 6. Valores críticos $C_{mn}(\alpha)$ de la estadística KL_{mn} , para la distribución Pareto obtenidos mediante simulación Monte Carlo $\alpha = 0.05$

n	m																		Nivel de Significancia 0.05
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
5	1.30779	0.99797	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
6	1.19622	0.86706	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
7	1.05818	0.77743	0.79113	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
8	0.98725	0.71050	0.70015	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
9	0.91860	0.68092	0.63346	0.68721	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
10	0.87705	0.61465	0.59078	0.62700	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
11	0.82216	0.58627	0.56243	0.57027	0.61291	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
12	0.78487	0.56048	0.52926	0.52698	0.57729	0.57135	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
13	0.75932	0.53196	0.49548	0.50171	0.53089	0.52678	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
14	0.72777	0.49697	0.46201	0.46123	0.48768	0.52678	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
15	0.69373	0.48163	0.44165	0.44594	0.46701	0.49200	0.54364	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
16	0.67050	0.47474	0.42436	0.43080	0.43597	0.46797	0.50634	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
17	0.66009	0.45920	0.41309	0.40622	0.41134	0.43749	0.47246	0.50167	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
18	0.64354	0.44224	0.39499	0.38712	0.39585	0.42029	0.44223	0.47205	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
19	0.61896	0.43359	0.37682	0.37330	0.37562	0.39876	0.42002	0.44623	0.48061	NA	NA	NA	NA	NA	NA	NA	NA	NA	
20	0.61303	0.40575	0.37305	0.35902	0.36877	0.37630	0.39923	0.42484	0.44849	NA	NA	NA	NA	NA	NA	NA	NA	NA	
30	0.51916	0.33398	0.28611	0.27080	0.26642	0.27464	0.28340	0.29140	0.30542	0.31358	0.32999	0.34713	0.36197	0.37377	NA	NA	NA	NA	
40	0.47481	0.29383	0.24285	0.22626	0.22430	0.21889	0.22657	0.22836	0.23820	0.24583	0.24736	0.26070	0.27260	0.27495	0.29375	0.30136	0.30819	0.31769	
50	0.44364	0.26876	0.21865	0.19619	0.18886	0.18606	0.19287	0.19225	0.19890	0.20246	0.20774	0.21543	0.21904	0.22466	0.22832	0.23988	0.24592	0.24971	
60	0.42156	0.25205	0.19954	0.17883	0.16857	0.16704	0.16459	0.16605	0.17008	0.17305	0.17615	0.18132	0.18604	0.19134	0.19428	0.19930	0.20460	0.21197	
70	0.40495	0.24068	0.18765	0.16602	0.15345	0.14717	0.14700	0.14955	0.15000	0.15195	0.15499	0.15752	0.16286	0.16841	0.16885	0.17294	0.17658	0.18144	
80	0.39474	0.22807	0.17665	0.15439	0.14295	0.13860	0.13467	0.13370	0.13397	0.13678	0.13750	0.14130	0.14207	0.14687	0.15229	0.15304	0.15902	0.16025	
90	0.38864	0.22108	0.16808	0.14747	0.13590	0.12907	0.12503	0.12299	0.12435	0.12489	0.12605	0.12790	0.13079	0.13378	0.13604	0.13948	0.13924	0.14347	
100	0.37905	0.21303	0.16343	0.13929	0.12774	0.11933	0.11794	0.11454	0.11473	0.11491	0.11561	0.11812	0.12003	0.12375	0.12402	0.12757	0.12762	0.13338	

