



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

**PRUEBAS DE BONDAD DE AJUSTE PARA LA
DISTRIBUCIÓN NORMAL ASIMÉTRICA
BASADAS EN TRANSFORMACIONES DE DATOS**

WALDENIA COSMES MARTÍNEZ

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, ESTADO DE MÉXICO
2019



COLEGIO DE POSTGRADUADOS
INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

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

En adición al beneficio ético, moral y académico que he obtenido durante mis estudios en el Colegio de Postgraduados, la que suscribe, “ **Waldenia Cosmes Martínez**”, alumna de esta Institución, estoy de acuerdo en ser partícipe de las regalías económicas y/o académicas, de procedencia nacional e internacional, que se deriven del trabajo de investigación que realicé en esta institución, bajo la dirección del Profesor “ **Dra. Elizabeth González Estrada**”, por lo que otorgo los derechos de autor de mi tesis **Pruebas de bondad de ajuste para la distribución normal asimétrica basadas en transformaciones de datos**, y de los productos de dicha investigación al Colegio de Postgraduados. Las patentes y secretos industriales que se puedan derivar serán registrados a nombre del colegio de Postgraduados y las regalías económicas que se deriven serán distribuidas entre la Institución, El Consejero o Director de Tesis y el que suscribe, de acuerdo a las negociaciones entre las tres partes, por ello me comprometo a no realizar ninguna acción que dañe el proceso de explotación comercial de dichos productos a favor de esta Institución.

Montecillo, Mpio. de Texcoco, Edo. de México, a 8 de febrero de 2019.

Waldenia Cosmes Martínez

Vo. Bo. Dra. Elizabeth González Estrada

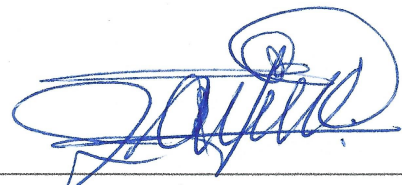
La presente tesis titulada: **Pruebas de bondad de ajuste para la distribución normal asimétrica basadas en transformaciones de datos**, realizada por la alumna: **Waldenia Cosmes Martínez**, bajo la dirección del Consejo Particular indicado ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

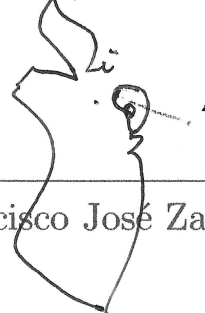
MAESTRO EN CIENCIAS

SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO 
Dra. Elizabeth González Estrada

ASESOR 
Dr. José Aurelio Villaseñor Alva

ASESOR 
Dr. Francisco José Zamudio Sánchez

Montecillo, Texcoco, Estado de México, enero de 2019

Pruebas de bondad de ajuste para la distribución normal asimétrica basadas en transformaciones de datos

Waldenia Cosmes Martínez M. en C.

Colegio de Postgraduados, 2018

RESUMEN

En los últimos años se han desarrollado diversos métodos estadísticos con base en datos provenientes de la familia de distribuciones Normal Asimétrica. De aquí la importancia de contar con pruebas de bondad de ajuste que permitan decidir si un conjunto de datos proviene de dicha familia. En el presente trabajo se proponen cuatro pruebas de bondad de ajuste, la primera está basada en una transformación de variables normales asimétricas a variables chi-cuadrada, cuyos valores críticos de prueba se obtienen maximizando la probabilidad del error tipo I sobre el espacio de parámetros de la distribución normal asimétrica. La segunda prueba propuesta consiste en transformar variables normales asimétricas a variables chi-cuadrada utilizando bootstrap paramétrico para obtener los valores críticos de la prueba. La tercer propuesta se basa en aplicar la prueba de Anderson-Darling directamente sobre variables normales asimétricas, utilizando Bootstrap paramétrico para el cálculo de los valores críticos de la prueba. Por último, la cuarta propuesta se basa en transformar variables normales asimétricas a variables normales, para probar la hipótesis de normalidad sobre las variables transformadas se utiliza la prueba de Shapiro-Wilk. Se presentan resultados de estudios de simulación para la estimación del nivel y la potencia de las pruebas frente a diversas distribuciones alternativas, las cuales fueron escogidas por sus características e importancia en la modelación de datos con asimetría. Se realizó un estudio de potencia comparativo con otras pruebas existentes, los resultados muestran que las pruebas propuestas tienen mayor potencia contra algunas distribuciones alternativas asimétricas de cola pesada.

Palabras clave: Transformaciones de datos; Bootstrap paramétrico; Distribuciones asimétricas; Prueba de Shapiro-Wilk, Prueba de Anderson-Darling.

Goodness-of-fit Tests for skew Normal Distribution Based on Transformations of data

Waldenia Cosmes Martínez M. en C.

Colegio de Postgraduados, 2018

ABSTRACT

In recent years, various statistical methods that involve data from the family of skew Normal distributions have been published. Hence the importance of having goodness-of-fit tests that allow us to decide if a set of data comes from that family of distributions. In this work, four goodness of fit tests are proposed. The first is based on a transformation from skew normal variables to approximately chi-square variables, the critical values of the test are obtained maximizing the Type I error probability. The second test consists on a transformation from skew normal variables to approximately chi-square variables using parametric bootstrap to obtain the critical values of the test. The third proposal is based on applying the Anderson-Darling test directly on the sample, and using Parametric Bootstrap to calculate the critical values of the test. Finally, the fourth proposal is based on a transformation from skew normal variables to approximately normal variables. For testing the hypothesis of normality on the transformed data, the Shapiro-Wilk test is used. Results of simulation studies are presented for the estimation of the size and power of the tests under diverse alternative distributions, which were chosen due to their characteristics and importance in the modeling of data with asymmetry. A comparison was made with other existing tests, the results show that the proposed tests have greater power against some skewed heavy tailed alternative distributions.

Key words: Data transformations; Parametric Bootstrap; skew distributions; Shapiro-Wilk test, Anderson-Darling test.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico brindado durante la realización de mis estudios de maestría.

Al Colegio de Postgraduados y al posgrado en Estadística, por brindarme la oportunidad de continuar mi formación académica, dándome las facilidades necesarias para concluir satisfactoriamente una maestría en ciencias.

A la Dra. Elizabeth González Estrada, por la dedicación de tiempo, esfuerzo y paciencia al realizar esta investigación, así como su guía y consejos durante mi estancia en esta maestría.

A los doctores José Villaseñor Alva y Francisco José Zamudio Sánchez por sus valiosas contribuciones a este trabajo y todos los conocimientos impartidos a lo largo de mi formación.

A mis profesores del Colegio de Postgraduados por compartir conmigo parte de los conocimientos teóricos y prácticos que poseen para hacer mi formación más completa.

DEDICATORIA

A mis padres **Florinda Martínez Sánchez y Francisco Javier Cosmes**
por su cariño y apoyo incondicional.

A mis **hermanos** por su compañía y apoyo durante toda mi vida.

A toda mi **familia** por todo lo que cada uno ha contribuido en mi
desarrollo y crecimiento como persona.

CONTENIDO

RESUMEN	iv
ABSTRACT	v
LISTA DE TABLAS	xii
LISTA DE FIGURAS	xiv
1. Introducción	1
2. Objetivos	4
2.1. General	4
2.2. Específicos	4
3. Marco teórico	5
3.1. Pruebas de hipótesis	5
3.1.1. Prueba de Anderson-Darling	7
3.1.2. Prueba de Shapiro-Wilk	8

CONTENIDO

3.1.3. Bootstrap	10
3.2. Distribución normal asimétrica	10
3.2.1. Propiedades de la distribución normal asimétrica	13
3.2.2. Otros trabajos sobre pruebas de hipótesis para la distribución normal asimétrica	15
3.3. Métodos de estimación	18
3.3.1. Máxima verosimilitud	18
3.3.2. Estimadores de máxima verosimilitud modificados	19
3.3.3. Estimadores de Momentos	20
3.3.4. Estimadores de máxima verosimilitud penalizados	22
3.3.5. Otros estimadores	24
4. Metodologías propuestas	25
4.1. Prueba A: Transformación de variables aleatorias normales asimétricas a va- riables aleatorias χ^2	26
4.2. Prueba B: Anderson-Darling para la distribución normal asimétrica usando bootstrap	30
4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a va- riables aleatorias normales	32
5. Estudio de simulación, resultados y discusión	38
5.1. Obtención del tamaño y potencia de la prueba	38
5.2. Otras pruebas para la distribución SN	45

CONTENIDO

5.3. Tamaño estimado de prueba	47
5.4. Potencia estimada de las pruebas	52
6. Aplicación	61
7. Conclusiones	65
Referencias	65
Anexos	70
Tamaño de las pruebas para $n = 50$ y $n = 500$	70
Potencia de las pruebas conforme n crece	72
Códigos en el programa R	76
Datos utilizados en la Sección Aplicación	83

LISTA DE TABLAS

4.1. Valores críticos para la prueba A (1) para distintos tamaños de muestra (n) y diferentes tamaños de prueba (α)	29
5.1. Tamaño estimado de las pruebas para diferentes valores del parámetro de forma, correspondiente a muestras aleatorias de tamaño n=100 y $\alpha = 0.05$. .	51
5.2. Potencia estimada de las pruebas para muestras de tamaño n=50, $\alpha = 0.05$.	53
5.3. Potencia estimada de las pruebas para muestras de tamaño n=100, $\alpha = 0.05$.	55
5.4. Potencia estimada de las pruebas para muestras de tamaño n=300, $\alpha = 0.05$.	57
5.5. Potencia estimada de las pruebas para muestras de tamaño n=500, $\alpha = 0.05$.	58
5.6. Número de alternativas en donde las pruebas estudiadas tienen la mayor potencia.	59
6.1. Descriptivos básicos de datos de diámetros para cuatro edades distintas . . .	62
6.2. Parámetros estimados de las distribuciones SN y Weibull3 para datos de diámetro normal a los 2.5, 4.7, 6 y 12 años.	63
6.3. Pruebas de bondad de ajuste para la distribución normal asimétrica aplicada a datos de diámetro normal para cuatro edades distintas	64
A.1. Tamaño estimado de las pruebas propuestas con n= 50 y $\alpha=0.05$	70

LISTA DE TABLAS

A.2. Tamaño estimado de las pruebas propuestas con $n= 500$ y $\alpha=0.05$	71
A.3. Potencia estimada de la prueba A (1) con $n= 50, 100, 300, 500$ y $\alpha=0.05$. . .	72
A.4. Potencia estimada de la prueba A (2) con $n= 50, 100, 300, 500$ y $\alpha=0.05$. . .	73
A.5. Potencia estimada de la prueba B con $n= 50, 100, 300, 500$ y $\alpha=0.05$	74
A.6. Potencia estimada de la prueba C con $n= 50, 100, 300, 500$ y $\alpha=0.05$	75
A.7. Diámetros de árboles de una plantación comercial de eucalipto para cuatro edades diferentes	83

LISTA DE FIGURAS

4.1. Distribución nula de la estadística de Anderson-Darling usada en la prueba A para distintos valores de λ	27
4.2. Distribución nula de la estadística de Anderson-Darling usada en la Prueba B para distintos valores de λ	31
4.3. Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud y muestras de tamaño $n=50$	35
4.4. Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud penalizados y muestras de tamaño $n=50$	35
4.5. Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud y muestras de tamaño $n=500$	36
4.6. Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud penalizados y muestras de tamaño $n=500$	37
5.1. Tamaño estimado de la prueba A(1) para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$	48
5.2. Tamaño estimado de la prueba A (2) para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$	48

LISTA DE FIGURAS

5.3. Tamaño estimado de la prueba B para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$	49
5.4. Tamaño estimado de la prueba C para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$	50
6.1. Ajuste de una distribución Weibull3 y de una distribución SN al diámetro normal a los 2.5, 4.7, 6 y 12 años.	62
6.2. Ajuste de la función de distribución acumulada de una distribución Weibull3 y de una distribución SN al diámetro normal a los 2.5, 4.7, 6 y 12 años.	63

Capítulo 1

Introducción

La familia normal asimétrica (SN por sus siglas en inglés) fue introducida de manera formal por [Azzalini \(1985\)](#) como una generalización de la familia normal. Además de los parámetros de localidad y escala, esta familia incluye un parámetro de forma, lo cual le permite ser conformada por una amplia variedad de densidades asimétricas.

La distribución normal asimétrica tiene soporte sobre todos los valores de la recta real al igual que su parámetro de forma, el cual le permite una amplia flexibilidad. Además, esta distribución posee muchas de las propiedades estudiadas en la distribución normal, las cuales son bastante atractivas en el caso de las aplicaciones, ([Pérez-Rodríguez y Villaseñor, 2010](#)). En este sentido, esta distribución es una alternativa viable cuando se pretende analizar datos que presentan un comportamiento normal pero con cierta asimetría. Sin embargo, no se debe perder de vista que esta distribución sirve para modelar datos que pueden ser simétricos o no, por lo que sus aplicaciones son diversas en diferentes áreas, ([Meintanis, 2007](#)).

[Gupta *et al.* \(2004\)](#) presenta una caracterización de esta distribución que incluye sus principales propiedades probabilísticas así como sus semejanzas con las propiedades conocidas para la distribución normal. Además, debido al interés generado alrededor de esta distribución, [Azzalini y Dalla-Valle \(1996\)](#) realizaron una extensión al caso multivariado, el cual también ha sido ampliamente estudiado por diversos autores como [Azzalini y Capitanio \(1999\)](#), [Arnold *et al.* \(2002\)](#) y [Arellano-Valle y Azzalini \(2008\)](#) entre otros.

Existen además otros dos trabajos que concentran mucha de la teoría matemática y estadística sobre esta distribución, los cuales fueron escritos por [Genton \(2004\)](#) y [Azzalini y](#)

1. Introducción

[Capitanio \(2014\)](#). Estos escritos recopilan los avances en el desarrollo teórico no solo de esta distribución sino de toda la línea de investigación relacionada con distribuciones asimétricas. Al mismo tiempo muestran ejemplos ilustrativos de aplicaciones sobre datos reales, lo que muestra el alcance final de la teoría.

[Figueiredo y Gomes \(2013\)](#) presentan un ejemplo de aplicación de la distribución SN en control de calidad, donde resalta la importancia del ajuste de los datos al modelo correcto, (en su caso normal o SN). Los datos son medidas de longitud de tapones de corcho. Para probar si los datos provenían de una distribución normal o normal asimétrica, se realizaron una prueba gráfica y posteriormente la prueba de Shapiro-Wilk para normalidad y la prueba de Kolmogorov-Smirnov para probar la distribución SN. El resultado fue que, aunque los datos parecen tener un ajuste normal gráficamente, su distribución se ajusta mejor al comportamiento de una normal asimétrica.

Otro trabajo donde se hace uso de esta distribución es el presentado por [Vernic \(2006\)](#), en el que es de interés comprobar si el riesgo y el reparto de capital siguen un comportamiento normal o por el contrario son mejor ajustados por una SN. El trabajo utiliza estas dos variables para modelar el riesgo en seguros.

[Azzalini y Capitanio \(1999\)](#) extienden el uso de la distribución SN al caso multivariado y mencionan que dicha distribución presenta bastante flexibilidad en el caso de ajuste a datos reales. Particularmente la teoría asociada a formas lineales y cuadráticas para la distribución normal es en gran medida válida para la SN multivariada.

[Meintanis \(2007\)](#) presenta un ejemplo de aplicación en datos médicos colectados en 1971, cuyo objetivo era estudiar el IQ de hombres utilizando la distribución normal asimétrica.

Otro conjunto de datos que se ha utilizado para ejemplificar el uso de pruebas de bondad de ajuste es el que presentan [Gupta y Chen \(2001\)](#), que consiste en modelar la edad a la que gemelos humanos contraen un resfriado por primera vez.

[Kazemi y Noorizadeh \(2015\)](#) utilizan datos de pérdidas por indemnizaciones en Estados Unidos y realizan pruebas de bondad de ajuste utilizando los estadísticos de Kolmogorov-Smirnov, criterio de log-verosimilitud y el AIC.

Sin embargo, se debe considerar un paso anterior a las aplicaciones, ya que es necesario tener la mayor certeza posible que los datos provienen de una distribución normal asimétrica

1. Introducción

pues de ello dependerán las metodologías que se utilicen en su análisis y la fiabilidad de los resultados obtenidos. Las pruebas de bondad de ajuste son una herramienta útil en este sentido.

Existen trabajos enfocados a la distribución normal asimétrica los cuales están basados en algunas de sus propiedades y abarcan un amplio espectro de hipótesis, desde la hipótesis simple de contrastar si una muestra aleatoria proviene de una distribución normal asimétrica con parámetros conocidos, hasta la hipótesis compuesta donde la hipótesis alternativa no es especificada y los parámetros de la distribución en la nula son desconocidos.

Trabajos como los de [Gupta y Chen \(2001\)](#), utilizan tablas de la distribución acumulada de la distribución normal asimétrica para distintos valores del parámetro de forma para implementar dos pruebas conocidas, kolmogorov-Smirnov y la prueba χ^2 asumiendo que se conoce el parámetro de forma, es decir que la distribución en la hipótesis nula está completamente especificada. Otra propuesta realizada por [Dalla-Valle \(2007\)](#) utiliza herramientas computacionales para generar tablas de la distribución con distintos tamaños de muestra, lo anterior para probar hipótesis compuestas. También se han desarrollado trabajos como los de [Meintanis \(2007\)](#), [Mateu-Figueras *et al.* \(2007\)](#), [Pérez-Rodríguez y Villaseñor \(2010\)](#) y [Sariso *et al.* \(2014\)](#) quienes proponen pruebas para el caso en donde los parámetros son desconocidos y toda la información que se posee debe ser extraída de la muestra. Utilizan distintos estimadores como los obtenidos por el método de momentos o máxima verosimilitud y recurren a métodos Monte Carlo para la obtención de constantes críticas.

El presente trabajo se centra en desarrollar una prueba de bondad de ajuste para la hipótesis nula de que una muestra aleatoria proviene de una población normal asimétrica con parámetros desconocidos. Se hacen cuatro propuestas que hacen uso de las estadísticas de Anderson-Darling y de Shapiro-Wilk, dos de las cuales utilizan bootstrap paramétrico y en cuanto a tiempo de ejecución son comparables con las metodologías de otros autores. Las otras dos propuestas no utilizan herramientas Monte Carlo al momento de evaluar las observaciones de una muestra aleatoria y el tiempo de cómputo para su ejecución es menor que para otras pruebas.

Además, se presenta el método de estimación utilizado para estimar los parámetros y la implementación del bootstrap para realizar la aproximación de la distribución del estadístico de prueba de dos de ellas. Para evaluar el desempeño de las propuestas se presenta un estudio de su tamaño y potencia, obtenido mediante simulación Monte Carlo.

Capítulo 2

Objetivos

2.1. General

- Proponer pruebas de bondad de ajuste para probar si las observaciones de una muestra aleatoria provienen de una distribución normal asimétrica univariada cuando los parámetros son desconocidos.

2.2. Específicos

- Investigar si algunas propiedades intrínsecas de la distribución normal asimétrica permiten o ayudan a obtener una prueba eficiente para discernir si una muestra aleatoria proviene o no de dicha distribución.
- Estudiar mediante simulación Monte Carlo el nivel y la potencia de la prueba para distintos escenarios de interés.
- Realizar un análisis comparativo sobre el nivel y la potencia de la prueba propuesta contra otras pruebas existentes bajo diferentes distribuciones alternativas.

Capítulo 3

Marco teórico

3.1. Pruebas de hipótesis

En esta sección se muestran algunos conceptos de pruebas de hipótesis que son útiles para el desarrollo de la metodología propuesta.

Una hipótesis estadística es una aseveración o conjetura sobre la distribución de una o más variables aleatorias. Si la hipótesis estadística especifica completamente la distribución de la que puede provenir una muestra aleatoria, entonces es llamada hipótesis simple, de otro modo es llamada hipótesis compuesta.

¿Qué son las pruebas de hipótesis?

La idea general de una prueba de hipótesis es que consiste en una regla específica con la cual se decide rechazar o no rechazar una afirmación acerca de una población sujeta a la información proporcionada por una muestra aleatoria.

Dada $\underline{X} = (X_1, X_2, \dots, X_n)$ una muestra aleatoria de una distribución $F_X(x)$, se desea probar un juego de hipótesis H_0 vs H_a para alguna hipótesis nula y alterna H_0 y H_a determinadas.

Definición: Sea χ el conjunto de realizaciones posibles de \underline{X} . Entonces una prueba es un subconjunto (cualquiera) χ_R de χ llamado región de rechazo tal que si una realización \underline{x} de \underline{X} es tal que $\underline{x} \in \chi_R$ entonces se decide rechazar H_0 . Si $\chi_A = \chi - \chi_R$ y $\underline{x} \in \chi_A$ entonces se

3.1. Pruebas de hipótesis

decide no rechazar H_0 .

Una función de prueba $\phi : \chi \rightarrow \{0, 1\}$ es una función tal que

$$\phi(\underline{x}) = \begin{cases} 1 & \text{si } \underline{x} \in \chi_R \\ 0 & \text{si } \underline{x} \notin \chi_R. \end{cases}$$

Entonces si $\phi(\underline{x}) = 1$ se rechaza H_0 . Por tanto, a cada función $\phi : \chi \rightarrow \{0, 1\}$ le corresponde una prueba e inversamente. Es decir, la prueba ϕ asociada a dicha función rechaza H_0 si $\phi(\underline{x}) = 1$ y no rechaza si $\phi(\underline{x}) = 0$.

Generalmente se utiliza una estadística para definir una prueba. Sea la estadística $T = t(\underline{X})$ con τ el conjunto de las realizaciones posibles de T . Se define la prueba con base en T : $\tau_R = \{t \in R \subset \mathbb{R}\}$ es la región de rechazo de la prueba y $\tau_A = \{t \in R^c \subset \mathbb{R}\}$. Donde R depende de la estadística T . En este caso se dice que T es la estadística de prueba.

Al tomar una decisión con base en una prueba estamos expuestos a cometer dos tipos de errores.

Error tipo I: Rechazar H_0 cuando es verdadera usando una prueba ϕ .

Error tipo II: Aceptar H_0 cuando es falsa usando una prueba ϕ .

Definición: La función de potencia de una prueba ϕ , denotada por β_ϕ , es tal que $\beta_\phi : \Omega \rightarrow [0, 1]$ donde:

$$\beta_\phi = P(\phi(\underline{x}) = 1 \mid H_0 \text{ es verdadera}).$$

Definición: El tamaño de la prueba ϕ para probar H_0 vs H_a se define como

$$\text{máx}(\beta_\phi) = \alpha^*$$

Definición: La prueba ϕ es de nivel α si $\alpha^* \leq \alpha$.

El concepto de bondad de ajuste hace referencia a lo bien que un modelo estadístico emula el comportamiento de las observaciones de una muestra. Existen diversas medidas de ajuste, las cuales son utilizadas en la creación de pruebas de bondad de ajuste.

3.1. Pruebas de hipótesis

Definición: Una prueba de bondad de ajuste es un procedimiento estadístico para contrastar la hipótesis nula de que una muestra aleatoria proviene de una distribución dada contra la hipótesis alternativa de que no proviene de dicha distribución.

Sea Y_1, Y_2, \dots, Y_n una muestra aleatoria de una población con alguna distribución F_Y desconocida. Una prueba de bondad de ajuste contrasta el siguiente tipo de juego de hipótesis:

$$H_0 : F_Y(y) = F(y)$$

vs

$$H_a : F_Y(y) \neq F(y),$$

donde F es alguna distribución conocida.

Como en cualquier prueba de hipótesis, el uso de diversas pruebas estadísticas es posible. [D' Agostino y Stephens \(1986\)](#) realizan un agrupamiento de algunas de éstas, basado en sus características. Las estadísticas de tipo Cramér-von Mises, que contienen a las estadísticas de Cramér-von Mises, de Watson y de Anderson-Darling, y las estadísticas de tipo Kolmogorov-Smirnov que contiene a la estadística de Kolmogorov-Smirnov y de Kuiper entre otras.

3.1.1. Prueba de Anderson-Darling

Dentro del grupo de pruebas que utilizan la función de distribución empírica de los datos, se encuentra la de Anderson-Darling (AD), utilizada para probar si una muestra proviene de una distribución con una forma conocida $F_0(x; \underline{\theta})$ donde $\underline{\theta}$ es el vector de parámetros asociado a dicha distribución y puede ser conocida o desconocida.

Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución $F(x)$, se desea probar el siguiente juego de hipótesis

$$H_0 : F(x) = F_0(x; \underline{\theta})$$

vs

$$H_a : F(x) \neq F_0(x; \underline{\theta}).$$

3.1. Pruebas de hipótesis

La prueba de Anderson-Darling se basa en la estadística AD, también conocida como A^2 , la cual tiene la siguiente definición:

$$AD = n \int_{-\infty}^{\infty} \frac{[F_n(x) - F_0(x, \underline{\theta})]^2}{F_0(x, \underline{\theta}) [1 - F_0(x, \underline{\theta})]} dF_0(x, \underline{\theta}) \quad (3.1)$$

Esta prueba rechaza la hipótesis nula H_0 para un tamaño de prueba $\alpha \in (0, 1)$ si la estadística AD es mayor que $k_{1-\alpha}$, donde $F_n(x)$ es la función de distribución empírica de las x'_i s, $F_0(x, \underline{\theta})$ es la función de distribución acumulada de las x'_i s bajo la hipótesis nula y $k_{1-\alpha}$ es el cuantil $(1 - \alpha)100\%$ de la distribución de AD bajo la hipótesis nula.

Para llevar a cabo el cálculo de la estadística, se utiliza la siguiente fórmula:

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

donde $x_{(i)}$ es la i -ésima estadística de orden de la muestra aleatoria de tamaño n , ([D' Agostino y Stephens, 1986](#)).

$F_0(x; \underline{\theta})$ puede ser cualquier distribución de interés y el vector de parámetros puede estar completamente especificado. Cuando esto ocurre, la distribución de la estadística de prueba bajo H_0 es independiente de $F_0(x; \underline{\theta})$. Cuando el vector de parámetros es desconocido al menos en un componente, la distribución de la estadística de prueba bajo H_0 depende del método de estimación utilizado para obtener dichos parámetros, de las herramientas numéricas, del tamaño de muestra y de la distribución que se probará.

Existen diversas adecuaciones para hacer que el cálculo de la estadística sea más preciso tal es el caso de [Marsaglia y Marsaglia \(2004\)](#), trabajo en el que se basa el paquete “*ADGofTest*” ([Bellosta, 2011](#)), de R para llevar a cabo la prueba de Anderson-Darling.

3.1.2. Prueba de Shapiro-Wilk

La prueba de Shapiro-Wilk es una de las más utilizadas para contrastar la hipótesis de normalidad. Esta prueba fue propuesta por [Shapiro y Wilk \(1965\)](#) y se muestra a continuación una breve descripción.

3.1. Pruebas de hipótesis

Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución $F(x)$, se desea probar el siguiente juego de hipótesis

$$H_0 : F(x) = N(\mu, \sigma^2) \quad (3.2)$$

vs

$$H_a : F(x) \neq N(\mu, \sigma^2)$$

donde $N(\mu, \sigma^2)$ denota una distribución normal con parámetros μ y σ^2 desconocidos.

Sea $m' = (m_1, m_2, \dots, m_n)$ el vector de esperanzas de las estadísticas de orden de una muestra aleatoria de variables normal estándar, y sea $V = (v_{ij})$ la matriz de covarianzas correspondiente.

Sea $x' = (x_{(1)}, x_{(2)}, \dots, x_{(n)})$ una realización ordenada de la muestra aleatoria, entonces la estadística de prueba W está dada por la siguiente ecuación:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2} \quad (3.3)$$

donde

$$a' = (a_1, a_2, \dots, a_n) = \frac{m'v^{-1}}{(m'v^{-1}v^{-1}m)^{1/2}}$$

[Shapiro y Wilk \(1965\)](#) construyeron tablas de valores de a' para distintos tamaños de muestra n así como tablas de valores críticos de la prueba para distintos niveles de significancia y tamaños de muestra.

Se rechaza la hipótesis nula en (3.2) si el valor calculado de la estadística de prueba W es menor que el valor crítico correspondiente al tamaño de muestra para un nivel de significancia $\alpha \in [0, 1]$.

Algunos autores como [Patrício et al. \(2017\)](#) y [Romao et al. \(2010\)](#) han realizado estudios comparativos de diferentes pruebas de bondad de ajuste para normalidad en donde el hallazgo general es que la prueba de Shapiro-Wilk resulta ser más potente y más sensible a alternativas asimétricas que pruebas como el tercer y cuarto momento estándares, Kolmogorov-Smirnov, Cramer-Von Mises (que incluye la prueba de Anderson-Darling), Lilliefors, Shapiro-Francia,

3.2. Distribución normal asimétrica

Jarque-Bera, chi-cuadrada y de rango estudentizado.

3.1.3. Bootstrap

Como se mencionó anteriormente, existen muchas pruebas que están basadas en la función de distribución empírica. Sin embargo, muchas veces los parámetros involucrados en el cálculo de la estadística de prueba no están completamente especificados. Por lo anterior, comúnmente se recurre a la obtención de estimadores. En estos casos, la distribución de la estadística de prueba no es factible de obtener ya que existe dependencia en las estimaciones obtenidas y los tamaños de muestra utilizados, de tal manera que es necesario obtener dicha distribución de manera aproximada.

Una alternativa para aproximar la distribución de la estadística de prueba bajo la hipótesis nula, es el uso del método bootstrap paramétrico.

Sea $X_1^*, X_2^*, \dots, X_n^*$ una muestra aleatoria de \hat{F} donde \hat{F} es una estimación paramétrica de la función de distribución F basada en la muestra X_1, X_2, \dots, X_n . Entonces $X_1^*, X_2^*, \dots, X_n^*$ es llamada muestra bootstrap. A través de la obtención de B muestras bootstrap, es posible calcular B valores de la estadística de prueba bajo la hipótesis nula, los cuales son utilizados para aproximar la distribución de dicha estadística de prueba bajo la hipótesis nula.

[Babu y Rao \(2004\)](#) han demostrado que, bajo condiciones bastante generales, la operación resultante de restar la función de distribución teórica con parámetros estimados de la distribución empírica, tiene el mismo límite débil que restar la distribución obtenida bajo el método bootstrap paramétrico de la distribución empírica. Esta afirmación es bastante útil para el desarrollo del presente trabajo ya que, de hecho asegura la validez del método bootstrap, es decir, que dicho método estima consistentemente las distribuciones nulas de varias pruebas de bondad de ajuste, entre las que destacan las pruebas de tipo Cramér-von Mises y Kolmogorov-Smirnov.

3.2. Distribución normal asimétrica

La familia de distribuciones normal asimétrica está constituida por distribuciones que contienen un parámetro de asimetría λ (cuyo valor es cero en distribuciones normales). Esta dis-

3.2. Distribución normal asimétrica

tribución es un caso particular de una clase de distribuciones de la forma $f(x) = 2\varphi(x)\phi(\lambda x)$ donde $\varphi(\cdot)$ es alguna función de densidad simétrica alrededor de cero y $\phi(x)$ es la función de distribución acumulada de $\varphi(\cdot)$.

Esta distribución tiene parámetros de localidad y escala. Se puede verificar que, cuando el parámetro de forma es igual a cero, se trata de una distribución normal y de la distribución mitad-normal cuando dicho parámetro tiende a infinito. Además, el valor absoluto del coeficiente de asimetría incrementa conforme el valor absoluto del parámetro de asimetría incrementa. La distribución es asimétrica a la derecha si el parámetro de asimetría es mayor que cero y a la izquierda si el parámetro de asimetría es menor que cero.

En las últimas tres décadas, esta familia de distribuciones ha sido ampliamente estudiada debido a que muchas de sus propiedades coinciden con las de la familia normal, lo que amplía su utilidad a diversas áreas de aplicación.

En la literatura se mencionan diversos estudios en donde aparece esta familia, [Pourahmadi \(2000\)](#) señala que la primera vez que apareció la distribución normal asimétrica fue en [Birnbaum \(1950\)](#) en un estudio sobre pruebas educativas, seguido de los trabajos de [Weinstein \(1964\)](#) y [Nelson \(1964\)](#) quienes se centraron en encontrar la distribución de la suma de una variable aleatoria normal y una variable normal truncada.

Otros autores mencionados como pioneros del estudio de esta familia son [Roberts \(1966\)](#), [O'hagan y Leonard \(1976\)](#) y [Azzalini \(1985\)](#). Este último autor fue quien estudió sus principales propiedades, la estimación de sus parámetros y realizó investigaciones sobre la distribución normal asimétrica multivariada.

Definición: Se dice que la variable aleatoria Z tiene distribución normal asimétrica (SN) si su función de densidad está dada por:

$$f(z) = 2\phi(z)\Phi(\lambda z)I_{(-\infty, \infty)}(z), \quad (3.4)$$

donde $\phi(\cdot)$ y $\Phi(\cdot)$ son las funciones de densidad y de distribución Normal estándar respectivamente y $\lambda \in \mathbb{R}$.

Además, cuando se involucran los parámetros de localidad y escala surge la siguiente definición.

Definición: Una variable aleatoria $X = \xi + \omega Z$ tiene distribución normal asimétrica (SN)

3.2. Distribución normal asimétrica

con parámetro de localidad ξ , de escala ω y forma λ , si su función de densidad está dada por:

$$f(x) = \frac{2}{\omega} \phi\left(\frac{x-\xi}{\omega}\right) \Phi\left(\lambda\left(\frac{x-\xi}{\omega}\right)\right) I_{(-\infty, \infty)}(x), \quad (3.5)$$

donde $\xi \in \mathbb{R}$ y $\omega \in \mathbb{R}^+$.

Note que si $\xi = 0$ y $\omega = 1$ entonces 3.5 se reduce a 3.4, la cual es conocida como distribución normal asimétrica estándar.

La función en 3.5 es una función de densidad y es fácil verificarlo utilizando el siguiente Lema de Azzalini (1985).

Lema. Sea f una función de densidad simétrica alrededor de 0, sea G una distribución absolutamente continua tal que G' es simétrica alrededor de 0. Entonces $2 f(y)G(\lambda y)$ con $-\infty < y < \infty$, es una función de densidad para cualquier λ .

La función de distribución acumulada de una variable aleatoria $X \sim SN(\xi, \omega, \lambda)$ está definida como:

$$F(x; \xi, \omega, \lambda) = \frac{2}{\omega} \phi\left(\frac{x-\xi}{\omega}\right) T\left(\frac{x-\xi}{\omega}, \lambda\right) I_{(-\infty, \infty)}(x), \quad (3.6)$$

donde $\xi, \lambda \in \mathbb{R}$ y $\omega \in \mathbb{R}^+$. $T(h, b)$ es la función T de Owen (integral de la densidad normal estándar bivariada delimitada por $x = h$, $y = 0$ y $y = bx$).

La función generadora de momentos de una variable aleatoria $SN(\xi, \omega, \lambda)$ está dada por la siguiente ecuación:

$$M_X(t) = 2 \exp\left(\xi t + \frac{\omega^2 t^2}{2}\right) \Phi\left(\frac{\omega \lambda t}{\sqrt{1 + \lambda^2}}\right), t \in \mathbb{R}. \quad (3.7)$$

La esperanza y la varianza pueden obtenerse de la ecuación anterior y están dadas por las siguientes ecuaciones respectivamente:

$$E(X) = \xi + \frac{\omega \lambda}{\sqrt{1 + \lambda^2}} \sqrt{\frac{2}{\pi}}. \quad (3.8)$$

$$Var(X) = \omega^2 \left(1 - \frac{2\lambda^2}{\pi(1 + \lambda^2)}\right). \quad (3.9)$$

3.2. Distribución normal asimétrica

Los coeficientes de asimetría y curtosis son los siguientes respectivamente:

$$\gamma_1\{X\} = \gamma_1\{Z\} = \frac{4 - \pi}{2} \frac{\left[\sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1+\lambda^2}} \right]^3}{\left[1 - \frac{2}{\pi} \left(\frac{\lambda}{\sqrt{1+\lambda^2}} \right)^2 \right]^{3/2}}. \quad (3.10)$$

$$\gamma_2\{X\} = \gamma_2\{Z\} = 2(\pi - 3) \frac{\left[\sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1+\lambda^2}} \right]^4}{\left[1 - \frac{2}{\pi} \left(\frac{\lambda}{\sqrt{1+\lambda^2}} \right)^2 \right]^2}, \quad (3.11)$$

en donde $-\sqrt{2}(4 - \pi)/(\pi - 2)^{3/2} < \gamma_1 < \sqrt{2}(4 - \pi)/(\pi - 2)^{3/2} \approx 0.9953$ y $0 \leq \gamma_2 < 8(\pi - 3)/(\pi - 2)^2 \approx 0.869$.

3.2.1. Propiedades de la distribución normal asimétrica

Note que, si $Y \sim SN(\xi, \omega, \lambda)$ entonces $X = \frac{Y - \xi}{\omega} \sim SN(\lambda)$.

Algunas propiedades básicas de la distribución $SN(\lambda)$ dada por [Azzalini \(1985\)](#) son:

1. $SN(0) = N(0, 1)$.
2. Si $Z \sim SN(\lambda)$ entonces $-Z \sim SN(-\lambda)$.
3. Cuando $\lambda \rightarrow \pm\infty$, $SN(\lambda)$ tiende a la distribución mitad-normal, es decir, la distribución de $\pm |X|$ cuando $X \sim N(0, 1)$.
4. Si $Z \sim SN(\lambda)$ entonces $Z^2 \sim \chi_1^2$.
5. Si U y V son variables aleatorias independientes e idénticamente distribuidas $N(0,1)$ entonces

$$\frac{\lambda}{\sqrt{1 - \lambda^2}} |U| + \frac{\lambda}{\sqrt{1 + \lambda^2}} V \sim SN(\lambda), \text{ (Henze, 1986).}$$

6. Sea (X_1, X_2) un vector aleatorio normal bivariado entonces

$$(X_1 | X_2 > 0) \sim SN(\lambda),$$

donde $\lambda = \frac{\rho}{\sqrt{1 - \rho^2}}$.

3.2. Distribución normal asimétrica

7. Si X_1 y X_2 son variables aleatorias independientes $N(0,1)$, entonces

$$(X_1 \mid \lambda X_1 > 0) \sim SN(\lambda), \text{ (Azzalini y Dalla-Valle, 1996).}$$

8. Sea $X \sim SN(\xi, \omega, \lambda)$ entonces $Y = |X - \xi|$ tiene una distribución mitad-normal (HN) con parámetros $(0, \omega)$, denotada como $HN(0, \omega)$.

9. La distribución SN está únicamente determinada por su secuencia de momentos. Es decir, cuando conocemos los momentos de todo orden de una distribución digamos F y éstos coinciden con los de la SN, entonces F es la distribución SN.

10. Sean X_1, X_2 variables aleatorias *i.i.d.* provenientes de F , una distribución no especificada que admite la obtención de sus momentos de todo orden. Entonces $X_1^2 \sim \chi_1^2, X_2^2 \sim \chi_1^2$ y $\frac{1}{2} \cdot (X_1 + X_2)^2 \sim G_0(\lambda)$ si y solo si $F = SN(\lambda)$ o $F = SN(-\lambda)$ donde $G_0(\lambda)$ es la distribución de $\frac{1}{2} \cdot (X + Y)^2$ cuando X y Y son *i.i.d.* $SN(\lambda)$.

11. Sea $G_0(\lambda)$ la distribución de $(Y + a)^2$ donde $Y \sim SN(\lambda)$ y $a \neq 0$ una constante dada. Sea X una variable aleatoria que admite la obtención de momentos de todo orden. Entonces $X^2 \sim \chi_1^2$ y $(X + a)^2 \sim G_0(\lambda)$ si y solo si $X \sim SN(\lambda)$ para alguna λ .

Con el siguiente lema se puede verificar fácilmente la propiedad 4.

Lema (Roberts, 1966) $W^2 \sim \chi_1^2$ si y solo si la f.d.p. de W tiene la forma $f(w) = h(w) \exp\left(\frac{-w^2}{2}\right)$ donde $h(w) + h(-w) = \sqrt{\frac{2}{\pi}}$.

Una demostración de la propiedad 8. es la siguiente. Note que

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(|X - \xi| \leq y) = P(-y \leq X - \xi \leq y) \\ &= P(-y + \xi \leq X \leq y + \xi) = F_X(y + \xi) - F_X(-y + \xi) \end{aligned}$$

3.2. Distribución normal asimétrica

$$\begin{aligned}\therefore f_Y(y) &= \frac{dF_Y(y)}{dy} = f_X(y + \xi) + f_X(-y + \xi) \\ &= \frac{2}{\omega} \phi\left(\frac{y}{\omega}\right) \Phi\left(\lambda\left(\frac{y}{\omega}\right)\right) + \frac{2}{\omega} \phi\left(-\frac{y}{\omega}\right) \Phi\left(\lambda\left(-\frac{y}{\omega}\right)\right) \\ &= \frac{2}{\omega} \phi\left(\frac{y}{\omega}\right) \left(\Phi\left(\lambda\left(\frac{y}{\omega}\right)\right) + \left(1 - \Phi\left(\lambda\left(\frac{y}{\omega}\right)\right)\right)\right) \\ &= \frac{2}{\omega} \phi\left(\frac{y}{\omega}\right) = \frac{2}{\omega} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{y}{\omega}\right)^2\right) = \frac{\sqrt{2}}{\omega\sqrt{\pi}} \exp\left(-\frac{y^2}{2\omega^2}\right),\end{aligned}$$

la cual es la función de densidad de una variable aleatoria $HN(0, \omega)$.

El uso de la propiedad 4, fue ampliamente estudiado por [Roberts \(1966\)](#), dicha propiedad tiene implicaciones directas, por ejemplo, que la distribución de $|X|$, X^2 y cualquier función par de X no depende del parámetro de asimetría λ .

Por último, como mencionan [Figueiredo y Gomes \(2013\)](#), se debe considerar que esta familia de distribuciones posee propiedades que proporcionan mayor amplitud a los métodos estadísticos en cuanto a las características que una muestra debe tener para utilizarlos, ya que, como se mencionó, emula muchas de las propiedades de la familia normal y es una extensión de ésta, pero puede modelar mejor datos cuya distribución empírica tiene una forma similar a la normal, y además exhiben una ligera asimetría. Incluso cuando se piensa que indudablemente los datos son normalmente distribuidos, existe la posibilidad de alteraciones, y la familia de distribuciones normal asimétrica puede describir los datos del proceso de una manera más fiable y sólida.

3.2.2. Otros trabajos sobre pruebas de hipótesis para la distribución normal asimétrica

Existen diversas propuestas sobre pruebas de hipótesis que se enfocan en esta distribución, las cuales pueden utilizar los métodos convencionales como son las tablas de distribución acumulada, o métodos computacionales conocidos como Monte Carlo. La mayor parte de estos trabajos abordan el caso en donde los parámetros de la distribución de la muestra son desconocidos y por lo tanto también se han propuesto distintos métodos de estimación para su desarrollo.

3.2. Distribución normal asimétrica

Pruebas de hipótesis paramétricas para la distribución normal asimétrica

Uno de los contrastes más sencillos que se presentan es cuando se pretende probar si una muestra proviene de una población con distribución normal o por el contrario proviene de una distribución normal asimétrica.

[Figueiredo y Gomes \(2013\)](#) presentan un trabajo de aplicación, en donde se realizan pruebas de razón de verosimilitudes para contrastar si los datos de una muestra aleatoria provienen de una población con distribución normal o de una población con distribución normal asimétrica.

Se utiliza el siguiente juego de hipótesis:

$$H_0 : X \sim SN(\xi, \omega, \lambda = 0)$$

vs

$$H_1 : X \sim SN(\xi, \omega, \lambda \neq 0)$$

Y la estadística de prueba siguiente:

$$\Lambda = \frac{L_{SN}(\hat{\xi}, \hat{\omega}, \lambda = 0)}{L_{SN}(\hat{\xi}, \hat{\omega}, \hat{\lambda})}$$

Hay que recordar que, bajo la hipótesis nula, $-2\log(\Lambda)$ se distribuye chi-cuadrada.

De la misma manera [Jiménez-Gamero et al. \(2015\)](#) proponen una prueba de hipótesis sobre el parámetro de asimetría de una distribución normal asimétrica, es decir, si el parámetro de asimetría es igual a cero, lo que conduciría a proponer a la distribución normal como modelo del cual provienen los datos. Esta propuesta utiliza la función de distribución tabulada por [Azzalini \(1985\)](#).

Pruebas de bondad de ajuste para la distribución normal asimétrica

Una de las pruebas de bondad de ajuste existentes fue propuesta por [Gupta y Chen \(2001\)](#). Ellos utilizan dos pruebas conocidas, Kolmogorov-Smirnov y χ^2 de Pearson a través de las cuales examinan si la distribución normal asimétrica es un modelo apropiado para un

3.2. Distribución normal asimétrica

conjunto de datos. Para el cálculo de las estadísticas de prueba de Kolmogorov-Smirnov y χ^2 , ellos presentan tablas de distribución acumulada de la distribución normal asimétrica para distintos valores del parámetro de forma λ y los valores críticos para cada prueba son tomados de otros trabajos donde se han desarrollado tablas para la distribución de dichas estadísticas.

Dalla-Valle (2007) utiliza una modificación de la prueba de Anderson-Darling y propone una tabla de valores de significancia a partir de los cuales se rechaza o no la hipótesis de normalidad asimétrica para distintos tamaños de muestra.

Otras pruebas de bondad de ajuste conocidas se presentan en los trabajos de Meintanis (2007) y Mateu-Figueras *et al.* (2007). El primero propone una prueba de tipo Kolmogorov-Smirnov basada en el comportamiento de la función generatriz de momentos empírica, en tanto que el segundo trabajo realiza un estudio de cinco pruebas conocidas, Cramér-von Mises, Watson, Anderson-Darling, Kolmogorov-Smirnov y Kuiper utilizando métodos Monte Carlo en el cálculo de las constantes críticas. Ambos trabajos estiman los parámetros a partir de una muestra aleatoria, el primero con el método de momentos y el segundo con máxima verosimilitud.

Pérez-Rodríguez y Villaseñor (2010) proponen dos pruebas de bondad de ajuste basadas en el coeficiente de correlación y en las propiedades de la distribución normal asimétrica. Ellos obtienen los valores críticos de la prueba mediante métodos Monte Carlo para distintos tamaños de muestra y niveles de significancia. Además, comparan el funcionamiento de éstas con respecto a los dos trabajos expuestos en el párrafo anterior, evidenciando una mejoría respecto a los resultados que ahí se reportan.

Por último el trabajo realizado por Sarisoy *et al.* (2014) utiliza la prueba de Liao y Shimokawa, basada en la estadística de prueba

$$L_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\max \left[\frac{i}{n} - F(x_{(i)}), F(x_{(i)}) - \frac{i-1}{n} \right]}{\sqrt{F(x_{(i)}) [1 - F(x_{(i)})]}}$$

donde $x_{(i)}$ es el i -ésimo estadístico de orden de una muestra aleatoria de tamaño n y $F(\cdot)$ es la función de distribución acumulada bajo la hipótesis nula, la cual mide el promedio de todas las distancias ponderadas sobre todo el rango de valores de x . La prueba rechaza cuando el valor calculado de la estadística de prueba es mayor que el valor crítico correspondiente,

3.3. Métodos de estimación

reportado en su trabajo. También realiza una comparación con [Mateu-Figueras *et al.* \(2007\)](#), estos autores concluyen que la prueba que utilizan es más potente para la hipótesis de que la muestra proviene de una distribución normal asimétrica para algunas distribuciones alternativas específicas.

3.3. Métodos de estimación

3.3.1. Máxima verosimilitud

Respecto a los estimadores de parámetros para la distribución $SN(\xi, \omega, \lambda)$, solo es posible obtener los estimadores de Máxima verosimilitud numéricamente, es decir, estos estimadores no tienen una forma analítica cerrada para la muestra.

El cálculo de dichos estimadores se puede realizar como sigue:

Sea Y_1, Y_2, \dots, Y_n una muestra aleatoria de tamaño n de una distribución $SN(\xi, \omega, \lambda)$. La función de verosimilitud está dada por

$$L(\xi, \omega, \lambda; y) = \left(\frac{2}{\omega}\right)^n \prod_{i=1}^n \phi\left(\frac{y_i - \xi}{\omega}\right) \Phi\left[\lambda\left(\frac{y_i - \xi}{\omega}\right)\right] I_{(-\infty, \infty)}(y).$$

Y la log-verosimilitud está dada por

$$\ln(L(\xi, \omega, \lambda)) = n \ln(2) - n \ln(\omega) + \sum_{i=1}^n \ln\left(\phi\left(\frac{y_i - \xi}{\omega}\right)\right) + \sum_{i=1}^n \ln\left(\Phi\left(\lambda\left(\frac{y_i - \xi}{\omega}\right)\right)\right).$$

De la ecuación anterior, los estimadores de máxima verosimilitud de ξ, ω y λ están dados por la solución del siguiente sistema de ecuaciones

$$\begin{aligned} \sum_{i=1}^n \left(\frac{y_i - \xi}{\omega}\right) - \lambda \sum_{i=1}^n \alpha_i &= 0 \\ \sum_{i=1}^n \left(\frac{y_i - \xi}{\omega}\right)^2 - \lambda \sum_{i=1}^n \alpha_i \left(\frac{y_i - \xi}{\omega}\right) - n &= 0 \end{aligned}$$

3.3. Métodos de estimación

$$\sum_{i=1}^n \alpha_i \left(\frac{y_i - \xi}{\omega} \right) = 0$$

donde

$$\alpha_i = \frac{\phi \left[\lambda \left(\frac{y_i - \xi}{\omega} \right) \right]}{\Phi \left[\lambda \left(\frac{y_i - \xi}{\omega} \right) \right]}$$

Es importante mencionar que para cualquier valor de λ dado, los valores de $\hat{\xi}$ y $\hat{\omega}$ están bien determinados, es decir, las ecuaciones para obtener $\hat{\xi}$ y $\hat{\omega}$ tienen solución única simple.