Tabla 7. Valores críticos $C_{mn}(\alpha)$ de la estadística KL_{mn} , para la distribución Pareto obtenidos mediante simulación Monte Carlo $\alpha=0.10$

n	m																		Nivel de Significancia 0.10
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
5	1.134927	0.8834	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
6	1.024422	0.76688	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
7	0.91384	0.68755	0.70204	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
8	0.847069	0.64188	0.61552	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
9	0.814471	0.59699	0.55867	0.61906	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
10	0.76646	0.55135	0.52832	0.55841	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
11	0.724774	0.52219	0.50016	0.51062	0.55123	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
12	0.700871	0.49428	0.46828	0.47423	0.51628	0.477	0.51535	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
13	0.678105	0.47236	0.43903	0.44846	0.477	0.51535	0.47334	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
14	0.646099	0.44886	0.41457	0.41538	0.43568	0.47334	0.47334	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
15	0.615436	0.42696	0.39422	0.39924	0.41785	0.44306	0.4818	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
16	0.598753	0.41742	0.37939	0.38236	0.39324	0.41616	0.44985	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
17	0.588911	0.40587	0.37288	0.36475	0.36953	0.39739	0.42432	0.45167	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
18	0.568142	0.38896	0.35219	0.34362	0.35665	0.37842	0.40029	0.42379	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
19	0.557129	0.38572	0.33677	0.33638	0.34235	0.3594	0.37602	0.40389	0.42894	NA	NA	NA	NA	NA	NA	NA	NA	NA	
20	0.545839	0.36335	0.33236	0.32256	0.33345	0.34362	0.35951	0.38354	0.40259	NA	NA	NA	NA	NA	NA	NA	NA	NA	
30	0.47455	0.30107	0.25681	0.24507	0.24086	0.24879	0.25463	0.26562	0.27434	0.28214	0.2973	0.31129	0.322652	0.33323	NA	NA	NA	NA	
40	0.43293	0.26913	0.22065	0.20381	0.20096	0.19676	0.20498	0.2079	0.21355	0.22187	0.22613	0.23711	0.243428	0.24944	0.26193	0.26885	0.27507	0.28224	
50	0.407223	0.24523	0.19886	0.17804	0.17191	0.16825	0.17232	0.17214	0.17916	0.18296	0.18787	0.19361	0.196068	0.20278	0.20528	0.21291	0.21711	0.225282	
60	0.392922	0.23049	0.18307	0.16335	0.15345	0.15111	0.15004	0.15081	0.1531	0.1562	0.15941	0.16336	0.168993	0.17407	0.1756	0.1794	0.18283	0.188975	
70	0.379806	0.22173	0.17202	0.15217	0.14129	0.13497	0.13372	0.13619	0.13599	0.13758	0.14093	0.14217	0.147315	0.15161	0.1521	0.15574	0.1594	0.161053	
80	0.371368	0.213	0.16221	0.14114	0.12966	0.12553	0.12282	0.12185	0.12301	0.12363	0.12596	0.12796	0.12748	0.13262	0.13706	0.13802	0.14157	0.142719	
90	0.364512	0.20502	0.15717	0.1348	0.12336	0.11783	0.11408	0.11276	0.11419	0.11419	0.11476	0.11527	0.117575	0.11975	0.12194	0.12436	0.12458	0.129643	
100	0.358832	0.19823	0.15133	0.12822	0.11701	0.10945	0.10691	0.10547	0.10441	0.10544	0.10543	0.10812	0.108383	0.1112	0.11062	0.11467	0.11377	0.118613	

Tabla 8. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=10$, $\alpha=0.10$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)		0.851			0.885
Weibull	(1,1)		0.8414			//
Gamma	(2,1)		0.7904			0.84
Weibull	(2,1)		0.8362			0.884
Gamma	(3,1)		0.7774			0.8084
Weibull	(3,1)		0.8344			0.8828
Gamma	(3,439,1)		0.767			0.8006
Weibull	(3,439,1)		0.8334			0.8844
Gamma	(1,3)		0.8472			0.8864
Weibull	(1,3)		0.8392			//
Gamma	(2,3)		0.8012			0.8504
Weibull	(2,3)		0.8436			0.8886
Lognormal	(1,1)		0.6692			0.6808
Lognormal	(1,0,4)		0.671			0.6758
Lognormal	(5,3)		0.7308			//
Lognormal	(10,0,2)		0.6646			0.669