Sin embargo para valores desconocidos de λ , estos estimadores pueden presentar ciertas dificultades. Como lo menciona [Azzalini \(2005\)](#), para muestras pequeñas e incluso para muestras de tamaño medio o moderado, la función de verosimilitud puede ser monótona creciente o decreciente en λ , lo que trae como consecuencia que el estimador de máxima verosimilitud, $\hat{\lambda}$, esté dado por $\pm\infty$. Además, cuando $\lambda \rightarrow 0$, la matriz de información de Fisher se hace singular por lo que no es posible obtener una estimación única.

Tomando en cuenta lo anterior, un modo de proceder consiste en proponer un valor de λ como valor inicial y resolver las ecuaciones mediante métodos numéricos para obtener los valores de $\hat{\xi}$, $\hat{\omega}$ y $\hat{\lambda}$. Estos métodos por lo general son lentos para alcanzar convergencia, además se recomienda repetir este paso para un rango razonable de valores de λ de tal manera que no haya complicaciones con máximos locales.

Como una alternativa, [Azzalini \(1985\)](#) propone una reparametrización a través de la cual se obtienen los estimadores de máxima verosimilitud.

3.3.2. Estimadores de máxima verosimilitud modificados

La parametrización “centrada” propuesta por [Azzalini \(1985\)](#), consiste en la siguiente transformación, sea $X \sim SN(\xi, \omega, \lambda)$ entonces,

$$Y = \mu + \sigma \left(\frac{X - E(X)}{\sqrt{Var(X)}} \right)$$

tiene una distribución $SN(\mu, \sigma, \gamma_1)$.

3.3. Métodos de estimación

Como resultado de esta re-parametrización $E(Y) = \mu$, $\text{Var}(Y) = \sigma^2$ y γ_1 es el coeficiente de asimetría tanto de X como de Y .

De esta manera se tiene una ventaja, se resuelve el problema de la singularidad de la matriz de información de Fisher en el cálculo de los estimadores de máxima verosimilitud. Además se controla la irregularidad en el comportamiento de la función de verosimilitud.

Hay que recordar que, al trabajar sobre esta parametrización centrada, obtendremos los estimadores de μ, σ y γ_1 . Para obtener los estimadores de los parámetros originales es necesario calcularlos a partir de las siguientes ecuaciones

$$\xi = \mu - \sigma \left(\frac{2\gamma_1}{(4 - \pi)} \right)^{1/3}$$

$$\omega = \sigma \left(1 + \left(\frac{2\gamma_1}{(4 - \pi)} \right)^{2/3} \right)^{1/2}$$

$$\lambda = \left(\frac{2\gamma_1}{(4 - \pi)} \right)^{1/3} \left(\frac{2}{\pi} + \left(\frac{2\gamma_1}{(4 - \pi)} \right)^{2/3} \left(\frac{2 - \pi}{\pi} \right) \right).$$

De acuerdo con [Azzalini \(1985\)](#), al realizar inspecciones visuales a los gráficos de la verosimilitud, es posible determinar el tamaño de muestra a partir del cual aproximadamente los estimadores de máxima verosimilitud aún para λ existirán. En experiencia de dicho autor, esto se debe hacer con muestras de tamaño menor a 30 ya que en muestras grandes los estimadores de máxima verosimilitud no presentan problemas.

3.3.3. Estimadores de Momentos

De la parametrización anterior, es posible utilizar el estimador de momentos para obtener γ_1 , a través del índice de asimetría

3.3. Métodos de estimación

$$\tilde{\gamma}_1 = \frac{\sqrt{n} \sum_{i=1}^n (Y_i - \bar{Y})^3}{\left(\sum_{i=1}^n (Y_i - \bar{Y})^2 \right)^{3/2}}.$$

Se debe recordar de la ecuación 3.11 que $\gamma_1 \in (-0.9953, 0.9953)$, sin embargo, también es posible que la asimetría muestral no se encuentre dentro de dicho intervalo. En estos casos el estimador $\hat{\lambda}$ no se puede obtener por lo que generalmente se incluye una restricción a la hora de realizar la estimación. Pérez-Rodríguez y Villaseñor (2010) presentan un procedimiento para obtener $\tilde{\gamma}_1$ cuando ocurre este problema.

Si $|\tilde{\gamma}_1| > 0.995$ entonces el estimador estará dado por

$$\tilde{\gamma}_1 = 0.95 \text{sign}(\tilde{\gamma}_1) 0.995.$$

De esta manera $\tilde{\gamma}_1$ será aproximado a un valor dentro de los límites permitidos y conservará el signo correspondiente a su asimetría. A través de la expresión del índice de asimetría en términos del parámetro original λ , es posible obtener el estimador de momentos de λ , es decir, $\tilde{\lambda}$.

Coefficiente de asimetría

$$\gamma_1 = \frac{\kappa_3}{k_2^{3/2}} = \frac{4 - \pi}{2} \frac{\left(\sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1+\lambda^2}} \right)^3}{\left(1 - \frac{2}{\pi} \frac{\lambda^2}{1+\lambda^2} \right)^{3/2}}.$$

Invirtiendo esta ecuación se obtiene el estimador de momentos, $\tilde{\lambda}$

$$\tilde{\lambda} = \text{sign}(\tilde{\gamma}_1) \left(\frac{2}{4 - \pi} \right)^{1/3} |\tilde{\gamma}_1|^{1/3} \left(\frac{2}{\pi} + |\tilde{\gamma}_1|^{2/3} \left(\frac{2}{4 - \pi} \right)^{2/3} \left(\frac{2}{\pi} - 1 \right) \right)^{-1/2}.$$

Meintanis (2007) presenta las siguientes expresiones para obtener los estimadores de momentos de los tres parámetros de la distribución $SN(\xi, \omega, \lambda)$

$$\tilde{\xi} = \bar{Y}_n - \tilde{\varphi}_n \sqrt{\frac{2}{\pi}}$$

3.3. Métodos de estimación

$$\tilde{\omega} = \left(S_n^2 + \left(\frac{2}{\pi} \right) \tilde{\varphi}_n^2 \right)^{1/2}$$

$$\tilde{\delta} = \left(\frac{\bar{Y}_n - \tilde{\xi}}{\tilde{\omega}} \right) \sqrt{\frac{\pi}{2}}$$

donde

$$\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}, S_n^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2, \tilde{\varphi}_n = \left(\frac{m_3}{4 - \pi} \sqrt{\frac{\pi^3}{2}} \right)^{1/3} \text{ y } m_3 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^3.$$

Es importante recordar que estos estimadores poseen propiedades generales deseables como que son asintóticamente insesgados, consistentes en error cuadrado medio y son asintóticamente normales.

3.3.4. Estimadores de máxima verosimilitud penalizados

Para evitar problemas con la matriz de información al momento de estimar los parámetros de la distribución normal asimétrica, existe una metodología que realiza un ajuste a los estimadores de máxima verosimilitud.

[Sartori \(2006\)](#) propuso una modificación de las ecuaciones de verosimilitud que [Firth \(1993\)](#) presentó como una técnica general de reducción de sesgo. Para el caso en donde sólo un parámetro debe ser estimado, el método de Firth reemplaza la función de score $S(\theta)$ por una forma modificada $S^*(\theta)$:

$$S^*(\theta) = S(\theta) - I(\theta)b(\theta) = 0, \tag{3.12}$$

donde $b(\theta)$ es el término principal de sesgo del MLE de θ , típicamente $O(n^{-1})$. $I(\theta)$ es la matriz de información de Fisher esperada. El término extra $-I(\theta)b(\theta)$ se elige de tal manera que la estimación obtenida tenga un sesgo de orden de magnitud $O(n^{-2})$.

Si se utiliza la ecuación anterior para obtener el estimador del parámetro de forma de una muestra proveniente de la distribución $SN(0, 1, \lambda)$, Sartori demostró que siempre admite una solución finita.

3.3. Métodos de estimación

Sin embargo este enfoque tiene una dificultad, ya que no es fácilmente implementable en situaciones más complejas, como es el caso en donde los tres parámetros son desconocidos, debido a la dificultad en el momento de obtener las expresiones analíticas del término de ajuste en (3.12). Además, incluso para el caso simple, existe la desventaja práctica de que cada evaluación de la ecuación de verosimilitud requiere calcular dos integrales numéricamente.

Por lo anterior es posible considerar el funcionamiento de una formulación relacionada pero algo distinta propuesta por [Azzalini y Arellano-Valle \(2013\)](#), quienes proponen la función de verosimilitud penalizada

$$\ell_p(\theta) = \ell(\theta) - Q(\theta),$$

donde θ denota el conjunto de parámetros que se estimarán y la función de penalización $Q(\theta)$ que no depende de n , satisface que

$$Q(\theta) \geq 0, \quad Q(\theta)|_{\lambda=0} = 0, \quad \lim_{|\lambda| \rightarrow \infty} Q(\theta) = +\infty.$$

Bajo estas suposiciones, $\ell_p(\theta)$ puede ser maximizada en un punto finito, digamos $\tilde{\theta}$, que se conoce como estimación máxima de verosimilitud penalizada (MPLE).

Entonces, bajo las condiciones mencionadas y algunas condiciones de regularidad estándar se puede demostrar que la diferencia entre $\tilde{\theta}$ y $\hat{\theta}$ desaparece cuando n tiende a infinito. Esto implica que las propiedades asintóticas de $\tilde{\theta}$ son las mismas que las de $\hat{\theta}$ ([Azzalini y Capitanio, 2014](#)).

[Azzalini y Arellano-Valle \(2013\)](#), proponen la siguiente ecuación para $Q(\theta)$ tanto para el caso de la distribución normal asimétrica estándar (donde solo se desconoce el parámetro de forma) como para el caso donde los tres parámetros son desconocidos, ya que la penalización se lleva a cabo únicamente sobre el parámetro de forma

$$Q = c_1 \cdot \log(1 + c_2 \lambda^2),$$

donde $c_1 \approx 0.87591$ y $c_2 \approx 0.85625$.

Estos estimadores pueden obtenerse utilizando la función `sn.mple` que se encuentra contenida en el paquete `sn` ([Azzalini, 2018](#)) del software estadístico R ([R Core Team, 2018](#)).

3.3. Métodos de estimación

3.3.5. Otros estimadores

Otra alternativa para la obtención de estimadores de los parámetros es la que proponen [Pérez-Rodríguez y Villaseñor \(2008\)](#) quienes utilizan el algoritmo Esperanza-Maximización Generalizado (GEM) para reducir la dimensionalidad del problema y obtener valores de estimación mediante el paquete estadístico R.

Por último, [Pérez-Rodríguez *et al.* \(2017\)](#), proponen una alternativa de estimación mediante métodos bayesianos para los parámetros de localidad, escala y forma, cuando la estimación de máxima verosimilitud de este último tiende a infinito (o menos infinito) con probabilidad positiva, es decir, no puede obtenerse.

[Azzalini y Capitanio \(1999\)](#) desarrollaron un paquete en el software estadístico R en donde se encuentran programadas muchas de las rutinas para inferencia en la distribución normal asimétrica. Esta versión ha sido modificada 22 veces, agregando nuevas contribuciones y actualmente se encuentra en uso la versión 1.5-2 ([Azzalini, 2018](#)).

Capítulo 4

Metodologías propuestas

A continuación se proponen cuatro pruebas para la hipótesis de normalidad asimétrica contra la hipótesis de que la muestra proviene de cualquier otra distribución.

Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución $F(x)$, con densidad $f(x)$, soporte en \mathbb{R} y media finita cuyos parámetros son desconocidos.

Se desea probar el siguiente juego de hipótesis:

$$H_0 : f(x) = f_X(x; \xi, \omega, \lambda) \tag{4.1}$$

vs

$$H_a : f(x) \neq f_X(x; \xi, \omega, \lambda),$$

donde $f_X(x; \xi, \omega, \lambda) = 2\frac{1}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left[\lambda\left(\frac{x-\xi}{\omega}\right)\right]I_{(-\infty, \infty)}(x)$ es la densidad de una variable aleatoria normal asimétrica y los valores de los parámetros son desconocidos.

4.1. Prueba A: Transformación de variables aleatorias normales asimétricas a variables aleatorias χ^2

4.1. Prueba A: Transformación de variables aleatorias normales asimétricas a variables aleatorias χ^2

La primera propuesta utiliza una de las propiedades de la subsección Propiedades de la distribución normal asimétrica, la cual se muestra a continuación:

$$\text{Si } X \sim SN(x; \xi, \omega, \lambda) \text{ entonces } Y = \left(\frac{X-\xi}{\omega}\right)^2 \sim \chi_1^2 .$$

En el caso de pruebas de hipótesis estadísticas, se utiliza información que proviene de una muestra aleatoria cuyos parámetros son desconocidos. Por lo anterior, no es inmediato hacer uso de la propiedad mencionada.

Sin embargo, se puede demostrar que, si los valores de los parámetros son sustituidos por estimadores consistentes, la siguiente expresión es cierta:

$$\text{Si } X^* \sim SN(x; \hat{\xi}, \hat{\omega}, \hat{\lambda}) \text{ entonces } Y^* = \left(\frac{X^*-\hat{\xi}}{\hat{\omega}}\right)^2 \xrightarrow{d} \chi_1^2 \text{ cuando } n \rightarrow \infty .$$

La prueba propuesta consiste en transformar los datos de una muestra aleatoria X_1, X_2, \dots, X_n con n suficientemente grande, con el fin de obtener una muestra de valores transformados $Y_1^*, Y_2^*, \dots, Y_n^*$.

Bajo la hipótesis nula, se espera que dicha muestra transformada se comporte como una muestra proveniente de una distribución χ_1^2 . Entonces, para probar la hipótesis nula en la ecuación (4.1) se recurre a probar el siguiente juego de hipótesis.

$$H_0^* : f_{Y^*}(y) = f_Y(y; 1) \tag{4.2}$$

vs

$$H_a^* : f_{Y^*}(y) \neq f_Y(y; 1),$$

donde $f_Y(y; 1)$ es la función de densidad de una variable aleatoria χ_1^2 .

4.1. Prueba A: Transformación de variables aleatorias normales asimétricas a variables aleatorias χ^2

Estas hipótesis pueden ser probadas mediante cualquier prueba de hipótesis conocida cuya eficiencia haya sido probada. En esta propuesta se utiliza la prueba de Anderson-Darling debido a su utilidad para probar las distribuciones χ_n^2 con tamaños de muestra grandes (Thas, 2010). La prueba de Anderson-Darling ha sido descrita en la subsección 3.1.1. La constante crítica de la prueba se obtiene de la distribución de la estadística de prueba bajo la hipótesis nula en 4.2.

Para la obtención de las variables aleatorias Y no es necesario involucrar el parámetro de forma λ . Además, la distribución de Y no depende de dicho parámetro, sin embargo, no ocurre lo mismo con las variables Y^* pues su distribución así como la distribución de la estadística de prueba AD obtenida con estas variables, dependen del parámetro de forma como se muestra en la Figura 4.1.

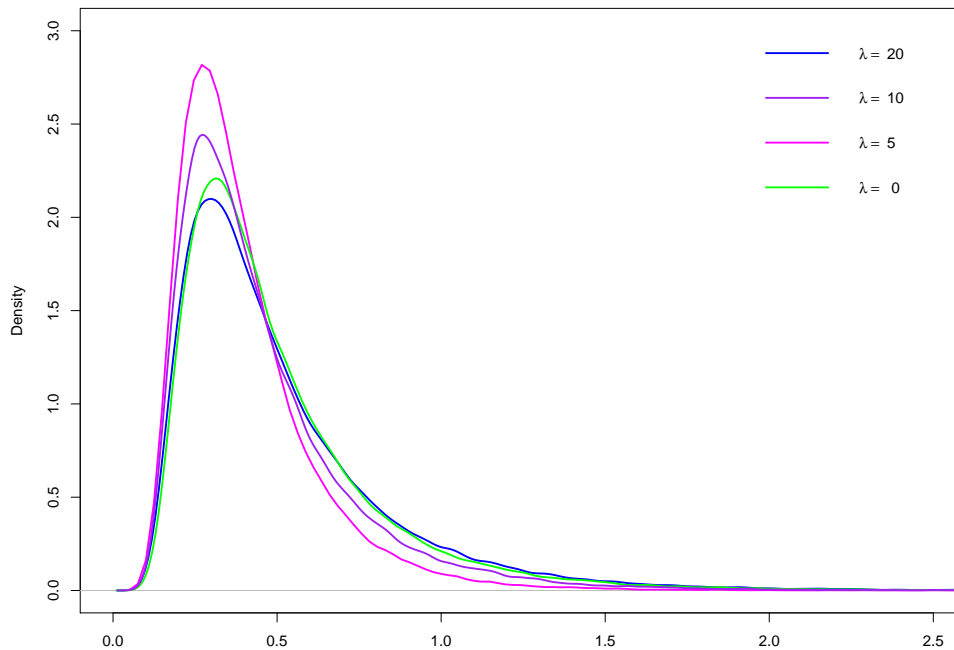


Figura 4.1: Distribución nula de la estadística de Anderson-Darling usada en la prueba A para distintos valores de λ .

Como se puede observar, la distribución del estadístico de prueba depende fuertemente del valor del parámetro de forma. Por lo anterior, para proponer una prueba y obtener la constante crítica k_α a partir de la cual se rechaza o no la hipótesis nula con un tamaño de prueba $\alpha \in [0, 1]$ especificado, es necesario resolver la siguiente ecuación:

4.1. Prueba A: Transformación de variables aleatorias normales asimétricas a variables aleatorias χ^2

$$\alpha = \sup_{\lambda \in \mathbb{R}} P(AD > k_\alpha | H_0^*)$$

Dada una muestra aleatoria, la prueba basada en la constante crítica k_α rechazaría H_0^* si AD_{calc} es mayor que dicha constante, donde AD_{calc} es la estadística AD calculada usando la muestra aleatoria.

Es importante notar que la distribución de la estadística de prueba no depende del signo de λ debido a una extensión de la propiedad 2 que dice que si $X \sim SN(\xi, \omega, \lambda)$ entonces $-X \sim SN(-\xi, \omega, -\lambda)$. Por lo tanto el valor crítico se puede determinar sobre la parte positiva del conjunto de posibles valores de λ , es decir, k_α es tal que:

$$\alpha = \sup_{\lambda \geq 0} P(AD > k_\alpha | H_0^*) . \quad (4.3)$$

Además, como lo mencionan [Pérez-Rodríguez y Villaseñor \(2010\)](#), se sabe que el comportamiento probabilístico de una variable aleatoria SN con parámetro de forma $|\lambda| \geq 20$ es similar al de una variable aleatoria con distribución HN. Esto implica que el comportamiento probabilístico del estadístico de prueba cuando $\lambda > 20$ es similar al que ocurre con $\lambda = 20$.

Prueba A (1)

Para aproximar la distribución de la estadística AD bajo H_0^* se utilizó simulación Monte Carlo. A partir de muestras aleatorias de 15 tamaños distintos, y con 5,000 repeticiones, se aproximó la distribución de la estadística de prueba para diferentes valores de λ . Los resultados de la simulación muestran que el máximo de la probabilidad del Error Tipo I ocurre cuando $\lambda = 20$. Por lo tanto, k_α dada en la ecuación (4.3) es el cuantil $(1 - \alpha)100\%$ de la distribución de AD obtenida cuando $X_i \sim SN(\lambda = 20)$.

La Tabla 4.1 presenta los valores críticos de la prueba para distintos valores de α y distintos tamaños de muestra.

Prueba A (2)

Alternativamente, para aproximar el valor crítico k_α de la prueba A, se propone usar el método de bootstrap paramétrico. Este método ha sido empleado en otros trabajos ([Meintanis y Bassiakos \(2007\)](#); [Meintanis \(2007\)](#); [Pérez-Rodríguez y Villaseñor \(2010\)](#); [Meintanis et](#)

4.1. Prueba A: Transformación de variables aleatorias normales asimétricas a variables aleatorias χ^2

Tabla 4.1: Valores críticos para la prueba A (1) para distintos tamaños de muestra (n) y diferentes tamaños de prueba (α)

$n \backslash \alpha$	0.01	0.05	0.1
10	1.81295	1.10544	0.85449
20	1.35727	0.92694	0.70312
30	1.41455	0.88701	0.74156
50	1.51083	0.96059	0.77379
75	1.57295	1.01117	0.83141
100	1.50746	1.03178	0.82226
150	1.7093	1.06297	0.85481
200	1.59484	1.09328	0.8849
300	1.69282	1.123	0.89984
500	1.73138	1.17442	0.94144
1000	1.75513	1.17205	0.98179
2000	1.88148	1.21858	0.98932
5000	1.98468	1.28747	1.04373
7500	1.96267	1.28295	1.02886
10000	1.99895	1.29965	1.05395

al. (2014); Villaseñor y González-Estrada (2015)) ya que generalmente tiene como resultado una mayor precisión para alcanzar un tamaño de prueba α , lo que influye directamente en la mejora de la potencia de la prueba. Además, como se menciona en capítulos anteriores, existe evidencia de su validez al aproximar distribuciones de ciertas estadísticas de prueba, entre las que se encuentra Anderson-Darling (Babu y Rao, 2004).

Dada una muestra aleatoria X_1, X_2, \dots, X_n , los pasos que se siguen para realizar el procedimiento de la prueba propuesta son los siguientes:

1. Cálculo de los estimadores de máxima verosimilitud $\hat{\xi}, \hat{\omega}$ y $\hat{\lambda}$, a través la función `sn.mple` del paquete `sn` (Azzalini, 2018) de R.
2. Obtener las variables transformadas $Y_i^* = \left(\frac{X_i - \hat{\xi}}{\hat{\omega}}\right)^2$, $i = 1, \dots, n$ a partir de la muestra.
3. Obtener la estadística de prueba AD basada en las variables Y_i^* 's.
4. Obtener B muestras bootstrap de la distribución SN con los parámetros estimados de la muestra en el paso 1.
5. Obtener B valores del estadístico de prueba a partir de las muestras bootstrap del inciso anterior.

4.2. Prueba B: Anderson-Darling para la distribución normal asimétrica usando bootstrap

6. Utilizando los B valores anteriores se aproxima la distribución de la estadística de prueba mediante el uso de frecuencias y se obtiene el valor crítico de la prueba de acuerdo con un nivel de significancia α . La prueba es de cola derecha, por tanto el valor crítico $k_{1-\alpha}$ es aproximado por el cuantil empírico del $(1 - \alpha)100\%$ de los B valores simulados de AD.
7. Se rechaza la hipótesis nula H_0^* en (4.2) para un tamaño de prueba $\alpha \in [0, 1]$ si la estadística AD es mayor que $k_{1-\alpha}$. Además, rechazar H_0^* en (4.2) implica rechazar H_0 en (4.1) dada la construcción de la prueba, por lo tanto se rechaza H_0 si la estadística AD es mayor que $k_{1-\alpha}$.

4.2. Prueba B: Anderson-Darling para la distribución normal asimétrica usando bootstrap

Como segunda propuesta se implementó una versión modificada de la prueba de Anderson-Darling. Dicha versión combina el uso de la prueba convencional con el método de remuestreo bootstrap paramétrico, utilizado en la propuesta anterior.

Dada una muestra aleatoria X_1, X_2, \dots, X_n , la prueba propuesta está basada en la obtención de la estadística de prueba AD sobre dicha muestra, la cual será denotada como AD_2 . Esta estadística usa la función de distribución empírica de la muestra y toma la diferencia entre ésta y la función de distribución acumulada de una variable aleatoria normal asimétrica con parámetros estimados por máxima verosimilitud.

Para llevar a cabo el cálculo de la estadística AD_2 se estiman los parámetros usando la muestra aleatoria suponiendo que los datos provienen de una distribución normal asimétrica. Posteriormente, se utiliza el paquete “*ADGofTest*” (Bellosta, 2011), de R para obtener la estadística, la cual tiene la expresión siguiente:

$$AD_2 = n \int_{-\infty}^{\infty} \frac{\left[F_n(x) - F_X(x, \hat{\xi}, \hat{\omega}, \hat{\lambda}) \right]^2}{F_X(x, \hat{\xi}, \hat{\omega}, \hat{\lambda}) \left[1 - F_X(x, \hat{\xi}, \hat{\omega}, \hat{\lambda}) \right]} dF_X(x, \hat{\xi}, \hat{\omega}, \hat{\lambda}) \quad (4.4)$$

donde $F_X(x, \hat{\xi}, \hat{\omega}, \hat{\lambda})$ es la función de distribución acumulada de una variable aleatoria $SN(\hat{\xi}, \hat{\omega}, \hat{\lambda})$

4.2. Prueba B: Anderson-Darling para la distribución normal asimétrica usando bootstrap

dada en la ecuación (3.6).

Como se mencionó en la subsección 3.1.1, cuando los parámetros ξ , ω y λ no son conocidos (como es el caso) y se utilizan estimadores para el cálculo de la estadística AD_2 , la distribución de dicha estadística bajo la hipótesis nula no puede ser conocida de forma explícita incluso para muestras grandes y depende de la distribución de la muestra. Sin embargo, dicha distribución puede ser aproximada mediante métodos Monte Carlo. Se propone específicamente utilizar el método bootstrap paramétrico. La Figura 4.2 presenta la distribución nula de AD_2 para distintos valores del parámetro de forma.

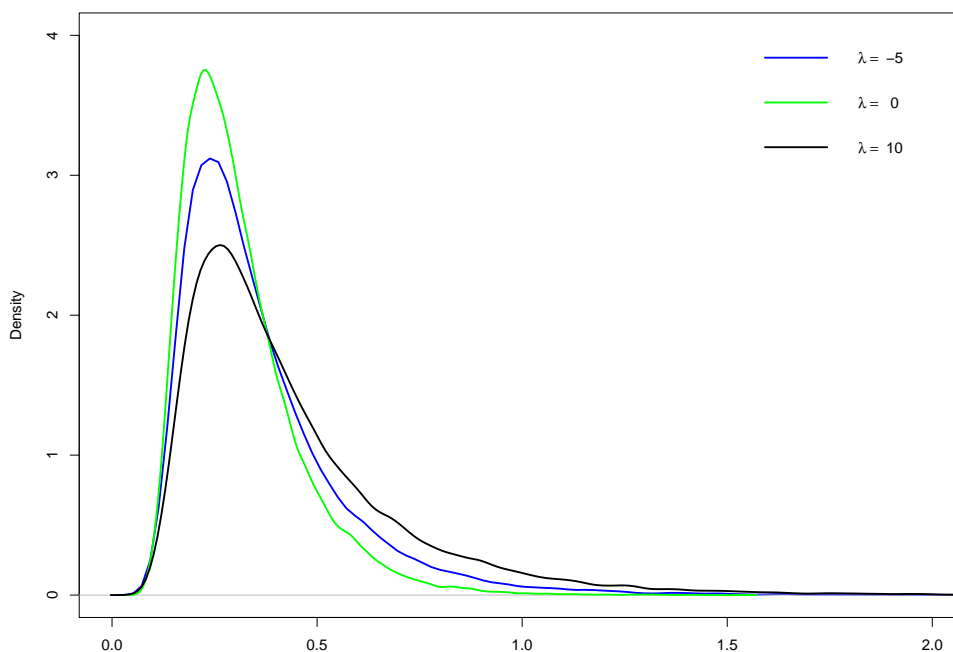


Figura 4.2: Distribución nula de la estadística de Anderson-Darling usada en la Prueba B para distintos valores de λ .

Para calcular la constante crítica k_α a partir de la cual se rechaza o no la hipótesis nula para un tamaño de prueba $\alpha \in [0, 1]$ especificado, es necesario resolver la siguiente ecuación:

$$\alpha = P(AD_2 > k_\alpha | H_0 : X \sim SN(\hat{\xi}, \hat{\omega}, \hat{\lambda})). \quad (4.5)$$

En este punto es necesario aclarar que los valores críticos de la prueba se calculan para cada muestra aleatoria a la que se desea aplicar la prueba de bondad de ajuste.

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

A continuación se muestran los pasos a seguir para probar el juego de hipótesis en (4.1) mediante la Prueba B propuesta:

1. Calcular los estimadores de máxima verosimilitud $\hat{\xi}$, $\hat{\omega}$ y $\hat{\lambda}$, a través la función `sn.mple` del paquete `sn` (Azzalini, 2018) de R.
2. Calcular la estadística de prueba de Anderson-Darling utilizando los datos de la muestra, X_1, X_2, \dots, X_n .
3. Obtener B muestras bootstrap de la distribución SN con los parámetros estimados de la muestra en el paso 1.
4. Obtener B valores del estadístico de prueba a partir de las muestras bootstrap del inciso anterior.
5. Utilizando los B valores anteriores, aproximar la distribución de la estadística de prueba mediante frecuencias y obtener el valor crítico de la prueba de acuerdo con un nivel de significancia α . La prueba es de cola derecha, por tanto el valor crítico $k_{1-\alpha}$ es aproximado por el cuantil empírico del $(1 - \alpha)100\%$ de los B valores simulados de AD_2 .
6. Se rechaza la hipótesis nula H_0 en (4.1) para un tamaño de prueba $\alpha \in [0, 1]$ si la estadística AD_2 es mayor que $k_{1-\alpha}$.

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

La tercera propuesta que se plantea en el presente trabajo está basada en la propiedad 8. de la subsección 3.2.1, la cual establece que si $X \sim SN(\xi, \omega, \lambda)$ entonces $V = |X - \xi| \sim HN(0, \omega)$. Haciendo uso de esta propiedad probabilística y de una variable aleatoria auxiliar U , independiente de V , con densidad

$$f_U(u) = \begin{cases} \frac{1}{2} & \text{si } u = -1, 1 \\ 0 & \text{d.o.m.} \end{cases} \quad (4.6)$$

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

podemos obtener la siguiente transformación:

$$W = UV$$

La variable $W = UV$ tiene una distribución $N(0, \omega^2)$. La demostración de este hecho se puede ver a través de la función generatriz de momentos de W como sigue:

$$M_W(t) = E\{e^{t(UV)}\} = \sum_u \int_0^\infty e^{t(uv)} f_V(v) f_U(u) dv = \sum_u \int_0^\infty e^{t(uv)} \frac{2}{\omega} \frac{1}{\sqrt{2\pi}} e^{(\frac{-v^2}{2\omega^2})} \frac{1}{2} dv$$

Realizando un cambio de variable

$$z = \frac{v}{\omega} \Rightarrow v = \omega z$$

$$dz = \frac{dv}{\omega} \Rightarrow dv = \omega dz$$

Entonces

$$\begin{aligned} M_W(t) &= \sum_u \int_0^\infty e^{t(u\omega z)} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz = \sum_u \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(z^2 - 2tu\omega z + t^2 u^2 \omega^2 - t^2 u^2 \omega^2)} dz \\ &= \sum_u e^{\frac{1}{2}(t^2 u^2 \omega^2)} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(z - tu\omega)^2} dz = \sum_u e^{\frac{1}{2}(t^2 u^2 \omega^2)} [1 - \Phi(-tu\omega)] \\ &= e^{\frac{1}{2}(t^2 \omega^2)} [1 - \Phi(-t\omega)] + e^{\frac{1}{2}(t^2 \omega^2)} [1 - \Phi(t\omega)] \\ &= e^{\frac{1}{2}(t^2 \omega^2)} [1 - \Phi(t\omega) + 1 - (1 - \Phi(t\omega))] \\ &= e^{\frac{1}{2}(t^2 \omega^2)} \end{aligned}$$

la cual es la función generatriz de momentos de una variable aleatoria $N(0, \omega^2)$, ([Villaseñor et al., 2017](#)).

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n . Considere las variables aleatorias

$$W_i^* = U_i V_i^*,$$

donde U_i y V_i^* son independientes, $V_i^* = |X_i - \hat{\xi}|$ y U_i es un valor simulado de U , $i = 1, \dots, n$.

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

Por la propiedad anterior, para probar H_0 en la ecuación (4.1) se propone probar el siguiente juego de hipótesis:

$$H_0^* : W^* \sim \text{Normal} \tag{4.7}$$

vs

$$H_a^* : W^* \not\sim \text{Normal}.$$

En esta propuesta se emplea la prueba de Shapiro-Wilk (cuya estadística de prueba será denotada como SW) para probar H_0^* en (4.7), debido a que ha sido ampliamente estudiada y resulta ser efectiva para probar normalidad en muestras grandes y pequeñas especialmente contra distribuciones alternativas asimétricas en donde otras pruebas han resultado ser poco sensibles (Romao *et al.* (2010), Noughabi y Arghami (2011) y Patrício *et al.* (2017)).

Para verificar si la distribución de la estadística de prueba SW bajo la hipótesis nula en (4.7) depende de λ , se realizó el cálculo de dicha estadística para 50000 muestras de tamaño $n = 50$ y $n = 500$, con 5 valores distintos del parámetro λ y utilizando dos tipos de estimadores, de máxima verosimilitud y de máxima verosimilitud penalizados. Se muestran los resultados obtenidos a través de la Figura (4.3) y la Figura (4.4).

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

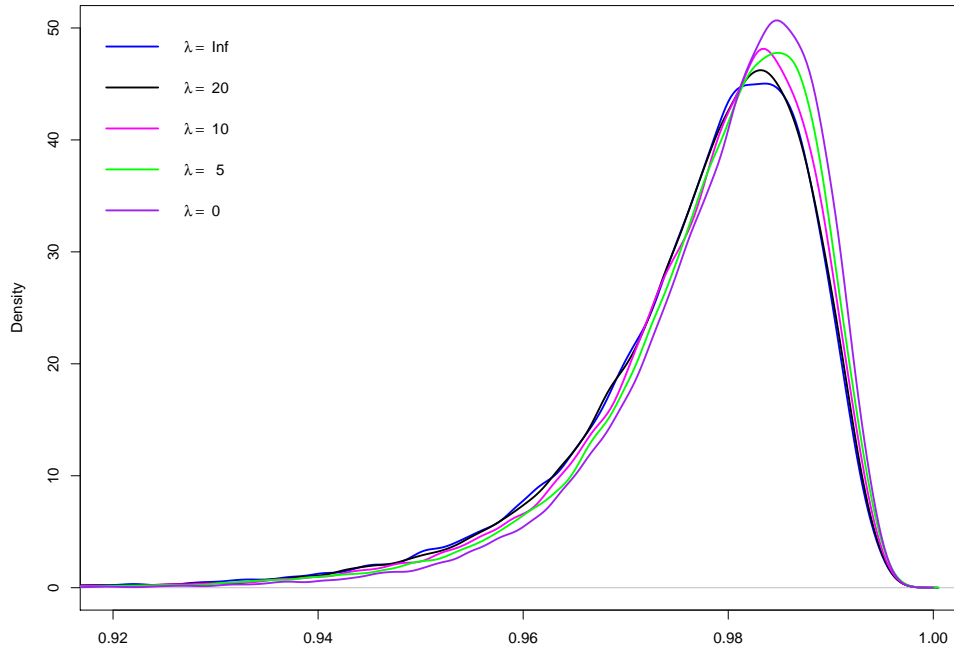


Figura 4.3: Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud y muestras de tamaño $n=50$.

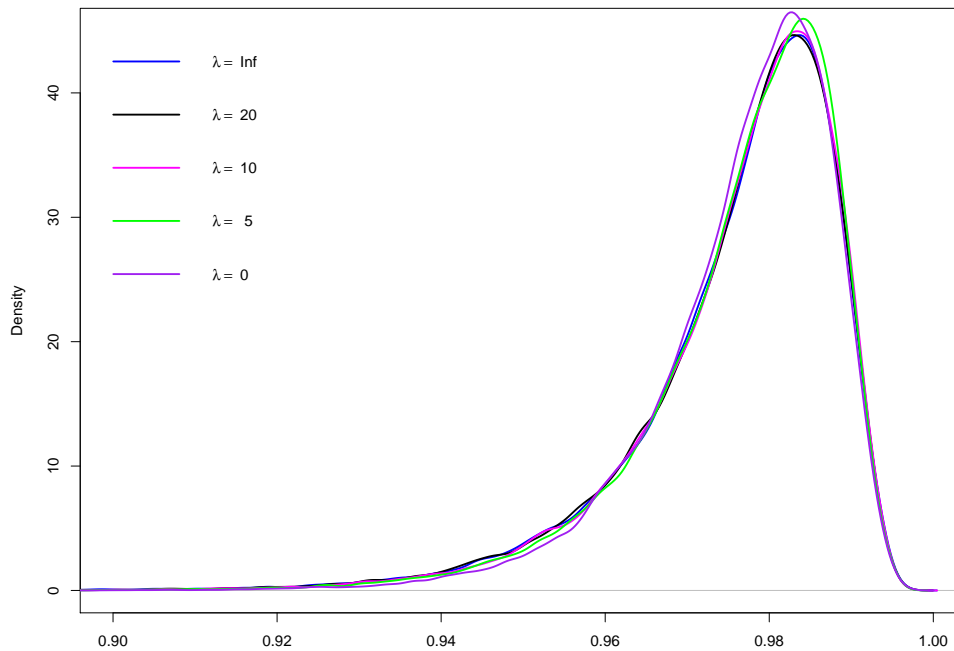


Figura 4.4: Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud penalizados y muestras de tamaño $n=50$.

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

Como se puede observar en la Figura (4.3), la distribución de la estadística de prueba depende del valor del parámetro λ cuando se obtienen las transformaciones usando estimadores de máxima verosimilitud. Sin embargo es diferente cuando se utilizan estimadores de máxima verosimilitud penalizados como se observa en la Figura (4.4) en donde la distribución de la estadística de prueba bajo la hipótesis nula no depende del valor de λ .

En la Figura (4.5) y la Figura (4.6) se muestra también el comportamiento de la distribución de la estadística SW para muestras de tamaño $n = 500$ para verificar empíricamente que el tamaño de muestra no influye en el comportamiento de dicha distribución.

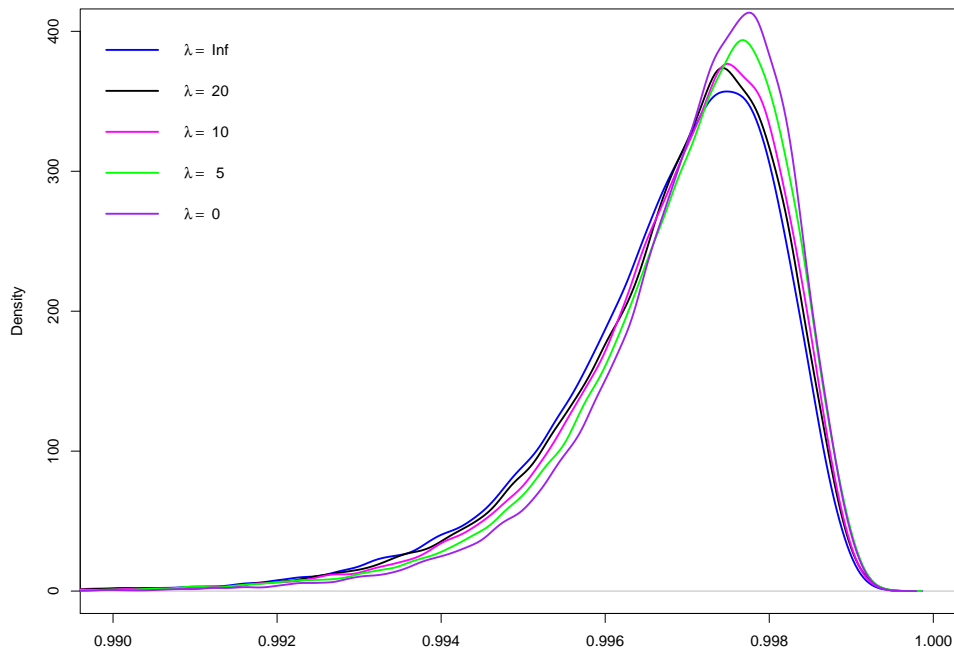


Figura 4.5: Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud y muestras de tamaño $n=500$.

4.3. Prueba C: Transformación de variables aleatorias normales asimétricas a variables aleatorias normales

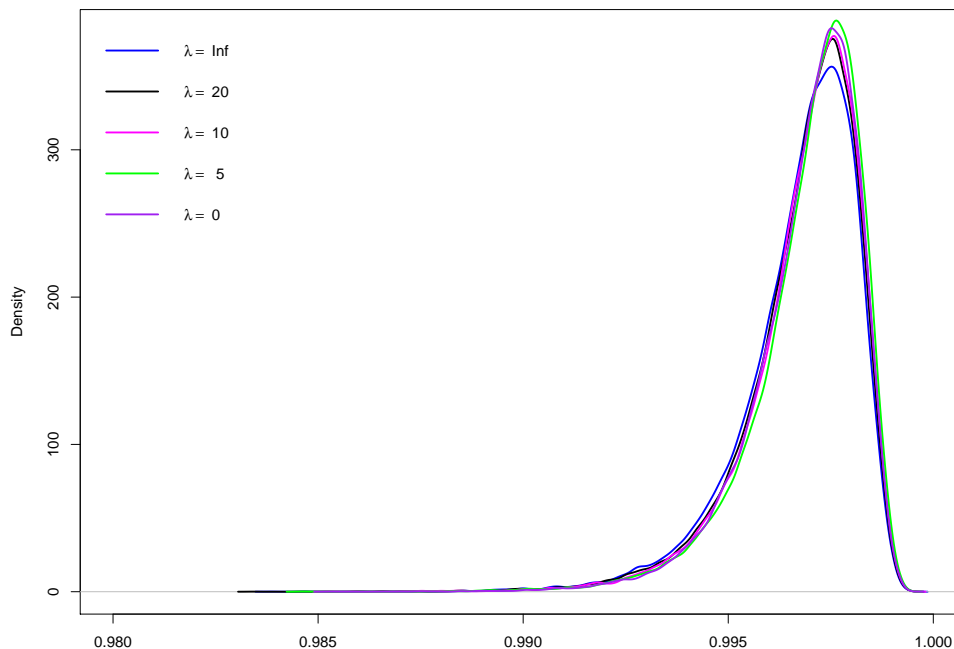


Figura 4.6: Distribución nula de la estadística Shapiro-Wilk usada en la Prueba C para distintos valores de λ usando estimadores de máxima verosimilitud penalizados y muestras de tamaño $n=500$.

Por lo anterior, se recomienda el uso de estimadores de máxima verosimilitud penalizados para llevar a cabo la propuesta de la prueba, ya que no es necesario implementar técnicas adicionales como bootstrap para obtener un nivel de prueba especificado. A continuación se muestran los pasos para realizar la prueba.

A partir de una muestra aleatoria de tamaño n , realizar los siguientes pasos:

1. Obtener el estimador de máxima verosimilitud penalizado $\hat{\xi}$ a través de la función `sn.mple` del paquete `sn` (Azzalini, 2018) de R.
2. Obtener las variables transformadas $V_i^* = |X_i - \hat{\xi}|$, $i = 1, \dots, n$, a partir de la muestra.
3. Simular n valores de la variable U , U_1, \dots, U_n , cuya densidad está dada por la ecuación (4.6).
4. Calcular las variables $W_i^* = U_i V_i^*$, $i = 1, 2, \dots, n$.
5. Probar el juego de hipótesis en (4.7) usando la prueba de Shapiro-Wilk, (Shapiro y Wilk, 1965). Note que si se rechaza H_0^* en (4.7) con un tamaño de prueba $\alpha \in [0, 1]$ entonces se rechaza H_0 en (4.1) con el mismo tamaño de prueba.

Capítulo 5

Estudio de simulación, resultados y discusión

Se realizó un estudio mediante simulación Monte Carlo para evaluar las propiedades de tamaño y potencia de las pruebas de forma empírica.

5.1. Obtención del tamaño y potencia de la prueba

Cuando los parámetros desconocidos son de localidad y escala, la distribución muestral de la estadística de prueba de Anderson-Darling es invariante a cambios en dichos parámetros (Mateu-Figueras *et al.*, 2007). Lo anterior implica que, para el estudio de tamaño de prueba, en donde la muestra proviene de una distribución normal asimétrica, basta realizar el cálculo variando únicamente el parámetro de forma dejando fijo el valor de los parámetros de localidad y escala ya que la estadística de prueba será invariante a cambios de localidad y escala. Por conveniencia se utilizan $\xi = 0$ y $\omega = 1$.

A continuación se describe el procedimiento para estimar el tamaño de las pruebas mediante simulación Monte Carlo.

Dado el juego de hipótesis en (4.1),

1. Generar una muestra aleatoria de tamaño n de una distribución $SN(0, 1, \lambda)$, es decir,

5.1. Obtención del tamaño y potencia de la prueba

asumiendo H_0 verdadera.

2. Sobre dicha muestra llevar a cabo la prueba de bondad de ajuste de interés, ya sea A (1), A (2), B o C, y determinar si se rechaza o no H_0 , para un tamaño de prueba $\alpha \in [0, 1]$.
3. Generar N muestras Monte Carlo como en el paso 1 para llevar a cabo N veces el paso 2.
4. Calcular la proporción de rechazos obtenidos en las N repeticiones para un tamaño $\alpha \in [0, 1]$. La proporción anterior es una estimación del tamaño de la prueba.

Para el caso del presente estudio, se realizaron $N = 5000$ repeticiones Monte Carlo de muestras aleatorias de tamaño $n = 100$ para un nivel $\alpha=0.05$. Dado que las distribuciones de las estadísticas de prueba de las cuatro propuestas varían bajo cambios en el parámetro de forma λ , se realizó el cálculo del tamaño estimado de prueba para distintos valores de λ en donde es factible observar cambios significativos, es decir, para $\lambda \in [-20,20]$.

Por último, en la estimación del tamaño de las pruebas A (2) y B se utilizaron 200 muestras bootstrap puesto que estas pruebas requieren dicho método para su funcionamiento.

Para el cálculo de la potencia de las pruebas se sigue un algoritmo similar al anterior pero asumiendo H_0 falsa, se detalla a continuación:

Dado el juego de hipótesis en (4.1),

1. Generar una muestra aleatoria de tamaño n de una distribución diferente a la $SN(\xi, \omega, \lambda)$, es decir, asumiendo H_0 falsa.
2. Sobre dicha muestra llevar a cabo la prueba de bondad de ajuste de interés, ya sea A (1), A (2), B o C y determinar si se rechaza o no H_0 , para un tamaño de prueba $\alpha \in [0, 1]$.
3. Generar N muestras Monte Carlo como en el paso 1 para llevar a cabo N veces el paso 2.
4. Calcular la proporción de rechazos obtenidos en las N repeticiones para un tamaño $\alpha \in [0, 1]$. La proporción anterior es una estimación de la potencia de la prueba con N repeticiones, para muestras bajo la hipótesis alternativa de tamaño n para un tamaño de prueba $\alpha \in [0, 1]$.

5.1. Obtención del tamaño y potencia de la prueba

Para el caso del presente estudio, se realizaron $N = 5000$ repeticiones Monte Carlo de muestras aleatorias de tamaños $n = 50$, $n = 100$, $n = 300$ y $n = 500$ de distintas distribuciones alternativas, para un nivel $\alpha=0.05$. Como en el caso del tamaño, para las pruebas A (2) y B se utilizaron 200 muestras bootstrap puesto que requieren dicho método para su funcionamiento.

A continuación se describen brevemente las distribuciones alternativas utilizadas en el estudio de potencia, se presenta su función de densidad, la parametrización utilizada y el soporte de cada una.

Distribución t-student

La función de densidad de una variable aleatoria X con distribución *t – student*, $t(\nu)$, está dada por:

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad (5.1)$$

donde $\nu > 0$ son los grados de libertad y $x \in \mathbb{R}$.

Esta alternativa fue utilizada para evaluar el desempeño de las pruebas para $\nu = 2, 4, 5$, y $\nu = 12$. Para la generación de números aleatorios de esta distribución se utilizó el paquete “*stats*” de R.

Distribución logística

La función de densidad de una variable aleatoria X con distribución logística está dada por:

$$f_X(x) = \frac{e^{-\frac{x-\mu}{s}}}{s \left(1 + e^{-\frac{x-\mu}{s}}\right)^2}, \quad (5.2)$$

donde μ es el parámetro de localidad y $s > 0$ el de escala, y $x \in \mathbb{R}$.

Para el estudio de potencia, se utilizaron $\mu = 0$ y $s = 1$, es decir, su versión estándar. La generación de números aleatorios de esta distribución se llevó a cabo mediante el paquete “*stats*” de R.