Tabla 9. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=10$, $\alpha=0.05$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)		0.737			0.8222
Weibull	(1,1)		0.7568			0.8162
Gamma	(2,1)		0.687			0.7498
Weibull	(2,1)		0.7308			0.8096
Gamma	(3,1)		0.6686			0.716
Weibull	(3,1)		0.7264			0.8084
Gamma	(3,439,1)		0.6398			0.7014
Weibull	(3,439,1)		0.7308			0.8138
Gamma	(1,3)		0.7556			0.5008
Weibull	(1,3)		0.7466			//
Gamma	(2,3)		0.6844			0.7588
Weibull	(2,3)		0.7348			0.8244
Lognormal	(1,1)		0.5266			0.5542
Lognormal	(1,0,4)		0.533			0.541
Lognormal	(5,3)		0.05			//
Lognormal	(10,0,2)		0.5256			0.5278

Tabla 10. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler considerando distribuciones alternativas para $n=10$, $\alpha=0.01$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)	(1,1)	0.514		0.6052	
Weibull	(1,1)	(1,1)	0.497			//
Gamma	(2,1)	(2,1)	0.4344		0.5162	
Weibull	(2,1)	(2,1)	0.4968		0.6132	
Gamma	(3,1)	(3,1)	0.3816		0.4784	
Weibull	(3,1)	(3,1)	0.4854		0.605	
Gamma	(3.439,1)	(3.439,1)	0.381		0.4508	
Weibull	(3.439,1)	(3.439,1)	0.4808		0.601	
Gamma	(1,3)	(1,3)	0.5103		0.6284	
Weibull	(1,3)	(1,3)	0.5034		0.6234	
Gamma	(2,3)	(2,3)	0.4346		0.5284	
Weibull	(2,3)	(2,3)	0.4872		0.5958	
Lognormal	(1,1)	(1,1)	0.27		0.302	
Lognormal	(1,0.4)	(1,0.4)	0.2538		0.2898	
Lognormal	(5,3)	(5,3)	0.3398			//
Lognormal	(10,0.2)	(10,0.2)	0.2672		0.2804	

Tabla 11. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=30$, $\alpha=0.10$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)	(1,1)	0.9998		1	
Weibull	(1,1)	(1,1)	0.9994		1	
Gamma	(2,1)	(2,1)	0.9986		1	
Weibull	(2,1)	(2,1)	0.9998		1	
Gamma	(3,1)	(3,1)	0.999		0.9988	
Weibull	(3,1)	(3,1)	0.9992		0.9998	
Gamma	(3.439,1)	(3.439,1)	0.9976		0.9996	
Weibull	(3.439,1)	(3.439,1)	1		1	
Gamma	(1,3)	(1,3)	0.999		1	
Weibull	(1,3)	(1,3)	0.9994		1	
Gamma	(2,3)	(2,3)	0.9984		1	
Weibull	(2,3)	(2,3)	1		1	
Lognormal	(1,1)	(1,1)	0.9904		0.9932	
Lognormal	(1,0.4)	(1,0.4)	0.9896		0.9898	
Lognormal	(5,3)	(5,3)	0.989		0.9916	
Lognormal	(10,0.2)	(10,0.2)	0.991		0.9938	

Tabla 12. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=30$, $\alpha=0.05$

<i>Distribución</i>	<i>Parámetros</i>	<i>Kolmogorov</i>	<i>Smirnov</i>	<i>Kullback</i>	<i>Leibler</i>
	κ, θ				
Gamma	(1,1)	0.9998	1		
Weibull	(1,1)	0.9996	1		
Gamma	(2,1)	0.9962	0.996		
Weibull	(2,1)	0.999	0.9996		
Gamma	(3,1)	0.9948	0.9978		
Weibull	(3,1)	0.9986	1		
Gamma	(3.439,1)	0.9934	0.9984		
Weibull	(3.439,1)	0.9986	1		
Gamma	(1,3)	0.9978	1		
Weibull	(1,3)	0.999	0.9998		
Gamma	(2,3)	0.9964	0.9994		
Weibull	(2,3)	0.9984	0.9994		
Lognormal	(1,1)	0.9764	0.983		
Lognormal	(1,0.4)	0.9798	0.9816		
Lognormal	(5,3)	0.9758	0.9804		
Lognormal	(10,0.2)	0.976	0.9816		

Tabla 13. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=30$, $\alpha=0.01$