Distribución Laplace

5.1. Obtención del tamaño y potencia de la prueba

La función de densidad de una variable aleatoria X con distribución Laplace está dada por:

$$f_X(x) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right), \quad (5.3)$$

donde μ es el parámetro de localidad, $b > 0$ el de escala, y $x \in \mathbb{R}$.

Para el estudio de potencia, se utilizaron $\mu = 0$ y $b = 1$, es decir, su versión estándar. La generación de números aleatorios de esta distribución se llevó a cabo mediante el paquete “*smoothest*”, (Hennig, 2012), de R.

Distribución Uniforme

La función de densidad de una variable aleatoria X con distribución uniforme (a, b) está dada por:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{para } x \in [a, b] \\ 0 & \text{d.o.m.} \end{cases} \quad (5.4)$$

donde $-\infty < a < b < \infty$ y $x \in [a, b]$.

Para el estudio de potencia, se utilizaron $a = 0$ y $b = 1$. La generación de números aleatorios de esta distribución se llevó a cabo mediante el paquete “*stats*” de R.

Distribución Beta

La función de densidad de una variable aleatoria con distribución Beta está dada por:

$$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad (5.5)$$

donde $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $\alpha, \beta > 0$ son parámetros de forma, y $x \in (0, 1)$.

Se utilizaron dos pares diferentes de valores de los parámetros ($\alpha = 0.5, \beta = 0.5$) y ($\alpha = 1, \beta = 3$). La generación de números aleatorios de estas distribuciones se llevó a cabo mediante el paquete “*stats*” de R.

Distribución Cauchy

5.1. Obtención del tamaño y potencia de la prueba

La función de densidad de una variable aleatoria $X \sim Cauchy(m, k)$ está dada por:

$$f_X(x) = \frac{1}{\pi k \left[1 + \left(\frac{x-m}{k} \right)^2 \right]} \quad (5.6)$$

donde m es el parámetro de localidad y k el de escala, y $x \in \mathbb{R}$.

El estudio de potencia se realizó sobre muestras aleatorias de una distribución Cauchy con parámetros $m = 0$ y $k = 1$. La generación de números aleatorios de estas distribuciones se llevó a cabo mediante el paquete “stats” de R.

Distribución Gumbel

La función de densidad de una variable aleatoria $X \sim Gumbel(\xi, \theta)$ está dada por:

$$f_X(x) = \frac{1}{\theta} e^{-(z+e^{-z})}, \quad (5.7)$$

donde $z = \frac{x-\xi}{\theta}$, ξ es el parámetro de localidad y $\theta > 0$ el de escala, y $x \in \mathbb{R}$.

El estudio de potencia se realizó sobre muestras aleatorias con parámetros $\xi = 0$ y $\theta = 1$. La generación de números aleatorios de estas distribuciones se llevó a cabo mediante una función creada en el programa R que se muestra en los anexos de este documento.

Distribución Lognormal

La función de densidad de una variable aleatoria $X \sim lognormal(\mu, \sigma)$ está dada por:

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad (5.8)$$

donde $\mu \in \mathbb{R}$, $\sigma > 0$ y $x \in \mathbb{R}^+$.

Se utilizaron dos combinaciones de los parámetros de esta distribución para el estudio de potencia, $\mu = 0$ y $\sigma = 0.5$ y $\mu = 0$ y $\sigma = 1.5$. La generación de números aleatorios de estas distribuciones se llevó a cabo mediante el paquete “stats” de R.

Distribución exponencial

5.1. Obtención del tamaño y potencia de la prueba

La función de densidad de una variable aleatoria $X \sim \exp(\lambda)$ está dada por:

$$f_X(x) = \lambda e^{-\lambda x}, \quad (5.9)$$

donde $\lambda > 0$ es el inverso del parámetro de escala y $x \in \mathbb{R}^+$.

Se consideró un valor de $\lambda = 1$, es decir, su forma estándar. La generación de números aleatorios se llevó a cabo mediante el paquete “stats” de R.

Distribución gamma

La distribución gamma también fue considerada en el estudio como distribución alternativa. La función de densidad de una variable aleatoria $X \sim \text{gamma}(\alpha, \beta)$ está dada por:

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (5.10)$$

donde $\alpha > 0$ es el parámetro de forma, $\beta > 0$ el inverso del parámetro de escala y $x \in \mathbb{R}^+$.

Las potencias se estimaron para $\alpha = 0.5$ y $\beta = 1$. La generación de números aleatorios se llevó a cabo mediante el paquete “stats” de R.

Distribución Weibull

La función de densidad de una variable aleatoria $X \sim \text{Weibull}(k, \lambda)$ está dada por:

$$f_X(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, \quad (5.11)$$

donde $\lambda \in \mathbb{R}^+$ es el parámetro de escala, $k \in \mathbb{R}^+$ el parámetro de forma y $x \in \mathbb{R}^+$.

Las potencias se estimaron para $\lambda = 1$ y $k = 0.75$. La generación de números aleatorios se llevó a cabo mediante el paquete “stats” de R.

Distribución t asimétrica

La función de densidad de una variable aleatoria X con distribución t -asimétrica con parámetro de forma λ y ν grados de libertad, denotada como $st(\lambda, \nu)$ está dada por:

5.1. Obtención del tamaño y potencia de la prueba

$$f_X(x) = 2f_{T_\nu}(x)F(\lambda x) = 2\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}F_{T_{\nu+1}}\left(\lambda x\sqrt{\frac{1+\nu}{\nu(\nu+x^2)}}\right) \quad (5.12)$$

donde $\lambda \in \mathbb{R}$, $x \in \mathbb{R}$, $f_{T_\nu}(x)$ es la densidad de una distribución t con ν grados de libertad y $F_{T_{\nu+1}}(x)$ es la función de distribución de $f_{T_{\nu+1}}(x)$.

Las potencias se estimaron para $\lambda = 0.5$ y $\nu = 1, 5$ y 10 . La generación de números aleatorios se llevó a cabo mediante el paquete “*sn*” (Azzalini, 2018) de R.

Distribución Laplace asimétrica

Otra densidad asimétrica considerada es la Laplace asimétrica. La función de densidad de una variable aleatoria $X \sim sLaplace(\mu, \sigma, \lambda)$ está dada por:

$$f_X(x) = \begin{cases} \frac{\lambda e^{-\lambda(x-\mu/\sigma)}}{\sigma(1+\lambda^2)} & \text{para } x \geq \mu \\ \frac{\lambda e^{\lambda(x-\mu/\lambda\sigma)}}{\sigma(1+\lambda^2)} & \text{d.o.m.} \end{cases} \quad (5.13)$$

donde $\mu, \lambda \in \mathbb{R}$, $\sigma > 0$ y $x \in \mathbb{R}$.

Las potencias se estimaron para $\mu = 0$, $\sigma = 1$ y $\lambda = 0.5$. La generación de números aleatorios se llevó a cabo mediante el paquete “*rmutil*” (Swihart y Lindsey, 2017) de R.

Distribución Cauchy asimétrica

Otra densidad asimétrica considerada es la Cauchy asimétrica con parámetros de localidad y escala 0 y 1 respectivamente. La función de densidad de una variable aleatoria $X \sim sCauchy(0, 1, \lambda)$ está dada por:

$$f_X(x) = 2f(x)F(\lambda x) = \frac{[1 + 2\arctan(\lambda x)/\pi]}{\pi(x^2)} \quad (5.14)$$

donde $\lambda \in \mathbb{R}$ y $x \in \mathbb{R}$.

Las potencias se estimaron para $\lambda = 1$. La generación de números aleatorios se llevó a cabo mediante el paquete “*sn*” (Azzalini, 2018) de R.

Distribución Pareto generalizada

5.2. Otras pruebas para la distribución SN

Otra densidad considerada en el estudio de potencia es la Pareto generalizada. La función de densidad de una variable aleatoria $X \sim gp(\mu, \sigma, \xi)$ está dada por:

$$f_X(x) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-(1/\xi+1)}, \quad (5.15)$$

donde $x \geq \mu$ para $\xi \geq 0$ y $\mu \leq x \leq \mu - \sigma/\xi$ para $\xi < 0$, $\mu \in \mathbb{R}$ es el parámetro de localidad, $\sigma \in \mathbb{R}^+$ el de escala y $\xi \in \mathbb{R}$ el de forma.

Las potencias se estimaron para muestras aleatorias de las distribuciones $gp(0, 1, 0.15)$ y $gp(0, 1, 0.5)$. La generación de números aleatorios se llevó a cabo mediante el paquete “*SpatialExtremes*” (Ribatet, 2018) de R.

5.2. Otras pruebas para la distribución SN

Para tener una idea del funcionamiento de las pruebas propuestas, en el estudio de potencia realizado en este trabajo, se incluyó la prueba de correlación propuesta por Pérez-Rodríguez y Villaseñor (2010) quienes a su vez contrastan su trabajo con pruebas existentes anteriormente. Una idea general de esta prueba se muestra a continuación:

Dada una muestra aleatoria X_1, X_2, \dots, X_n con densidad $f(x)$, soporte en \mathbb{R} con media y varianzas finitas, se desea probar la hipótesis nula en (4.1).

Para un λ fijo, digamos λ_0 , $P(X \leq x) = F_{\lambda_0} \left(\frac{x-\xi}{\omega} \right)$, donde F_{λ_0} es la función de distribución acumulada de $Z \sim SN(\lambda_0)$. Además, la función de distribución empírica de la muestra $F_n(x)$ es un estimador consistente de $P(X \leq x)$, entonces, $F_{\lambda_0} \left(\frac{x-\xi}{\omega} \right) \approx F_n(x)$, por lo tanto, $v_i := F_{\lambda_0}^{-1} (F_n(x_i)) \approx \frac{x_i - \xi}{\omega}$. Por lo anterior, se espera que bajo la hipótesis nula exista una fuerte relación lineal entre las variables x'_i s y v'_i s.

La estadística de prueba utilizada es:

$$r_n^* = Corr(X, V) = \frac{\sum_{i=1}^n (X_i - \bar{X})(V_i - \bar{V})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (V_i - \bar{V})^2}}$$

5.2. Otras pruebas para la distribución SN

donde λ_0 para el cálculo de las variables V_i es reemplazado por el estimador de momentos de λ .

Se rechaza la hipótesis nula en (4.1) para un nivel de significancia $\alpha \in [0, 1]$ si $r_n^* \leq C_n(\alpha)$, donde $C_n(\alpha)$ es tal que:

La distribución de r_n^* bajo H_0 para distintos valores de λ se obtiene mediante bootstrap paramétrico. Se rechaza la hipótesis nula en (4.1) para un nivel de significancia $\alpha \in [0, 1]$ si $r_n^* \leq C_n(\alpha)$, el cual denota el cuantil $100\alpha\%$ de dicha distribución.

Esta prueba requiere que existan la media y la varianza de la población muestreada para que esté bien definido el coeficiente de correlación, el cual es la estadística de prueba utilizada.

Otra de las pruebas utilizadas en este estudio de potencia como referencia del funcionamiento de este trabajo es la que desarrolló Meintanis (2007).

Dada una muestra aleatoria X_1, X_2, \dots, X_n se desea probar la hipótesis nula

$$H_0 : \text{la distribución de } X_i \text{ es } SN(\xi, \omega, \lambda)$$

para algún $\lambda, \xi \in \mathbb{R}$ y $\omega > 0$

la estadística de prueba utilizada es

$$\tilde{T}_{n,a} = \sqrt{n} \sup_{-a \leq t \leq a} |\tilde{D}_n(t)| \quad (5.16)$$

la cual es calculada para valores de a cercanos a 0 y evaluando:

$$\tilde{D}_n(t) = \hat{M}'_n(t) - t\hat{M}_n(t) - \frac{\hat{\lambda}}{\sqrt{1 + \hat{\lambda}^2}} \sqrt{\frac{2}{\pi}} \exp\left[\frac{t^2}{2} \left(1 - \frac{\hat{\lambda}^2}{1 + \hat{\lambda}^2}\right)\right],$$

donde $\hat{\lambda}$ es un estimador consistente de λ , $\hat{M}_n(t) = \frac{1}{n} \sum_{i=1}^n \exp(t \frac{X_i - \hat{\xi}}{\hat{\omega}})$ para $i = 1, 2, \dots, n$ y $t \in \mathbb{R}$.

5.3. Tamaño estimado de prueba

Los estimadores utilizados son estimadores momentos y se muestran a continuación:

$$\begin{aligned}\tilde{\xi} &= \bar{Y}_n - \tilde{\varphi}_n \sqrt{\frac{2}{\pi}} \\ \tilde{\omega} &= \left(S_n^2 + \left(\frac{2}{\pi} \right) \tilde{\varphi}_n^2 \right)^{1/2} \\ \tilde{\delta} &= \left(\frac{\bar{Y}_n - \tilde{\xi}}{\tilde{\omega}} \right) \sqrt{\frac{\pi}{2}}\end{aligned}$$

donde

$$\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}, S_n^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2, \tilde{\varphi}_n = \left(\frac{m_3}{4 - \pi} \sqrt{\frac{\pi^3}{2}} \right)^{1/3} \text{ y } m_3 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^3.$$

El valor de a en donde se evalúa la ecuación (5.16) es $a = 0.2$ con pasos de $0.001a$ para poder evaluar en un número finito de puntos el intervalo $(-a, a)$.

La prueba resultante es de cola derecha y la distribución de la estadística de prueba es aproximada usando bootstrap paramétrico, la constante crítica asociada es el cuantil $(1 - \alpha)100\%$ de dicha distribución. Entonces, se rechaza H_0 si el valor calculado de la estadística de prueba es mayor que la constante crítica.

5.3. Tamaño estimado de prueba

Una de las características deseables de una prueba de bondad de ajuste consiste en que ésta preserve el tamaño deseado.

En las Figuras 5.1 a 5.4 se presentan los resultados del tamaño estimado de cada propuesta y su comportamiento para los distintos valores de λ considerados en el estudio, correspondientes a muestras aleatorias de tamaño $n = 100$ y un tamaño de prueba nominal $\alpha = 0.05$.

5.3. Tamaño estimado de prueba

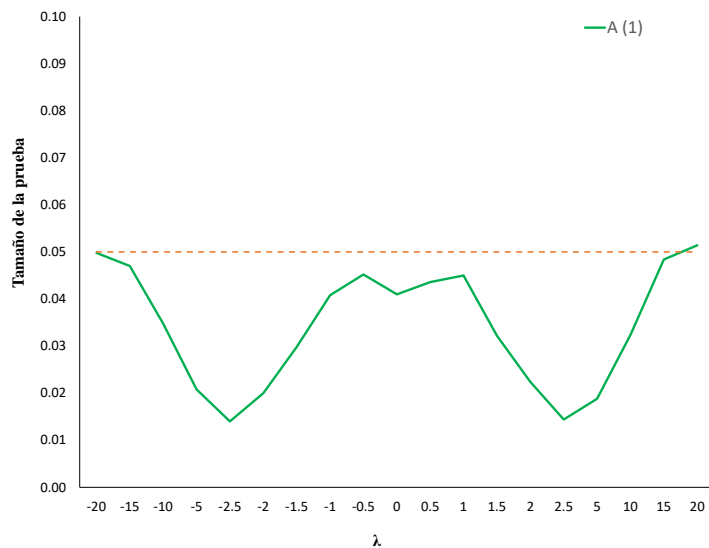


Figura 5.1: Tamaño estimado de la prueba A(1) para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$.

Para la prueba A (1) en la Figura 5.1 se puede observar que el tamaño de la prueba se mantiene por debajo del nivel establecido $\alpha=0.05$ excepto quizá para los valores del parámetro de forma cercanos a -20 y a 20, en donde el tamaño estimado está alrededor de 0.06. Lo anterior puede deberse exclusivamente a la variabilidad introducida por la simulación, sin embargo, es evidente que la prueba preserva el nivel deseado.

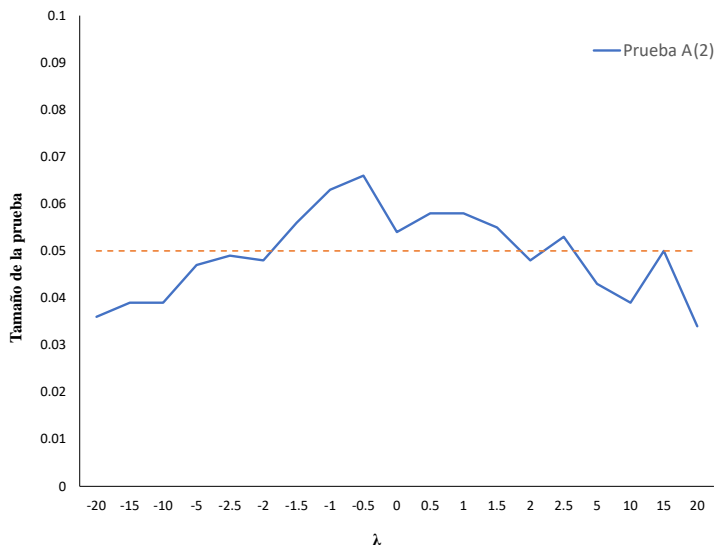


Figura 5.2: Tamaño estimado de la prueba A (2) para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$.

5.3. Tamaño estimado de prueba

En la Figura 5.2 se observa que el tamaño de la prueba A (2) se mantiene por debajo o cercano a $\alpha=0.05$. Los valores que sobrepasan el nivel nuevamente se encuentran alrededor de 0.06 aunque, como ya se mencionó, estos excedentes pueden ser explicados por el uso de simulación.

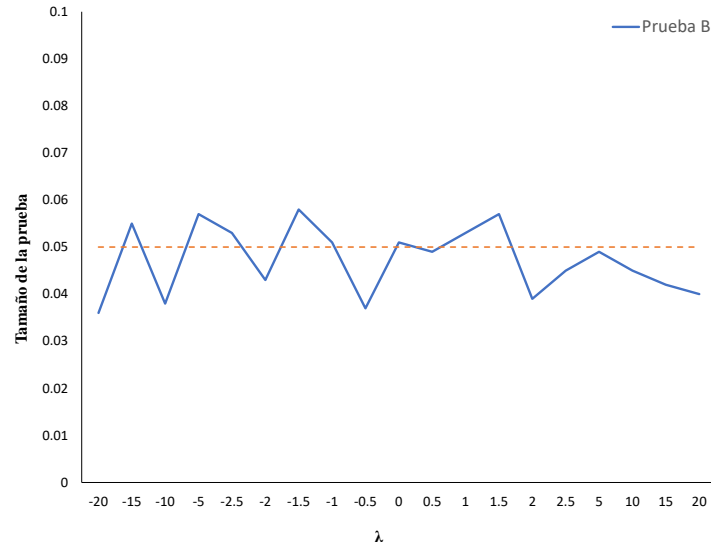


Figura 5.3: Tamaño estimado de la prueba B para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$.

En el caso de la prueba B el tamaño estimado de la prueba se comporta de una manera satisfactoria (ver Figura 5.3), ya que no presenta ninguna tendencia y todos los valores se encuentran cerca de $\alpha = 0.05$. Lo anterior indica que el nivel de la prueba se conserva para distintos valores del parámetro de forma.

5.3. Tamaño estimado de prueba

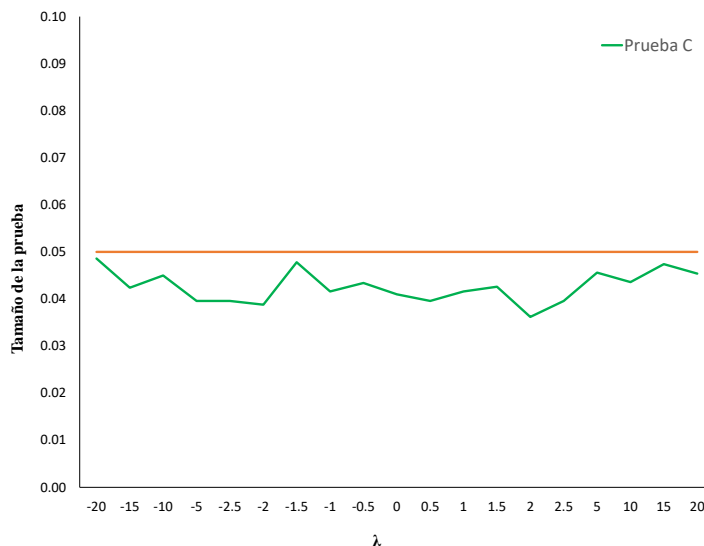


Figura 5.4: Tamaño estimado de la prueba C para valores del parámetro de forma en el intervalo $[-20, 20]$, $n=100$, $\alpha = 0.05$.

Para la última propuesta, la prueba C, el tamaño estimado para los distintos valores del parámetro de forma, no sobrepasa el nivel establecido $\alpha=0.05$ (ver Figura 5.4). Se observa una ligera disminución del tamaño alrededor del valor $\lambda = 0$, lo cual puede hacer que contra distribuciones con menor asimetría la prueba tenga menor potencia.

La Tabla 5.1 presenta los tamaños de prueba estimados para distintos valores de λ . Se incluyen además los resultados de las pruebas de bondad de ajuste r_n^* , desarrollada por Pérez-Rodríguez y Villaseñor (2010), y la prueba de tipo Kolmogorov-Smirnov, D_n , propuesta por Meintanis (2007). En el apartado de Anexos se reportan los tamaños de prueba estimados con distintos tamaños de muestra ($n = 50$ y $n = 500$), en donde se puede verificar que las estimaciones no dependen de n . Los tamaños de prueba reportados en la Tabla 5.1 se encuentran redondeados a dos cifras decimales para facilitar su lectura.

Para el tamaño de prueba, se observan valores muy cercanos al nivel de significancia $\alpha=0.05$ para todas las pruebas, lo que confirma que tiene sentido llevar a cabo una comparación en su desempeño ante otras distribuciones alternativas. En el caso de la prueba de Meintanis (2007) existe cierto descontrol para valores del parámetro mayores que 2.5. Por lo anterior, esta prueba no será incluida en el estudio de potencia.

Una vez que se ha confirmado que cada prueba propuesta preserva el nivel deseado, es importante revisar otra de las características más importantes, que es la potencia.

5.3. Tamaño estimado de prueba

Tabla 5.1: Tamaño estimado de las pruebas para diferentes valores del parámetro de forma, correspondiente a muestras aleatorias de tamaño $n=100$ y $\alpha = 0.05$.

Lambda	A (1)	A (2)	B	C	r_n^*	D_n
-20	0.05	0.04	0.04	0.05	0.05	0.01
-15	0.05	0.04	0.06	0.04	0.04	0.01
-10	0.03	0.04	0.04	0.05	0.04	0.01
-5	0.02	0.05	0.06	0.04	0.03	0.01
-2.5	0.01	0.05	0.05	0.04	0.03	0.02
-2	0.02	0.05	0.04	0.04	0.02	0.02
-1.5	0.03	0.06	0.06	0.05	0.01	0.04
-1	0.04	0.06	0.05	0.04	0.01	0.06
-0.5	0.05	0.07	0.04	0.04	0.01	0.05
0	0.04	0.05	0.05	0.04	0.01	0.05
0.5	0.04	0.06	0.05	0.04	0.01	0.06
1	0.05	0.06	0.05	0.04	0.01	0.04
1.5	0.03	0.06	0.06	0.04	0.02	0.02
2	0.02	0.05	0.04	0.04	0.02	0.03
2.5	0.01	0.05	0.05	0.04	0.03	0.03
5	0.02	0.04	0.05	0.05	0.03	0.17
10	0.03	0.04	0.05	0.04	0.04	0.29
15	0.05	0.05	0.04	0.05	0.04	0.33
20	0.05	0.03	0.04	0.05	0.05	0.33

5.4. Potencia estimada de las pruebas

A continuación se muestra el desempeño de cada propuesta a través de la estimación de su potencia para cuatro tamaños de muestra distintos. Se realizó el análisis con muestras aleatorias de distintas distribuciones alternativas, las cuales fueron seleccionadas de tal manera que se cubriera un espectro amplio de comportamientos probabilísticos, es decir, se incluyen distribuciones con parámetros de localidad y escala, de forma, de cola ligera, pesada, etcétera. Para mejor comprensión de los resultados se agregan el coeficiente de asimetría (CA) y el de curtosis (CK) de cada distribución alternativa.

La potencia estimada de las pruebas propuestas es alta bajo algunas distribuciones alternativas, por ejemplo, en el caso de la distribución Gamma (0.5, 1), en donde la potencia es 1.

Haciendo referencia a los valores de la Tabla 5.2, la cual presenta las potencias estimadas de las pruebas bajo las distribuciones alternativas consideradas con muestras aleatorias de tamaño $n = 50$ y $\alpha = 0.05$, a partir de la distribución Pareto generalizada (0.15) (con parámetro de localidad cero y escala uno), se puede observar un desempeño sobresaliente de las cuatro propuestas, lo que indica que, para distribuciones con un coeficiente de asimetría y curtosis elevado, cualquiera de las cuatro propuestas puede ser utilizada con buenos resultados, incluso mejores que las pruebas ya existentes.

Además, note que las distribuciones exponencial estándar y Laplace asimétrica son bien identificadas bajo las propuestas A (1), A (2) y B, sin embargo la propuesta C (transformación a variables aleatorias normales) tiene una menor potencia. También para estas distribuciones, las pruebas propuestas representan una alternativa mejor que las pruebas existentes.

Para el caso de las distribuciones asimétricas que tienen coeficientes de asimetría y curtosis menores como lognormal (0, 0.5), st (0.5,5), t asimétrica, Beta (1,3) y Gumbel estándar, las pruebas propuestas tienen una potencia menor a las pruebas ya existentes, las proporciones de rechazo bajo la hipótesis alternativa son bajas, es decir que dichas distribuciones pueden confundirse con una distribución normal asimétrica de manera frecuente con muestras aleatorias de tamaño 50. Lo anterior puede ser no del todo lógico ya que la distribución Beta, por ejemplo, tiene soporte acotado mientras que la normal asimétrica no, o la distribución log-normal, cuyo soporte son los reales positivos, sin embargo se debe recordar que la distribución normal asimétrica cuenta con un parámetro de forma, adicional a los de localidad

5.4. Potencia estimada de las pruebas

Tabla 5.2: Potencia estimada de las pruebas para muestras de tamaño $n=50$, $\alpha = 0.05$.

CA	CK	Distribución alterna	A (1)	A (2)	B	C	r_n^*
S							
0	3.8	t(12)	0.06	0.05	0.07	0.09	0.07
0	4.2	Logística estándar	0.07	0.08	0.12	0.13	0.12
0	9.0	t(5)	0.12	0.15	0.25	0.25	0.29
0	nd	t(4)	0.20	0.24	0.34	0.34	0.40
0	6.0	Laplace estándar	0.27	0.29	0.50	0.36	0.49
0	4.8	Uniforme(0,1)	0.42	0.33	0.40	0.57	0.31
nd	nd	t(2)	0.65	0.69	0.83	0.77	-
0	1.5	Beta(0.5,0.5)	0.96	0.87	0.89	0.85	0.89
nd	nd	Cauchy estándar	0.99	0.99	1.00	0.99	-
A							
1.1	5.4	Gumbel estándar	0.06	0.07	0.08	0.13	0.14
-0.9	8.5	st(0.5,10)	0.07	0.07	0.10	0.11	0.16
0.7	3.1	Beta(1,3)	0.20	0.12	0.10	0.04	0.02
-0.3	13.6	st(0.5,5)	0.13	0.16	0.23	0.25	0.38
1.8	8.9	Lognormal (0,0.5)	0.17	0.17	0.17	0.27	0.31
1.8	8.6	Laplace asimétrica (0.5)	0.64	0.65	0.66	0.41	0.49
2.0	9.0	Exponencial estándar	0.82	0.72	0.71	0.52	0.51
3.5	33.5	Pareto generalizada (0.15)	0.94	0.90	0.91	0.77	0.73
2.8	15.0	Gamma(0.5,1)	1.00	1.00	1.00	0.96	0.90
12.4	28.7	Weibull (0.75, 1)	1.00	0.99	0.99	0.95	0.89
nd	nd	st(0.5,1)	0.98	0.98	1.00	0.98	-
nd	nd	sCauchy(1)	0.99	0.99	1.00	0.98	-
33.47	10078	Lognormal (0,1.5)	1.00	1.00	1.00	1.00	0.99
nd	nd	Pareto Generalizada (0.5)	1.00	1.00	1.00	0.98	-

- Potencia no estimada ya que no existe el segundo momento de la distribución alternativa, CA=coeficiente de asimetría, CK=coeficiente de curtosis, S=simétrica, A=asimétrica, nd=no definido.

5.4. Potencia estimada de las pruebas

y escala, por lo que tiene una mayor flexibilidad y por lo tanto puede adoptar un mayor número de formas reconocibles.

Por su parte la distribución t asimétrica, como en el caso simétrico para la distribución normal, generalmente es una buena aproximación de la distribución normal asimétrica conforme los grados de libertad aumentan. Por lo anterior es de esperar que, conforme el número de grados de libertad aumente, las pruebas tengan menor potencia. Un claro ejemplo es el comportamiento de las pruebas bajo las alternativas $st(0.5,5)$ y $st(0.5,10)$. En estos casos la prueba r_n^* tiene mayor potencia que cualquiera de las propuestas en este trabajo.

Ahora, si se presta atención a las distribuciones simétricas, las distribuciones que mejor son diferenciadas por las pruebas propuestas son la Cauchy estándar, $Beta(0.5, 0.5)$ y $t(2)$. Las pruebas A (1) y C son preferibles contra las primeras dos, para la $t(2)$ la potencia más alta es de la prueba B (AD con bootstrap).

Para el caso de la distribución uniforme (0,1) la prueba que tiene mayor potencia es B. Esta distribución también es de soporte acotado sin embargo ninguna prueba logra diferenciarla de manera concluyente de una distribución normal asimétrica.

Las pruebas B y r_n^* alcanzan prácticamente la misma potencia con la distribución Laplace estándar en la hipótesis alternativa, a pesar de que esta potencia no es alta, son las mejores opciones si se sospecha que los datos pueden provenir de dicha distribución.

Por otro lado, como ya se mencionó, la distribución t puede aproximar el comportamiento de una distribución normal de manera más precisa conforme sus grados de libertad aumenten. Hay que recordar que la distribución normal es un caso particular de la familia normal asimétrica, por lo tanto se espera que, conforme los grados de libertad aumenten, las pruebas sean menos potentes. Se puede apreciar en los resultados que para las alternativas $t(4)$, $t(5)$ y $t(12)$ la potencia de todas las pruebas disminuye considerablemente, sin embargo la que tiene el mejor desempeño es la propuesta C seguida por la prueba r_n^* .

Para la distribución logística estándar las potencias también son bajas, siendo la propuesta C y B y la prueba r_n^* las que presentan los valores más altos.

La Tabla 5.3 presenta la potencia estimada de las pruebas para un tamaño de muestra $n=100$ y $\alpha = 0.05$. En esta tabla se puede visualizar como el aumento en el tamaño de muestra provoca un aumento en la potencia de cada prueba. El mayor crecimiento en potencia se

5.4. Potencia estimada de las pruebas

observa en la prueba r_n^* mientras que de las pruebas propuestas el aumento mayor fue en B.

Tabla 5.3: Potencia estimada de las pruebas para muestras de tamaño $n=100$, $\alpha = 0.05$.

CA	CK	Distribución alterna	A (1)	A (2)	B	C	r_n^*
S							
0	3.8	t(12)	0.05	0.07	0.13	0.12	0.10
0	4.2	Logística estándar	0.08	0.09	0.24	0.20	0.20
0	9.0	t(5)	0.19	0.23	0.45	0.41	0.47
0	nd	t(4)	0.30	0.36	0.62	0.55	0.62
0	6.0	Laplace estándar	0.46	0.48	0.83	0.58	0.76
0	4.8	Uniforme(0,1)	0.83	0.83	0.94	0.95	0.93
nd	nd	t(2)	0.88	0.89	0.98	0.95	-
0	1.5	Beta(0.5,0.5)	1.00	1.00	1.00	1.00	1.00
nd	nd	Cauchy estándar	1.00	1.00	1.00	1.00	-
A							
1.1	5.4	Gumbel estándar	0.07	0.11	0.12	0.16	0.22
-0.9	8.5	st(0.5,10)	0.06	0.08	0.14	0.14	0.25
0.7	3.1	Beta(1,3)	0.21	0.14	0.16	0.05	0.01
-0.3	13.6	st(0.5,5)	0.19	0.23	0.45	0.41	0.58
1.8	8.9	Lognormal (0,0.5)	0.26	0.28	0.30	0.39	0.45
1.8	8.6	Laplace asimétrica (0.5)	0.91	0.95	0.94	0.66	0.75
2.0	9.0	Exponencial estándar	0.96	0.92	0.93	0.80	0.76
3.5	33.5	Pareto generalizada (0.15)	1.00	1.00	0.99	0.96	0.92
2.8	15.0	Gamma(0.5,1)	1.00	1.00	1.00	1.00	0.99
12.4	28.7	Weibull (0.75, 1)	1.00	1.00	1.00	1.00	0.99
nd	nd	st(0.5,1)	1.00	1.00	1.00	1.00	-
nd	nd	sCauchy(1)	1.00	1.00	1.00	1.00	-
33.47	10078	Lognormal (0,1.5)	1.00	1.00	1.00	1.00	1.00
nd	nd	Pareto Generalizada (0.5)	1.00	1.00	1.00	1.00	-

- Potencia no estimada ya que no existe el segundo momento de la distribución alternativa, CA=coeficiente de asimetría, CK=coeficiente de curtosis, S=simétrica, A=asimétrica, nd=no definido.

En general el comportamiento de las potencias estimadas con muestras aleatorias de tamaño 50 se mantiene, sin embargo existen casos en donde no es así. Por ejemplo, la prueba A (1) presentaba la mayor potencia contra la distribución uniforme (0, 1) para $n=50$, no obstante la mayor potencia para $n=100$ se obtiene con la propuesta C. De igual forma para la distribución $t(2)$ la propuesta B es la que tiene un mejor desempeño. Para la distribución Beta (0.5, 0.5) y Cauchy estándar, cualquier prueba propuesta es igual de efectiva a partir de este tamaño de muestra.

5.4. Potencia estimada de las pruebas

Lo mismo ocurre con las distribuciones asimétricas que presentan mayor asimetría y curtosis (Pareto generalizada (0.15) hacia abajo en la Tabla 5.3), cualquier prueba puede ser utilizada y se obtendrán potencias altas. Para el resto de las distribuciones asimétricas con muestras aleatorias de tamaño 100, se mantiene el comportamiento observado con muestras de tamaño 50.

La Tabla 5.4 contiene la potencia estimada de cada prueba con un tamaño de muestra $n=300$ y $\alpha = 0.05$. Se observa un comportamiento creciente general de la potencia y a continuación se detallan los cambios en el orden de recomendación de uso de las pruebas.

Para el tamaño de muestra $n = 300$ y con alternativas simétricas, la propuesta B es la opción más recomendable. Para el caso de las alternativas asimétricas hay cinco distribuciones en donde ninguna prueba alcanza el valor de potencia máxima, Gumbel estándar, st (0.5, 10), Beta (1, 3), st (0.5, 5) y lognormal (0, 0.5). Para la alternativa lognormal (0, 0.5) las pruebas C y r_n^* funcionan mejor que las demás y para la distribución st (0.5, 10) la prueba r_n^* funciona mejor, al igual que para la st (0.5, 5), aunque para esta última, la propuesta B compite bastante bien con la prueba r_n^* . Finalmente para la distribución Beta (1, 3) la propuesta A (1) y B son recomendables y para la alternativa Gumbel estándar la propuesta B y C.

Por último la Tabla 5.5 presenta la potencia estimada de las pruebas para un tamaño de muestra $n=500$ y $\alpha = 0.05$, y se detallan los cambios en el comportamiento de la potencia de cada prueba.

Para un tamaño de muestra $n=500$ la mayor parte de las distribuciones asimétricas presentadas en este estudio, son distinguibles de una distribución normal asimétrica mediante cualquiera de las pruebas propuestas y ya existentes.

Para el caso de la distribución lognormal, la propuesta que mayor potencia tiene es la C, sin embargo las pruebas ya existentes son las que tienen un mejor desempeño.

La potencia registrada para la distribución alternativa st (0.5, 5) es aproximadamente de la misma magnitud entre la propuesta B y la prueba r_n^* que son las que tienen potencias más grandes.

En cuanto a las distribuciones Beta (1, 3), st (0.5, 10) y Gumbel estándar se observa que la potencia es mayor que con muestras de tamaño $n=300$, sin embargo no son valores elevados. Para la distribución Beta (1, 3) es recomendable usar la propuesta C o la A (1), r_n^* para la

5.4. Potencia estimada de las pruebas

Tabla 5.4: Potencia estimada de las pruebas para muestras de tamaño $n=300$, $\alpha = 0.05$.

CA	CK	Distribución alterna	A (1)	A (2)	B	C	r_n^*
S							
0	3.8	t(12)	0.07	0.08	0.25	0.22	0.13
0	4.2	Logística estándar	0.13	0.19	0.57	0.39	0.39
0	9.0	t(5)	0.39	0.45	0.88	0.77	0.88
0	nd	t(4)	0.66	0.72	0.96	0.90	0.76
0	6.0	Laplace estándar	0.88	0.90	1.00	0.91	0.92
0	4.8	Uniforme(0,1)	1.00	1.00	1.00	1.00	1.00
nd	nd	t(2)	1.00	1.00	1.00	1.00	-
0	1.5	Beta(0.5,0.5)	1.00	1.00	1.00	1.00	1.00
nd	nd	Cauchy estándar	1.00	1.00	1.00	1.00	-
A							
1.1	5.4	Gumbel estándar	0.11	0.28	0.30	0.32	0.26
-0.9	8.5	st(0.5,10)	0.08	0.12	0.35	0.28	0.55
0.7	3.1	Beta(1,3)	0.33	0.23	0.23	0.08	0.01
-0.3	13.6	st(0.5,5)	0.43	0.52	0.88	0.73	0.92
1.8	8.9	Lognormal (0,0.5)	0.59	0.68	0.66	0.72	0.71
1.8	8.6	Laplace asimétrica (0.5)	1.00	1.00	1.00	0.99	1.00
2.0	9.0	Exponencial estándar	1.00	1.00	1.00	1.00	0.88
3.5	33.5	Pareto generalizada (0.15)	1.00	1.00	1.00	1.00	1.00
2.8	15.0	Gamma(0.5,1)	1.00	1.00	1.00	1.00	1.00
12.4	28.7	Weibull (0.75, 1)	1.00	1.00	1.00	1.00	1.00
nd	nd	st(0.5,1)	1.00	1.00	1.00	1.00	-
nd	nd	sCauchy(1)	1.00	1.00	1.00	1.00	-
33.47	10078	Lognormal (0,1.5)	1.00	1.00	1.00	1.00	1.00
nd	nd	Pareto Generalizada (0.5)	1.00	1.00	1.00	1.00	-

- Potencia no estimada ya que no existe el segundo momento de la distribución alternativa, CA=coeficiente de asimetría, CK=coeficiente de curtosis, S=simétrica, A=asimétrica, nd=no definido.

5.4. Potencia estimada de las pruebas

Tabla 5.5: Potencia estimada de las pruebas para muestras de tamaño $n=500$, $\alpha = 0.05$.

CA	CK	Distribución alterna	A (1)	A (2)	B	C	r_n^*
S							
0	3.8	t(12)	0.08	0.11	0.39	0.31	0.15
0	4.2	Logística estándar	0.19	0.25	0.80	0.56	0.43
0	9.0	t(5)	0.62	0.69	0.98	0.92	0.94
0	nd	t(4)	0.85	0.90	1.00	0.99	0.81
0	6.0	Laplace estándar	0.98	0.98	1.00	0.99	1.00
0	4.8	Uniforme(0,1)	1.00	1.00	1.00	1.00	1.00
nd	nd	t(2)	1.00	1.00	1.00	1.00	-
0	1.5	Beta(0.5,0.5)	1.00	1.00	1.00	1.00	1.00
nd	nd	Cauchy estándar	1.00	1.00	1.00	1.00	-
A							
1.1	5.4	Gumbel estándar	0.17	0.42	0.44	0.45	0.47
-0.9	8.5	st(0.5,10)	0.11	0.17	0.53	0.37	0.72
0.7	3.1	Beta(1,3)	0.43	0.34	0.39	0.13	0.05
-0.3	13.6	st(0.5,5)	0.61	0.76	0.98	0.90	1.00
1.8	8.9	Lognormal (0,0.5)	0.79	0.89	0.88	0.90	0.99
1.8	8.6	Laplace asimétrica (0.5)	1.00	1.00	1.00	1.00	1.00
2.0	9.0	Exponencial estándar	1.00	1.00	1.00	1.00	1.00
3.5	33.5	Pareto generalizada (0.15)	1.00	1.00	1.00	1.00	1.00
2.8	15.0	Gamma(0.5,1)	1.00	1.00	1.00	1.00	1.00
12.4	28.7	Weibull (0.75, 1)	1.00	1.00	1.00	1.00	1.00
nd	nd	st(0.5,1)	1.00	1.00	1.00	1.00	-
nd	nd	sCauchy(1)	1.00	1.00	1.00	1.00	-
33.47	10078	Lognormal (0,1.5)	1.00	1.00	1.00	1.00	1.00
nd	nd	Pareto Generalizada (0.5)	1.00	1.00	1.00	1.00	-

- Potencia no estimada ya que no existe el segundo momento de la distribución alternativa, CA=coeficiente de asimetría, CK=coeficiente de curtosis, S=simétrica, A=asimétrica, nd=no definido.

5.4. Potencia estimada de las pruebas

st (0.5, 10) y r_n^* , B o C para Gumbel estándar.

En las distribuciones simétricas se observan potencias de prueba bastante buenas excepto quizá en los casos donde las alternativas son $t(12)$ y logística estándar, las cuales son las distribuciones que han presentado menor potencia para todos los tamaños de muestra estudiados. En estos casos la propuesta B presenta potencias elevadas.

A continuación, la Tabla 5.6 muestra el número veces que cada prueba es recomendable sobre las demás, lo anterior hablando exclusivamente de las nueve alternativas simétricas y las catorce asimétricas reportadas en las tablas anteriores. Note que las filas no suman nueve y catorce respectivamente, esto se debe a que algunas pruebas tienen potencias iguales para algunas alternativas y se cuenta como acierto para todas.

Tabla 5.6: Número de alternativas en donde las pruebas estudiadas tienen la mayor potencia.

n		A (1)	A (2)	B	C	r_n^*
50	Simétricas	2	0	4	0	5
	Asimétricas	7	3	6	2	4
100	Simétricas	2	2	8	2	5
	Asimétricas	9	8	6	7	7
300	Simétricas	4	4	9	4	5
	Asimétricas	10	9	10	9	11
500	Simétricas	4	4	9	4	5
	Asimétricas	9	9	9	10	13

Note que, para $n=50$, la prueba r_n^* es recomendable para contrastar la distribución normal asimétrica contra más alternativas simétricas que las otras propuestas aunque también la propuesta B es recomendable. Para el caso de las alternativas asimétricas es recomendable la propuesta A (1) y la B.

Para $n=100$ la propuesta B tiene mayor potencia para alternativas simétricas y las propuestas A (1) y A (2) para las asimétricas. Para $n=300$, las pruebas propuestas tienen un desempeño cercano a la prueba r_n^* contra distribuciones asimétricas, y para distribuciones simétricas el mejor desempeño es el de la propuesta B. Por último, para $n=500$ la prueba r_n^* tiene el mejor desempeño contrastando distribuciones asimétricas y la propuesta B las simétricas.

También se muestra en el apartado de Anexos, evidencia empírica de la consistencia de las

5.4. Potencia estimada de las pruebas

pruebas, es decir, el comportamiento de cada prueba conforme n crece para cada distribución alternativa. En dichas tablas se puede observar que, la potencia de las pruebas aumenta conforme aumenta el tamaño de muestra contra todas las alternativas consideradas.

Capítulo 6

Aplicación

Los datos presentados a continuación son medidas de diámetro normal (diámetro a 1.30 m del piso) obtenidos de una plantación forestal de Eucalipto a una edad de 2.5, 4.7, 6 y 12 años. La base de datos cuenta con 545 medidas de diámetro para la edad 2.5, 1200 para la edad 4.7, 447 para la edad 6 y 677 para la edad 12.

Llevar a cabo el ajuste de modelos probabilísticos a esta clase de datos es importante en el ámbito forestal, ya que dichos modelos son utilizados para optimizar la cosecha y calcular el valor de la producción.

Para modelar estas observaciones generalmente se utiliza la distribución Weibull de tres parámetros, la cual tiene el comportamiento de la distribución en la ecuación (5.11) pero con un parámetro de localidad. La elección de esta distribución se debe a que puede adecuarse bien a la forma de la distribución de los datos y a su asimetría.

Los datos también pueden ser modelados por otras distribuciones y debido a sus características es factible proponer como modelo la distribución normal asimétrica. Se muestra un análisis exploratorio simple de los conjuntos de datos de diámetros normales para las cuatro edades distintas.

Como se puede ver en la Tabla (6.1), los datos son positivos y presentan asimetría, además junto con la curtosis indican que la distribución asociada a los datos es de cola ligera.

Como ya se mencionó antes, este tipo de datos suelen modelarse bajo la distribución Weibull,

6. Aplicación

Tabla 6.1: Descriptivos básicos de datos de diámetros para cuatro edades distintas

Edad	Min.	1erQu.	Median	Media	3erQu.	Max.	CA	CK
2.5	1.1	7.5	10.3	9.959	12.5	18.5	-0.3	2.6
4.7	1.3	10.38	13.4	13.16	16.1	26.5	-0.2	2.8
6	2.5	11.1	14.6	14.78	18.55	27.4	0.1	2.5
12	8	18.6	22.3	22.45	26	39.7	0.1	2.9

CA=coeficiente de asimetría, CK=coeficiente de curtosis

por lo que la Figura (6.1) muestra el ajuste de dos curvas a los datos, la distribución Weibull y la distribución que deseamos probar, es decir, la normal asimétrica. Los parámetros de la distribución normal asimétrica fueron estimados mediante el método de máxima verosimilitud penalizado usando el paquete “*sn*” de R mientras que los parámetros de la distribución Weibull de tres parámetros fueron obtenidos por el método de L-momentos mediante el paquete “*lmom*” (Hosking, 2017), de R.

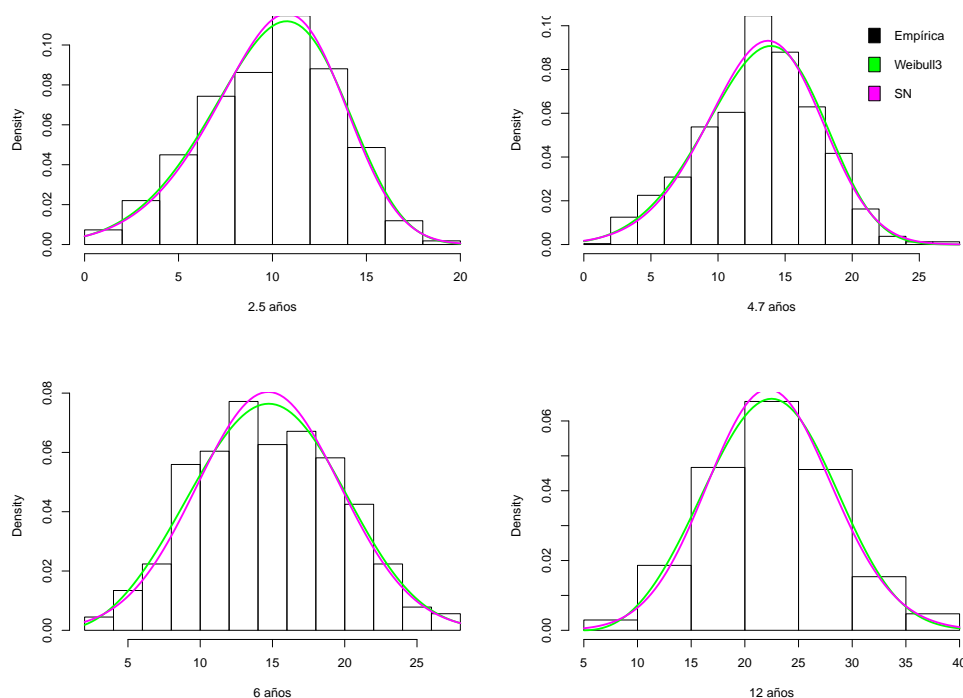


Figura 6.1: Ajuste de una distribución Weibull3 y de una distribución SN al diámetro normal a los 2.5, 4.7, 6 y 12 años.

La Figura 6.2 muestra el ajuste de las distribuciones Weibull y normal asimétrica sobre el histograma de los datos en los cuatro casos.