<i>Distribución</i>	<i>Parámetros</i>	<i>Kolmogorov</i>	<i>Smirnov</i>	<i>Kullback</i>	<i>Leibler</i>
	κ, θ				
Gamma	(1,1)	0.9922	0.9994		
Weibull	(1,1)	0.99	0.9992		
Gamma	(2,1)	0.9782	0.9954		
Weibull	(2,1)	0.99	0.9988		
Gamma	(3,1)	0.976	0.9918		
Weibull	(3,1)	0.9918	0.999		
Gamma	(3.439,1)	0.9716	0.9924		
Weibull	(3.439,1)	0.9904	0.9984		
Gamma	(1,3)	0.992	0.999		
Weibull	(1,3)	0.9918	0.9988		
Gamma	(2,3)	0.9822	0.9954		
Weibull	(2,3)	0.9902	0.9984		
Lognormal	(1,1)	0.9102	0.941		
Lognormal	(1,0.4)	0.9108	0.9338		
Lognormal	(5,3)	0.9192	0.9384		
Lognormal	(10,0.2)	0.9196	0.9334		

Tabla 14. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=50$, $\alpha=0.10$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)		1			1
Weibull	(1,1)		1			1
Gamma	(2,1)		1			1
Weibull	(2,1)		1			1
Gamma	(3,1)		1			1
Weibull	(3,1)		1			1
Gamma	(3.439,1)		1			1
Weibull	(3.439,1)		1			1
Gamma	(1,3)		1			1
Weibull	(1,3)		1			1
Gamma	(2,3)		1			1
Weibull	(2,3)		1			1
Lognormal	(1,1)		1			1
Lognormal	(1,0.4)		1			1
Lognormal	(5,3)		0.9998			1
Lognormal	(10,0.2)		1			0.9994

Tabla 15. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=50$, $\alpha=0.05$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)		1			1
Weibull	(1,1)		1			1
Gamma	(2,1)		1			1
Weibull	(2,1)		1			1
Gamma	(3,1)		1			1
Weibull	(3,1)		1			1
Gamma	(3.439,1)		1			1
Weibull	(3.439,1)		1			1
Gamma	(1,3)		1			1
Weibull	(1,3)		1			1
Gamma	(2,3)		1			1
Weibull	(2,3)		1			1
Lognormal	(1,1)		0.9994			0.9995
Lognormal	(1,0.4)		0.9984			0.9992
Lognormal	(5,3)		0.9998			0.9996
Lognormal	(10,0.2)		0.9998			0.9996

Tabla 16. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=50$, $\alpha=0.01$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)		0.9998			1
Weibull	(1,1)		1			1
Gamma	(2,1)		0.9996			1
Weibull	(2,1)		1			1
Gamma	(3,1)		0.9998			1
Weibull	(3,1)		1			1
Gamma	(3.439,1)		0.9998			1
Weibull	(3.439,1)		1			1
Gamma	(1,3)		1			1
Weibull	(1,3)		1			1
Gamma	(2,3)		1			1
Weibull	(2,3)		1			1
Lognormal	(1,1)		0.9974			0.999
Lognormal	(1,0.4)		0.9956			0.9984
Lognormal	(5,3)		0.9968			0.9962
Lognormal	(10,0.2)		0.9954			0.9978

Tabla 17. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=100$, $\alpha=0.10$

Distribución	Parámetros	κ, θ	Smirnov	Kolmogorov	Kullback	Leibler
Gamma	(1,1)		1			1
Weibull	(1,1)		1			1
Gamma	(2,1)		1			1
Weibull	(2,1)		1			1
Gamma	(3,1)		1			1
Weibull	(3,1)		1			1
Gamma	(3.439,1)		1			1
Weibull	(3.439,1)		1			1
Gamma	(1,3)		1			1
Weibull	(1,3)		1			1
Gamma	(2,3)		1			1
Weibull	(2,3)		1			1
Lognormal	(1,1)		1			1
Lognormal	(1,0.4)		1			1
Lognormal	(5,3)		1			1
Lognormal	(10,0.2)		1			1

Tabla 18. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=100$, $\alpha=0.05$