6. Aplicación

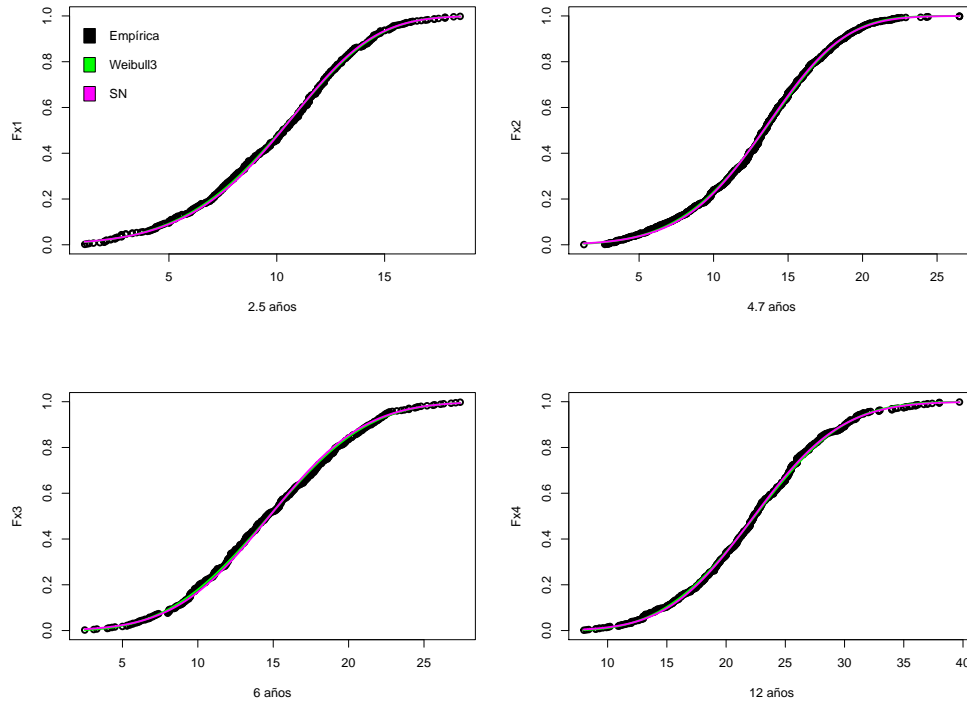


Figura 6.2: Ajuste de la función de distribución acumulada de una distribución Weibull3 y de una distribución SN al diámetro normal a los 2.5, 4.7, 6 y 12 años.

Los parámetros estimados de la muestra al ajustar una distribución $SN(\xi, \omega, \lambda)$ y al ajustar una distribución $Weibull3(\mu, \lambda, k)$ se muestran en la Tabla 6.2.

Tabla 6.2: Parámetros estimados de las distribuciones SN y Weibull3 para datos de diámetro normal a los 2.5, 4.7, 6 y 12 años.

Edad	Distribución SN			Distribución Weibull3		
	$\hat{\xi}$	$\hat{\omega}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\lambda}$	\hat{k}
2.5	13.45	4.98	-1.81	-7.73	19.12	5.72
4.7	16.76	5.64	-1.32	-5.65	20.50	4.95
6	12.87	5.32	0.50	-0.11	16.60	3.27
12	19.04	6.72	0.82	4.88	19.56	3.36

Como lo muestran las Figuras 6.1 y 6.2 la distribución normal asimétrica parece un modelo razonable para modelar los datos, ya que tiene un comportamiento parecido a la distribución empírica de las observaciones.

En la Tabla 6.3, se muestran los resultados de las pruebas de bondad de ajuste propuestas

6. Aplicación

en este trabajo. Se utilizó un $\alpha = 0.05$ para todas las pruebas y 10, 000 muestras bootstrap para el caso de las pruebas A (2) y B.

Tabla 6.3: Pruebas de bondad de ajuste para la distribución normal asimétrica aplicada a datos de diámetro normal para cuatro edades distintas

	2.5 años	4.7 años	6 años	12 años
Prueba	Decisión	Decisión	Decisión	Decisión
A (1)	No Rechazar	Rechazar	No rechazar	No rechazar
A (2)	No rechazar	Rechazar	No rechazar	No rechazar
B	No rechazar	Rechazar	Rechazar	No rechazar
C	No rechazar	Rechazar	No rechazar	No rechazar

Para tres de los cuatro conjuntos de datos las pruebas de bondad de ajuste no descartan el uso de la distribución normal asimétrica como modelo apropiado para los datos explorados. Sin embargo, para el caso de la edad 4.7 años, las pruebas rechazan que las observaciones provengan de una distribución normal asimétrica a un nivel $\alpha = 0.05$. Este conjunto de datos en particular tiene 1200 observaciones. En el caso del conjunto de datos de 6 años de edad, la prueba B rechaza que provengan una población con distribución normal asimétrica mientras que las otras pruebas no rechazan que los datos sean de una población normal asimétrica.

Capítulo 7

Conclusiones

De los resultados obtenidos en los estudios de simulación se puede concluir lo siguiente:

1. Con base en las propiedades de la distribución normal asimétrica se lograron obtener cuatro pruebas de bondad de ajuste cuyo desempeño es satisfactorio en términos de su potencia contra las distribuciones alternativas estudiadas.
2. Mediante simulación Monte Carlo fue posible obtener la potencia y el nivel de las pruebas propuestas, todas ellas preservan el nivel de prueba fijado en $\alpha = 0.05$.
3. Como parte de la simulación Monte Carlo, la implementación del bootstrap paramétrico genera buenos resultados como método de aproximación de la distribución bajo la hipótesis nula de la estadística de prueba, ya que se preserva el nivel de las pruebas y éstas alcanzan potencias elevadas en muchos casos.
4. Respecto al análisis comparativo con otros trabajos existentes, la potencia de las pruebas propuestas es más alta que la de las pruebas contra las que se compararon bajo muchas de las distribuciones alternativas consideradas, especialmente bajo distribuciones asimétricas con tamaños de muestra moderados (50 y 100).
5. Las propuestas A (1) y C son computacionalmente más eficientes en cuanto al tiempo de ejecución que las pruebas ya existentes r_n^* , D_n . Además estas propuestas tienen el mismo desempeño en términos de potencia que las otras pruebas al probar con las distribuciones alternativas asimétricas de cola pesada seleccionadas, para tamaños de muestra mayores a 300.

Referencias

- Arellano-Valle, R. B. y Azzalini, A. (2008). The centered parametrization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis*, 99, 7, 1362–1382.
- Arnold, B. C., Beaver, R. J., Azzalini, A., Balakrishnan, N., Bhaumik, A., Dey, D., Cuadras, C. y Sarabia, J. M. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *Test*, 11, 1, 7–54.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 2, 171–178.
- Azzalini, A. (2005). The Skew-normal Distribution and Related Multivariate Families. *Scandinavian Journal of Statistics*, 32, 159–188. ISSN 0303-6898.
- Azzalini, A. (2018). *The R package sn: The Skew-Normal and Related Distributions such as the Skew-t (version 1.5-2)*. Università di Padova, Italia.
- Azzalini, A. y Arellano-Valle, R. B. (2013). Maximum penalized likelihood estimation for skew-normal and skew-t distributions. *Journal of Statistical Planning and Inference*, 143, 2, 419–433.
- Azzalini, A. y Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61, 3, 579–602.
- Azzalini, A. y Capitanio, A. (2014). *The Skew-Normal and Related Families*. Institute of Mathematical Statistics Monographs. Cambridge University Press. ISBN 978-1-107-02927-9.
- Azzalini, A. y Dalla-Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, 83, 4, 715–726.
- Babu, G. J. y Rao, C. (2004). Goodness-of-fit tests when parameters are estimated. *Sankhya: The Indian Journal of Statistics*, 66, 1, 63–74.
- Bellosta, C. J. G. (2011). *ADGofTest: Anderson-Darling GoF test*. R package version 0.3.

Referencias

- Birnbaum, Z. W. (1950). Effect of linear truncation on a multinormal population. *The Annals of Mathematical Statistics*, 21, 2, 272–279.
- D' Agostino, R. B. y Stephens, M. A. (1986). Goodness-of-fit techniques. *Statistics: Textbooks and Monographs*, New York: Dekker, 1986, edited by D' Agostino, Ralph B.; Stephens, A.
- Dalla-Valle, A. (2007). A test for the hypothesis of skew-normality in a population. *Journal of Statistical Computation and Simulation*, 77, 1, 63–77.
- Figueiredo, F. y Gomes, M. I. (2013). The Skew-normal Distribution in SPC. *REVSTAT-Statistical Journal*, 11, 1, 83–104.
- Firth, D. (1993). Bias Reduction of Maximum Likelihood Estimates. *Biometrika*, 80, 1, 27–38.
- Genton, M. G. (2004). *Skew-elliptical distributions and their applications: a journey beyond normality*. CRC Press.
- Gupta, A. K. y Chen, T. (2001). Goodness-of-fit tests for the skew-normal distribution. *Communications in Statistics - Simulation and Computation*, 30, 4, 907–930.
- Gupta, A. K., Nguyen, T. T. y Sanqui, J. A. T. (2004). Characterization of the skew-normal distribution. *Annals of the Institute of Statistical Mathematics*, 56, 2, 351–360.
- Hennig, C. (2012). *smoothest: Smoothed M-estimators for 1-dimensional location*. R package version 0.1-2.
- Henze, N. (1986). A probabilistic representation of the 'skew-normal' distribution. *Scandinavian journal of statistics*, 271–275.
- Hosking, J. R. M. (2017). *L-Moments*. R package, version 2.6.
- Jiménez-Gamero, M. D., Alba-Fernández, M. V., Jodrá, P. y Chalco-Cano, Y. (2015). Testing for the symmetric component in skew distributions. *Mathematical Methods in the Applied Sciences*, 39, 16, 4713–4722.
- Kazemi, R. y Noorzadeh, M. (2015). A Comparison Between Skew-logistic and Skew-normal Distributions. *Matematika*, 31, 1, 15–24.
- Marsaglia, G. y Marsaglia, J. (2004). Evaluating the Anderson-Darling Distribution. *Journal of Statistical Software*, 19, i02.
- Mateu-Figueras, G., Puig, P. y Pewsey, A. (2007). Goodness-of-fit tests for the skew-normal distribution when the parameters are estimated from the data. *Communications in Statistics-Theory and Methods*, 36, 9, 1735–1755.

Referencias

- Meintanis, S. G. (2007). A Kolmogorov-Smirnov type test for skew normal distributions based on the empirical moment generating function. *Journal of Statistical Planning and Inference*, 137, 8, 2681–2688.
- Meintanis, S. G. y Bassiakos, Y. (2007). Data-transformation and test of fit for the generalized Pareto Hypothesis. *Communications in Statistics-Theory and Methods*, 36, 4, 833–849.
- Meintanis, S. G., Gamero, M. D. J. y Alba-Fernández, V. (2014). A class of goodness-of-fit tests based on transformation. *Communications in Statistics-Theory and Methods*, 43, 8, 1708–1735.
- Nelson, L. S. (1964). The sum of values from a normal and a truncated normal distribution. *Technometrics*, 6, 4, 469–471.
- Noughabi, H. A. y Arghami, N. R. (2011). Monte Carlo comparison of seven normality tests. *Journal of Statistical Computation and Simulation*, 81, 8, 965–972.
- O’hagan, A. y Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, 63, 1, 201–203.
- Patrício, M., Ferreira, F., Oliveiros, B. y Caramelo, F. (2017). Comparing the performance of normality tests with ROC analysis and confidence intervals. *Communications in Statistics-Simulation and Computation*, 46, 10, 7535–7551.
- Pérez-Rodríguez, P. y Villaseñor, J. A. (2008). Comparación de poblaciones normales asimétricas. En *Memoria del XXII Foro Nacional de Estadística*. INEGI, México.
- Pérez-Rodríguez, P. y Villaseñor, A. J. A. (2010). On testing the skew normal hypothesis. *Journal of Statistical Planning and Inference*, 140, 11, 3148–3159.
- Pérez-Rodríguez, P., Villaseñor, J. A., Pérez, S. y Suárez, J. (2017). Bayesian Estimation for the Centered Parameterization of the Skew-Normal Distribution. *Revista Colombiana de Estadística*, 40, 1, 123–140.
- Pourahmadi, M. (2000). *Construction of Skew-Normal Random Variables: Are They Linear Combinations of Normal and Half-Normal*. Division of Statistics, Northern Illinois University.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ribatet, M. (2018). *SpatialExtremes: Modelling Spatial Extremes*. R package version 2.0-6.
- Roberts, C. (1966). A correlation model useful in the study of twins. *Journal of the American Statistical Association*, 61, 316, 1184–1190.

Referencias

- Romao, X., Delgado, R. y Costa, A. (2010). An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation*, 80, 5, 545–591.
- Sarisoy, E. E., Potas, N. y Kara, M. (2014). A simulation study goodness-of-fit tests for the skewed normal distribution. En S. Banerjee y S. S. Ercetin, eds., *Chaos, Complexity and Leadership 2012*, cap. 36, 277–283. Springer.
- Sartori, N. (2006). Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions. *Journal of Statistical Planning and Inference*, 136, 12, 4259–4275.
- Shapiro, S. S. y Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, 3/4, 591–611.
- Swihart, B. y Lindsey, J. (2017). *rmutil: Utilities for Nonlinear Regression and Repeated Measurements Models*. R package version 1.1.0.
- Thas, O. (2010). *Comparing distributions*. Springer.
- Vernic, R. (2006). Multivariate skew-normal distributions with applications in insurance. *Insurance: Mathematics and Economics*, 38, 2, 413–426.
- Villaseñor, J. A. y González-Estrada, E. (2015). Tests of fit for Inverse Gaussian distributions. *Statistics & Probability Letters*, 105, 189–194.
- Villaseñor, J. A., González-Estrada, E. y Ochoa, A. (2017). On testing the inverse Gaussian distribution hypothesis. *Sankhya B*, 1–15.
- Weinstein, M. (1964). The sum of values from a normal and a truncated normal distribution. *Technometrics*, 6, 4, 104–105.

Anexos

Tamaño de las pruebas para $n = 50$ y $n = 500$

Tabla A.1: Tamaño estimado de las pruebas propuestas con $n= 50$ y $\alpha=0.05$.

Lambda	A (1)	A (2)	B	C
-20	0.06	0.05	0.04	0.06
-15	0.06	0.04	0.03	0.06
-10	0.04	0.04	0.03	0.05
-5	0.02	0.03	0.03	0.04
-2.5	0.04	0.04	0.03	0.05
-2	0.04	0.03	0.04	0.05
-1.5	0.05	0.06	0.04	0.05
-1	0.05	0.06	0.04	0.05
-0.5	0.04	0.06	0.03	0.05
0	0.04	0.05	0.05	0.04
0.5	0.05	0.06	0.03	0.04
1	0.05	0.06	0.04	0.04
1.5	0.05	0.05	0.03	0.05
2	0.03	0.04	0.03	0.04
2.5	0.03	0.04	0.03	0.05
5	0.02	0.03	0.04	0.05
10	0.05	0.04	0.03	0.06
15	0.07	0.04	0.03	0.05
20	0.06	0.04	0.04	0.05

Tabla A.2: Tamaño estimado de las pruebas propuestas con $n= 500$ y $\alpha=0.05$.

Lambda	A (1)	A (2)	B	C
-20	0.06	0.05	0.05	0.05
-15	0.05	0.05	0.03	0.04
-10	0.02	0.03	0.03	0.04
-5	0.02	0.03	0.03	0.04
-2.5	0.02	0.04	0.04	0.05
-2	0.03	0.03	0.04	0.05
-1.5	0.03	0.06	0.04	0.05
-1	0.04	0.05	0.04	0.06
-0.5	0.05	0.05	0.04	0.06
0	0.04	0.04	0.05	0.06
0.5	0.04	0.06	0.03	0.06
1	0.03	0.06	0.03	0.06
1.5	0.02	0.04	0.03	0.05
2	0.02	0.04	0.04	0.05
2.5	0.02	0.03	0.04	0.05
5	0.02	0.04	0.04	0.06
10	0.02	0.04	0.04	0.04
15	0.04	0.05	0.04	0.04
20	0.06	0.04	0.05	0.05

Potencia de las pruebas conforme n crece**Tabla A.3:** Potencia estimada de la prueba A (1) con $n= 50, 100, 300, 500$ y $\alpha=0.05$.

Distribución alternativa	n			
	50	100	300	500
t(12)	0.05	0.06	0.07	0.08
Logística estándar	0.07	0.08	0.13	0.19
t(5)	0.12	0.19	0.39	0.62
t(4)	0.2	0.3	0.66	0.85
Laplace estándar	0.27	0.46	0.88	0.98
Uniforme (0,1)	0.42	0.83	1.00	1.00
t(2)	0.65	0.88	1.00	1.00
Beta(0.5,0.5)	0.96	1.00	1.00	1.00
Cauchy estándar	0.99	1.00	1.00	1.00
Gumbel estándar	0.06	0.07	0.11	0.17
st(0.5,10)	0.06	0.07	0.08	0.11
Beta(1,3)	0.2	0.21	0.33	0.43
st(0.5,5)	0.13	0.19	0.43	0.61
Lognormal (0,0.5)	0.17	0.26	0.59	0.79
Laplace asimétrica (1/2)	0.64	0.91	1.00	1.00
Exponencial estándar	0.82	0.96	1.00	1.00
Pareto generalizada (0.15)	0.94	1.00	1.00	1.00
Gamma(0.5,1)	1.00	1.00	1.00	1.00
Weibull (0.75, 1)	1.00	1.00	1.00	1.00
st(0.5,1)	0.98	1.00	1.00	1.00
sCauchy(1)	0.99	1.00	1.00	1.00
Lognormal (0,1.5)	1.00	1.00	1.00	1.00
Pareto Generalizada (0.5)	1.00	1.00	1.00	1.00

Tabla A.4: Potencia estimada de la prueba A (2) con $n= 50, 100, 300, 500$ y $\alpha=0.05$.

Distribución alternativa	n			
	50	100	300	500
t(12)	0.05	0.07	0.08	0.11
Logística estándar	0.08	0.09	0.19	0.25
t(5)	0.15	0.23	0.45	0.69
t(4)	0.24	0.36	0.72	0.90
Laplace estándar	0.29	0.48	0.90	0.98
Uniforme (0,1)	0.33	0.83	1.00	1.00
t(2)	0.69	0.89	1.00	1.00
Beta(0.5,0.5)	0.87	1.00	1.00	1.00
Cauchy estándar	0.99	1.00	1.00	1.00
Gumbel estándar	0.07	0.11	0.28	0.42
st(0.5,10)	0.07	0.08	0.12	0.17
Beta(1,3)	0.12	0.14	0.23	0.34
st(0.5,5)	0.16	0.23	0.52	0.76
Lognormal (0,0.5)	0.17	0.28	0.68	0.89
Laplace asimétrica (1/2)	0.65	0.95	1.00	1.00
Exponencial estándar	0.72	0.92	1.00	1.00
Pareto generalizada (0.15)	0.90	1.00	1.00	1.00
Gamma(0.5,1)	1.00	1.00	1.00	1.00
Weibull (0.75, 1)	0.99	1.00	1.00	1.00
st(0.5,1)	0.98	1.00	1.00	1.00
sCauchy(1)	0.99	1.00	1.00	1.00
Lognormal (0,1.5)	1.00	1.00	1.00	1.00
Pareto Generalizada (0.5)	1.00	1.00	1.00	1.00

Tabla A.5: Potencia estimada de la prueba B con $n= 50, 100, 300, 500$ y $\alpha=0.05$.

Distribución alternativa	n			
	50	100	300	500
t(12)	0.07	0.13	0.25	0.39
Logística estándar	0.12	0.24	0.57	0.80
t(5)	0.25	0.45	0.88	0.98
t(4)	0.34	0.62	0.96	1.00
Laplace estándar	0.50	0.83	1.00	1.00
Uniforme (0,1)	0.40	0.94	1.00	1.00
t(2)	0.83	0.98	1.00	1.00
Beta(0.5,0.5)	0.89	1.00	1.00	1.00
Cauchy estándar	1.00	1.00	1.00	1.00
Gumbel estándar	0.08	0.12	0.30	0.44
st(0.5,10)	0.10	0.14	0.35	0.53
Beta(1,3)	0.10	0.16	0.23	0.39
st(0.5,5)	0.23	0.45	0.88	0.98
Lognormal (0,0.5)	0.17	0.30	0.66	0.88
Laplace asimétrica (1/2)	0.66	0.94	1.00	1.00
Exponencial estándar	0.71	0.93	1.00	1.00
Pareto generalizada (0.15)	0.91	0.99	1.00	1.00
Gamma(0.5,1)	1.00	1.00	1.00	1.00
Weibull (0.75, 1)	0.99	1.00	1.00	1.00
st(0.5,1)	1.00	1.00	1.00	1.00
sCauchy(1)	1.00	1.00	1.00	1.00
Lognormal (0,1.5)	1.00	1.00	1.00	1.00
Pareto Generalizada (0.5)	1.00	1.00	1.00	1.00

Tabla A.6: Potencia estimada de la prueba C con $n= 50, 100, 300, 500$ y $\alpha=0.05$.

Distribución alternativa	n			
	50	100	300	500
t(12)	0.06	0.08	0.15	0.22
Logística estándar	0.08	0.12	0.27	0.42
t(5)	0.17	0.31	0.67	0.88
t(4)	0.25	0.43	0.86	0.97
Laplace estándar	0.23	0.38	0.87	0.99
Uniforme (0,1)	0.22	0.79	1.00	1.00
t(2)	0.67	0.91	1.00	1.00
Beta(0.5,0.5)	0.92	1.00	1.00	1.00
Cauchy estándar	0.97	1.00	1.00	1.00
Gumbel estándar	0.10	0.14	0.30	0.44
st(0.5,10)	0.06	0.09	0.21	0.30
Beta(1,3)	0.16	0.21	0.32	0.44
st(0.5,5)	0.18	0.31	0.68	0.86
Lognormal (0,0.5)	0.21	0.34	0.70	0.90
Laplace asimétrica (1/2)	0.37	0.62	0.99	1.00
Exponencial estándar	0.54	0.80	1.00	1.00
Pareto generalizada (0.15)	0.93	0.99	1.00	1.00
Gamma(0.5,1)	1.00	1.00	1.00	1.00
Weibull (0.75, 1)	0.95	1.00	1.00	1.00
st(0.5,1)	0.97	1.00	1.00	1.00
sCauchy(1)	0.98	1.00	1.00	1.00
Lognormal (0,1.5)	1.00	1.00	1.00	1.00
Pareto Generalizada (0.5)	0.98	1.00	1.00	1.00

Códigos en el programa R

El código siguiente permite simular números pseudoaleatorios de una distribución Gumbel(ξ, θ).

```
rgumbel_w=function(n=100, xi=0, theta=1)
{
  u=runif(n)
  x=-theta*log(-log(u))+xi
  return(x)
}
```

Se muestran además los códigos necesarios para utilizar las pruebas propuestas en el presente trabajo. Para ejecutar los códigos siguientes es necesario instalar los paquetes “*sn*” y “*ADGofTest*” como sigue:

```
install.packages("sn")
install.packages("ADGofTest")
library(sn)
library(ADGofTest)
```

El código para ejecutar la prueba A (1) es el siguiente:

```
st_calc=function(n)
{
  muestra=rsn(n,0,1,20)
  cp_w=sn.mple(y=muestra, opt.method="nlminb")$cp
  w=cp2dp(cp_w, family="SN")
  Y_s=((muestra-w[1])/w[2])^2
  AD_calc=ad.test(Y_s, pchisq, 1)$statistic
  return(AD_calc)
}

F_k_alpha=function(n, alpha)
{
  AD_dist=replicate(5000, st_calc(n))
  orden=sort(AD_dist)
```

```
    cuant=5000*(1-alpha)
    k_alpha=orden [ cuant ]
    return(k_alpha)
}

pal=function(muestra , alpha)
{
  rechazar=0
  n=length(muestra)
  cp_w=sn.mple(y=muestra , opt.method="nlminb")$cp
  w=cp2dp(cp_w , family="SN")
  Y_s=((muestra-w[1])/w[2])^2
  AD_calc=ad.test(Y_s , pchisq , 1) $statistic
  k_alpha=F_k_alpha(n , alpha)
  if (AD_calc > k_alpha)
  {
    rechazar=1
  }
  return(rechazar)
}
```

El código para ejecutar la prueba A (2) es el siguiente:

```
st_calc=function(muestra)
{
  cp_w=sn.mple(y=muestra , opt.method="nlminb")$cp
  w=cp2dp(cp_w , family="SN")
  Y_s=((muestra-w[1])/w[2])^2
  AD_calc=ad.test(Y_s , pchisq , 1) $statistic
  return(AD_calc)
}

boot=function(muestra , alpha , B)
{
  rechazar=0
  n=length(muestra)
  cp_w=sn.mple(y=muestra , opt.method="nlminb")$cp
```

```
w=cp2dp(cp_w , family="SN")
Y_s=((muestra-w[1])/w[2])^2
AD_calc=ad.test(Y_s, pchisq, 1) $statistic
AD_dist=replicate(B, st_calc(rsn(n,w[1],w[2],w[3])))
orden=sort(AD_dist)
cuant=B*(1-alpha)
k_alpha=orden[ cuant ]
if (AD_calc > k_alpha)
{
  rechazar=1
}
return(rechazar)
}
```