Distribución	Parámetros	κ, θ		
		Smirnov	Kolmogorov	Kullback Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 19. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=100$, $\alpha=0.01$

Distribución	Parámetros	κ, θ		
		Smirnov	Kolmogorov	Kullback Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 20. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=200$, $\alpha=0.10$

Distribución	Parámetros	κ, θ		
		Smirnov	Kolmogorov	Kullback Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 21. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=200$, $\alpha=0.05$

Distribución	Parámetros	κ, θ		
		Smirnov	Kolmogorov	Kullback Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 22. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=200$, $\alpha=0.01$

Distribución	Parámetros	x, θ		
		Kolmogorov	Smirnov	Leibler
Gamma	(1,1)	0.9998	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	0.9996	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	0.9998	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	0.9998	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	0.9974	0.999	
Lognormal	(1,0.4)	0.9956	0.9984	
Lognormal	(5,3)	0.9968	0.9962	
Lognormal	(10,0.2)	0.9954	0.9978	

Tabla 23. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=225$, $\alpha=0.10$

Distribución	Parámetros	x, θ		
		Kolmogorov	Smirnov	Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 24. Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=225$, $\alpha=0.05$

Distribución	Parámetros	κ, θ		
		Smirnov	Kolmogorov	Kullback Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 25 Estimación de la potencia con las pruebas de Kolmogorov-Smirnov y Kullback-Leibler; considerando distribuciones alternativas para $n=225$, $\alpha=0.01$

Distribución	Parámetros	κ, θ		
		Smirnov	Kolmogorov	Kullback Leibler
Gamma	(1,1)	1	1	1
Weibull	(1,1)	1	1	1
Gamma	(2,1)	1	1	1
Weibull	(2,1)	1	1	1
Gamma	(3,1)	1	1	1
Weibull	(3,1)	1	1	1
Gamma	(3.439,1)	1	1	1
Weibull	(3.439,1)	1	1	1
Gamma	(1,3)	1	1	1
Weibull	(1,3)	1	1	1
Gamma	(2,3)	1	1	1
Weibull	(2,3)	1	1	1
Lognormal	(1,1)	1	1	1
Lognormal	(1,0.4)	1	1	1
Lognormal	(5,3)	1	1	1
Lognormal	(10,0.2)	1	1	1

Tabla 26. Estimación del Tamaño de la Prueba de bondad de ajuste de Kullback-Leibler considerando diferentes tamaños de muestra y niveles de significancia

κ, θ	α	n=10	n=30	n=50
(1,1)	0.1	0.094	0.116	0.1154
(1,1)	0.05	0.0512	0.0558	0.0534
(1,1)	0.01	0.0102	0.0102	0.0104
(3,1)	0.1	0.094	0.1064	0.0932
(3,1)	0.05	0.0448	0.0506	0.0486
(3,1)	0.01	0.0104	0.012	0.009
(5,3)	0.1	0.0916	0.1004	0.0982
(5,3)	0.05	0.047	0.047	0.0498
(5,3)	0.01	0.0078	0.0116	0.0108
(10,0.2)	0.1	0.091	0.0994	0.0972
(10,0.2)	0.05	0.054	0.052	0.0518
(10,0.2)	0.01	0.008	0.0092	0.0114

Tabla 27. Estimación del Tamaño de la Prueba de bondad de ajuste de Kullback-Leibler considerando diferentes tamaños de muestra y niveles de significancia

κ, θ	α	n=100	n=200	n=225
(1,1)	0.1	0.1054	0.1042	0.0994
(1,1)	0.05	0.049	0.0536	0.0564
(1,1)	0.01	0.0084	0.0096	0.0138
(3,1)	0.1	0.0944	0.0984	0.096
(3,1)	0.05	0.047	0.0478	0.057
(3,1)	0.01	0.0082	0.0096	0.0126
(5,3)	0.1	0.984	0.1046	0.1002
(5,3)	0.05	0.0476	0.0522	0.0554
(5,3)	0.01	0.0088	0.0086	0.013
(10,0.2)	0.1	0.1054	0.107	0.0976
(10,0.2)	0.05	0.0472	0.0512	0.0568
(10,0.2)	0.01	0.007	0.0098	0.0102