El código para ejecutar la prueba B es el siguiente:

```
AD_st_calc=function(muestra)
{
  cp_w=sn.mple(y=muestra, opt.method="nlminb") $cp
  w=cp2dp(cp_w, family="SN")
  AD_calc=ad.test(muestra, psn, w[1], w[2], w[3]) $statistic
  return(AD_calc)
}

AD_boot=function(muestra, alpha, B)
{
  rechazar=0
  n=length(muestra)
  cp_w=sn.mple(y=muestra, opt.method="nlminb") $cp
  w=cp2dp(cp_w, family="SN")
  AD_calc=ad.test(muestra, psn, w[1], w[2], w[3]) $statistic
  AD_dist=replicate(B, AD_st_calc(rsn(n,w[1],w[2],w[3])))
  orden=sort(AD_dist)
  cuant=B*(1-alpha)
  k_alpha=orden[ cuant ]
  if (AD_calc > k_alpha)
  {
```

```
rechazar=1
}
return(rechazar)
}
```

El código para ejecutar la prueba C es el siguiente:

```
u=function(n)
{
  al=replicate(n,runif(1))
  aleatorio=ifelse(al<0.5,-1,1)
  return(aleatorio)
}

prueba_SW=function(muestra,alpha)
{
  rechazar=0
  n=length(muestra)
  cp_w=sn.mple(y=muestra,opt.method="nlminb",penalty="Qpenalty")$cp
  w=cp2dp(cp_w,family="SN")
  Y_s=abs(muestra-w[1])
  un=u(n)
  X=Y_s*un
  prueba=shapiro.test(X)$p.value
  if (prueba < alpha)
  {
    rechazar=1
  }
  return(rechazar)
}
```

El código para ejecutar la prueba r_n^* es el siguiente:

```
semiparametric <- function(datos)
{
```

Códigos en el programa R

```
n=length(datos)
X =as.matrix(rep(1,n))
qrX=qr(X)
s=sqrt(sum(qr.resid(qrX, datos)^2)/n)
gammal=sum(qr.resid(qrX, datos)^3)/(n*s^3)
if(abs(gammal) > 0.99527) gammal=sign(gammal)*0.95
cp2dp(c(qr.coef(qrX, datos),s,gammal),family="SN")[3]
}
```

```
ccrsn=function(x)
{
  theta=semiparametric(x)
  y=sort(x)
  n=length(y)
  z=qsn((1:n)/(n+1),0,1,theta, solver = "RFB")
  cor(y,z)
}
```

```
Critico.Bootstrap=function(R=1000,x,alpha=0.05)
{
  theta=semiparametric(x)
  n=length(x)
  ccrsnboot=rep(0,R)
  for(i in 1:R)
  {
    muestraboot=rsn(n,0,1,theta)
    ccrsnboot[i]=ccrsn(muestraboot)
  }
  quantile(ccrsnboot,alpha)
}
```

```
Phix=function(datos,alpha=0.05)
{
  cmuestral=ccrsn(datos)
  if(cmuestral<Critico.Bootstrap(1000,datos,alpha)) return(1)
  else return(0)
}
```

El código para ejecutar la prueba \tilde{D}_n es el siguiente:

```
T_est=function(muestra,Vn)
{
  n=length(muestra)
  t=seq(-0.2,0.2,0.01)
  Dn_t=rep(NA,length(t))
  for(i in 1:length(t))
  {
    Dn_t[i]=((1/n)*(sum(muestra*exp(t[i]*muestra)))-
              (t[i]*(1/n)*(sum(exp(t[i]*muestra))))-
              (Vn*sqrt(2/pi)*exp(((t[i])^2)*(1-(Vn^2))/2)))
  }

  t_est=(sqrt(n))*(max(Dn_t))
  return(t_est)
}

Tst_calc=function(muestra)
{
  n=length(muestra)
  Y_barra=mean(muestra)
  Sn_2=(1/n)*(sum((muestra-Y_barra)^2))
  m3=(1/n)*(sum((muestra-Y_barra)^3))
  m3=ifelse(abs(m3)<(((4-pi)*sqrt(2))/(pi-2)^(3/2))*
            (Sn_2)^(3/2)), m3, sign(m3)*
            (((4-pi)*sqrt(2))/(pi-2)^(3/2))*(Sn_2)^(3/2)-0.01))
  fi=m3/(4-pi)*sqrt((pi^3)/2)
  dn=Y_barra-(sqrt(2/pi)*sign(fi)*(abs(fi))^(1/3))
  Cn= sqrt(Sn_2+((2/pi)*(sign(fi)*(abs(fi))^(1/3))^2))
  Vn=((Y_barra-dn)/Cn)*sqrt(pi/2)
  lambda=Vn/sqrt(1-(Vn^2))
  X=(muestra-dn)/Cn
  T_calc=T_est(X,Vn)
  return(T_calc)
}
```

Códigos en el programa R

```
T_boot=function (muestra , alpha ,B)
{
  rnr=0
  n=length (muestra)
  Y_barra=mean(muestra)
  Sn_2=(1/n)*(sum((muestra-Y_barra)^2))
  m3=(1/n)*(sum((muestra-Y_barra)^3))
  m3=ifelse (abs(m3)<(((4-pi)*sqrt(2))/(pi-2)^(3/2))*
            (Sn_2)^(3/2)), m3, sign(m3)*
            (((4-pi)*sqrt(2))/(pi-2)^(3/2))*(Sn_2)^(3/2)-0.01))
  fi=m3/(4-pi)*sqrt((pi^3)/2)
  dn=Y_barra-(sqrt(2/pi)*sign(fi)*(abs(fi))^(1/3))
  Cn= sqrt(Sn_2+((2/pi)*(sign(fi)*(abs(fi))^(1/3))^2))
  Vn=((Y_barra-dn)/Cn)*sqrt(pi/2)
  lambda=Vn/sqrt(1-(Vn^2))
  X=(muestra-dn)/Cn
  T_calc=T_est(X,Vn)
  T_dist=replicate(B, Tst_calc(rsn(n,0,1,lambda)))
  orden=sort(T_dist)
  cuant=B*(1-alpha)
  k_alpha=orden[cuant]+((1-alpha)*(orden[cuant+1]-orden[cuant]))
  if (T_calc > k_alpha)
  {
    rnr=1
  }
  return(rnr)
}
```


Datos utilizados en la Sección Aplicación

Tabla A.7: Diámetros de árboles de una plantación comercial de eucalipto para cuatro edades diferentes

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	8.6	4.7	12.7	4.7	12.5	6	16.7
2.5	6.1	4.7	16.3	4.7	12.3	6	9.8
2.5	1.2	4.7	14.1	4.7	16	6	11.1
2.5	4.5	4.7	12.1	4.7	22	6	19.4
2.5	4.2	4.7	4	4.7	19.5	6	13.9
2.5	4.1	4.7	16.6	4.7	14.6	6	9.5
2.5	2.5	4.7	14.3	4.7	15.4	6	24
2.5	5.9	4.7	12.2	4.7	3	6	16.5
2.5	4.8	4.7	7.9	4.7	18	6	18
2.5	6.9	4.7	13	4.7	9.4	6	22.6
2.5	8.5	4.7	15	4.7	18.6	6	14.4
2.5	8.3	4.7	16.2	4.7	15.6	6	17.7
2.5	7.4	4.7	11.7	4.7	4.7	6	15.5
2.5	1.1	4.7	11.4	4.7	19.7	6	19
2.5	8.4	4.7	9.8	4.7	10.5	6	18.2
2.5	8.1	4.7	9.2	4.7	24.3	6	19
2.5	8.9	4.7	16.1	4.7	11.3	6	19.8
2.5	5.3	4.7	11.6	4.7	13.5	6	20.5
2.5	4.8	4.7	13.9	4.7	16.3	6	19.2
2.5	9	4.7	14.2	4.7	11.8	6	17.4
2.5	10.9	4.7	14.6	4.7	7.6	6	16.2
2.5	2.1	4.7	8.3	4.7	13.9	6	20.8
2.5	7	4.7	13.2	4.7	17.9	6	25.6
2.5	6	4.7	14.3	4.7	13.9	6	14.1
2.5	4.6	4.7	14.6	4.7	14.8	6	18.7
2.5	4.7	4.7	13.2	4.7	14.5	6	19.2
2.5	6.5	4.7	14.9	4.7	5	6	23.5
2.5	8.7	4.7	11.1	4.7	12	6	27.4
2.5	10	4.7	14.5	4.7	7.6	6	16.1
2.5	4.3	4.7	9	4.7	12.4	6	6.8
2.5	6.7	4.7	13.7	4.7	6.4	6	21.1
2.5	2.8	4.7	12.7	4.7	4	6	17.8

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	5.5	4.7	15.7	4.7	14	6	19
2.5	8.1	4.7	13	4.7	17.4	6	22.4
2.5	10.1	4.7	15.3	4.7	7.5	6	24.6
2.5	6.4	4.7	12.3	4.7	15.1	6	19
2.5	7.7	4.7	14.5	4.7	10.7	6	20.5
2.5	8.3	4.7	13.7	4.7	7.8	6	19.6
2.5	8	4.7	10	4.7	12.9	12	19.7
2.5	8.5	4.7	14.3	4.7	20.4	12	16.2
2.5	6.3	4.7	12.3	4.7	8.4	12	19
2.5	5.8	4.7	12.5	4.7	5	12	17.5
2.5	9.6	4.7	14.2	4.7	6.2	12	30.5
2.5	6.5	4.7	13	4.7	12.4	12	25.6
2.5	11.2	4.7	13.8	4.7	13.3	12	23.4
2.5	12.2	4.7	10.8	4.7	16.4	12	31.2
2.5	10.3	4.7	13	4.7	18.9	12	20.1
2.5	9.3	4.7	14.3	4.7	11.6	12	23.4
2.5	11.5	4.7	12.7	4.7	4.2	12	17.5
2.5	10.6	4.7	16.7	4.7	3.4	12	24
2.5	6.1	4.7	14.5	4.7	12.7	12	16.5
2.5	14.3	4.7	7.4	4.7	12.3	12	21.3
2.5	11	4.7	13.7	4.7	15.9	12	17.5
2.5	10.5	4.7	13.7	4.7	13.4	12	23.5
2.5	8.6	4.7	17.8	4.7	8.2	12	19.2
2.5	8.4	4.7	14.3	4.7	13.1	12	20.3
2.5	9.5	4.7	10.2	4.7	18.5	12	15.4
2.5	5.6	4.7	15.2	4.7	14.3	12	19.2
2.5	14.5	4.7	14	4.7	3	12	15.3
2.5	11.9	4.7	14.9	4.7	13.3	12	20.2
2.5	12.1	4.7	15.5	4.7	10.7	12	23.3
2.5	13.1	4.7	16.4	4.7	17.9	12	14
2.5	10.1	4.7	16.3	4.7	3	12	22.2
2.5	14.8	4.7	13.9	4.7	14.5	12	25.5
2.5	11.2	4.7	14.8	4.7	12.8	12	34.5
2.5	11.3	4.7	21.2	4.7	14.9	12	11.5
2.5	6.1	4.7	15.3	4.7	10.9	12	27.4
2.5	13	4.7	8.2	4.7	9.4	12	12.1
2.5	7.5	4.7	20.2	4.7	16.3	12	21.3

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	12.8	4.7	7.8	4.7	15	12	26
2.5	10.6	4.7	11.7	4.7	6.1	12	27.3
2.5	8.1	4.7	16.5	4.7	15.5	12	25.3
2.5	11.9	4.7	13.6	4.7	11	12	17.4
2.5	9.9	4.7	8.4	4.7	12	12	22.8
2.5	12.7	4.7	15	4.7	8.8	12	18.2
2.5	10.9	4.7	18.54	4.7	13	12	12.4
2.5	10.3	4.7	13.3	4.7	9.6	12	16
2.5	12	4.7	20.8	4.7	5.9	12	26.2
2.5	11.5	4.7	19	4.7	14.6	12	13.2
2.5	9.9	4.7	3	4.7	12.7	12	16.2
2.5	6.7	4.7	14.5	4.7	20.2	12	21
2.5	9.1	4.7	19.7	4.7	4.1	12	28.2
2.5	12.3	4.7	14.8	4.7	13.8	12	28.2
2.5	11.2	4.7	16.2	4.7	16.4	12	12.3
2.5	6.7	4.7	10.7	4.7	13	12	22
2.5	9.4	4.7	16.8	4.7	14.7	12	18.2
2.5	4.4	4.7	17.9	4.7	10	12	23.1
2.5	11.1	4.7	9.4	4.7	26.5	12	28.2
2.5	7.9	4.7	4.2	4.7	7.5	12	25.6
2.5	6.8	4.7	16.4	4.7	5.5	12	25.7
2.5	7	4.7	15.6	4.7	15.7	12	17.8
2.5	7.1	4.7	19.1	4.7	7.1	12	30.3
2.5	7.9	4.7	13.9	4.7	9.5	12	23.1
2.5	5	4.7	11.2	4.7	8.3	12	11.2
2.5	4.7	4.7	14.9	4.7	8.2	12	18
2.5	2.6	4.7	11.8	4.7	6.9	12	17.6
2.5	8.9	4.7	9.9	4.7	9.6	12	20.9
2.5	9	4.7	1.3	4.7	12.4	12	15.5
2.5	6.2	4.7	8.6	4.7	6.1	12	24.6
2.5	2.3	4.7	9.9	4.7	7.9	12	12.6
2.5	9.7	4.7	15.7	4.7	10.3	12	27.2
2.5	10.2	4.7	13.4	4.7	13	12	18.8
2.5	8.5	4.7	9.1	4.7	13	12	24.5
2.5	8.9	4.7	6.2	4.7	12.9	12	30.8
2.5	5.3	4.7	14.3	4.7	4	12	15
2.5	8.9	4.7	3.5	4.7	10.9	12	24.3

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	7.5	4.7	17.8	4.7	7.4	12	17.2
2.5	8.7	4.7	2.7	4.7	9.8	12	15.9
2.5	6.9	4.7	10	4.7	13.7	12	24.3
2.5	7.6	4.7	8.7	4.7	3.5	12	22.6
2.5	8.7	4.7	13.4	4.7	13.2	12	25.4
2.5	10.5	4.7	12.9	4.7	14.6	12	20.9
2.5	5.1	4.7	14.1	4.7	3.7	12	20
2.5	5.1	4.7	9.8	4.7	9.8	12	19.2
2.5	8.3	4.7	10	4.7	11.6	12	27.1
2.5	7.9	4.7	15.1	4.7	16	12	22.3
2.5	5.4	4.7	12.4	4.7	19.3	12	25.4
2.5	11.7	4.7	18.4	4.7	15.6	12	20.9
2.5	9.2	4.7	18.8	4.7	15.6	12	22.7
2.5	7.4	4.7	14	4.7	13.4	12	22
2.5	7.1	4.7	16.4	4.7	9.8	12	30.5
2.5	9	4.7	10	4.7	17.9	12	32.9
2.5	11.5	4.7	13.7	4.7	13.2	12	25.6
2.5	7.2	4.7	12.1	4.7	5.5	12	23.2
2.5	9.7	4.7	17.7	4.7	17.3	12	19.7
2.5	5	4.7	10.9	4.7	6.6	12	20.3
2.5	7.2	4.7	18.5	4.7	18.2	12	18.6
2.5	6.2	4.7	6.8	4.7	13	12	15.9
2.5	5.4	4.7	17.8	4.7	15.9	12	17.1
2.5	4.4	4.7	16.9	4.7	11.7	12	25
2.5	6.4	4.7	11.2	4.7	18.1	12	28
2.5	6	4.7	17.5	4.7	5.8	12	30
2.5	5.3	4.7	13.6	4.7	22.1	12	19.1
2.5	5	4.7	11.4	4.7	10.1	12	17.9
2.5	7.8	4.7	11.7	4.7	14.3	12	22.3
2.5	7.3	4.7	16.6	4.7	18.2	12	31.2
2.5	9.6	4.7	22.3	4.7	9.8	12	35.7
2.5	7.8	4.7	12.3	4.7	21.8	12	30.9
2.5	7.7	4.7	17.7	4.7	9.5	12	25.5
2.5	7.1	4.7	15.8	4.7	21.9	12	20.4
2.5	5.2	4.7	16.5	4.7	5.5	12	19.3
2.5	10.3	4.7	14.4	4.7	12.4	12	27.6
2.5	7.7	4.7	17.6	4.7	4.8	12	26.5

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	3.3	4.7	16.7	4.7	18.2	12	20.9
2.5	7.3	4.7	14.1	4.7	5.7	12	25.6
2.5	5.9	4.7	18.3	4.7	7.6	12	22.2
2.5	7.1	4.7	13.6	4.7	4.2	12	20.7
2.5	10.4	4.7	17.4	4.7	17.5	12	15.2
2.5	10.3	4.7	13.5	4.7	3.5	12	24.7
2.5	7	4.7	16	4.7	7.2	12	31.6
2.5	4.3	4.7	11.7	4.7	17.6	12	18.6
2.5	5.1	4.7	13.9	4.7	14.9	12	11.2
2.5	3.3	4.7	13.3	4.7	4.4	12	12
2.5	5.8	4.7	12.2	4.7	8	12	18.2
2.5	7.3	4.7	14.4	4.7	11.1	12	20.5
2.5	5.4	4.7	13.8	4.7	9.9	12	10.2
2.5	6.5	4.7	10.8	4.7	20.3	12	14.1
2.5	14.7	4.7	16.1	4.7	15.9	12	26.2
2.5	11.2	4.7	5.6	4.7	12	12	24.9
2.5	8.5	4.7	16.8	4.7	6.3	12	30
2.5	14.4	4.7	16.6	4.7	10.9	12	27.7
2.5	13.3	4.7	5.4	4.7	20.2	12	17.7
2.5	7.7	4.7	11	4.7	10.9	12	13.2
2.5	10.8	4.7	15	4.7	15	12	23.2
2.5	13.6	4.7	14.6	4.7	8.2	12	12
2.5	14.5	4.7	7.2	4.7	16.1	12	28.8
2.5	12.5	4.7	16.4	4.7	14.1	12	20.7
2.5	15.4	4.7	15.6	4.7	3.5	12	22.6
2.5	13.9	4.7	5.7	4.7	8.2	12	17.2
2.5	12.2	4.7	17.5	4.7	15.5	12	14.2
2.5	11.1	4.7	14.7	4.7	16.8	12	21.6
2.5	13.6	4.7	5.1	4.7	5.5	12	23.9
2.5	13.6	4.7	11.6	4.7	15.5	12	18.3
2.5	9.9	4.7	13.1	4.7	19.7	12	13
2.5	7.5	4.7	13	4.7	11.3	12	20.4
2.5	12.6	4.7	16.5	4.7	15	12	30.9
2.5	11.2	4.7	13.4	4.7	16.4	12	25.5
2.5	13	4.7	9.8	4.7	4.8	12	30.5
2.5	12.3	4.7	17.4	4.7	12.3	12	24.6
2.5	10	4.7	12.4	4.7	18.9	12	32.2

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	8.6	4.7	8.2	4.7	5.8	12	26.3
2.5	10.8	4.7	16.9	4.7	13.1	12	20.3
2.5	8.6	4.7	15	4.7	17.3	12	25.4
2.5	8.1	4.7	17.5	4.7	17.5	12	27.9
2.5	12.3	4.7	15.4	4.7	15.2	12	16.1
2.5	13.5	4.7	14.8	4.7	17.9	12	22.4
2.5	12.8	4.7	14.4	4.7	24.4	12	22.9
2.5	10.6	4.7	10.4	4.7	8.5	12	25.3
2.5	13.3	4.7	10.1	4.7	8.5	12	25.6
2.5	12	4.7	3.4	4.7	17.2	12	31.3
2.5	13.7	4.7	14.9	4.7	18.9	12	22.9
2.5	9.7	4.7	15.3	4.7	14	12	19.2
2.5	12.4	4.7	8.5	4.7	20.1	12	20.9
2.5	11.4	4.7	15.7	4.7	5.5	12	20.1
2.5	12.2	4.7	7.7	4.7	13.5	12	22.4
2.5	11.5	4.7	5.9	4.7	19.2	12	25.3
2.5	2.3	4.7	14.7	4.7	13.9	12	22.6
2.5	1.8	4.7	16.5	4.7	5.1	12	29.8
2.5	2.8	4.7	10.9	4.7	5	12	30.5
2.5	2	4.7	18.6	4.7	13.6	12	35.2
2.5	2.8	4.7	16	4.7	14.8	12	21
2.5	2.5	4.7	20.4	4.7	20.1	12	19
2.5	8.2	4.7	6.7	4.7	5.6	12	23.5
2.5	2.6	4.7	13.6	4.7	17.1	12	28.5
2.5	2	4.7	15.6	4.7	18.8	12	20.9
2.5	2.5	4.7	17.4	4.7	12.4	12	22
2.5	5	4.7	6.4	4.7	18.4	12	20.7
2.5	6	4.7	13	4.7	15.7	12	23.6
2.5	8.5	4.7	13.3	4.7	14.6	12	21.1
2.5	2.8	4.7	18	4.7	15	12	20.9
2.5	3	4.7	12.3	4.7	15	12	27.4
2.5	4.1	4.7	7.6	4.7	16.8	12	19.9
2.5	2	4.7	8.2	4.7	6.6	12	21.8
2.5	4.2	4.7	10.4	4.7	19.3	12	17.8
2.5	7.6	4.7	15.6	4.7	9.8	12	26.3
2.5	8	4.7	11.2	4.7	16.4	12	12.6
2.5	11.5	4.7	5.6	4.7	11.3	12	21.3

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	7.5	4.7	17	4.7	13	12	24.7
2.5	3	4.7	13.2	4.7	16.7	12	13.6
2.5	10.2	4.7	22.6	4.7	16.2	12	16.7
2.5	1.5	4.7	14.8	4.7	13.9	12	27.7
2.5	12.6	4.7	15	4.7	16.1	12	21.4
2.5	8	4.7	21	4.7	19.8	12	23.6
2.5	7.3	4.7	11.6	4.7	12.8	12	19.8
2.5	2.1	4.7	12.3	4.7	7.5	12	17.6
2.5	2.1	4.7	15.6	4.7	12.8	12	22.4
2.5	8.7	4.7	12.2	4.7	19.7	12	14.2
2.5	10	4.7	14.4	4.7	19.5	12	16
2.5	6.3	4.7	11.2	4.7	22.4	12	20.8
2.5	2.8	4.7	10.6	4.7	13.6	12	15.7
2.5	8	4.7	16	4.7	10.3	12	26.5
2.5	4.4	4.7	16.8	4.7	19.8	12	25.2
2.5	9.2	4.7	16	4.7	10	12	13.2
2.5	13.4	4.7	10.4	4.7	20	12	19.5
2.5	13.9	4.7	18	4.7	10	12	26.8
2.5	16	4.7	12	4.7	14.7	12	22.8
2.5	10.1	4.7	13	4.7	22.2	12	17.7
2.5	11.9	4.7	19.8	4.7	10.9	12	22.4
2.5	17.3	4.7	12.2	4.7	9.7	12	8
2.5	7.5	4.7	14.2	4.7	15.8	12	31
2.5	13.5	4.7	8.1	4.7	19.4	12	19.9
2.5	14.8	4.7	5.1	4.7	16	12	26.6
2.5	7.1	4.7	12.5	4.7	6.3	12	22.4
2.5	16	4.7	12.4	4.7	17.6	12	17.1
2.5	14	4.7	13.7	4.7	13.5	12	15.4
2.5	13.2	4.7	13.1	4.7	26.5	12	25.3
2.5	15	4.7	12.6	4.7	20.7	12	22.1
2.5	8.7	4.7	7	4.7	13.4	12	19.3
2.5	13	4.7	3.6	4.7	14.3	12	21.2
2.5	12.6	4.7	4.5	4.7	11.6	12	25.9
2.5	11.1	4.7	6.6	4.7	20.5	12	27.1
2.5	13.3	4.7	7.2	4.7	16.5	12	30.5
2.5	11.9	4.7	8.7	4.7	20	12	32
2.5	13	4.7	13.2	4.7	22	12	21.5

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	8	4.7	7	4.7	10.5	12	22.7
2.5	11.8	4.7	12.3	4.7	13.5	12	27.1
2.5	12.3	4.7	14.6	4.7	15.2	12	30.3
2.5	8.5	4.7	9.5	4.7	18	12	32.9
2.5	10.3	4.7	6.3	4.7	22.9	12	19.8
2.5	11.1	4.7	11	4.7	13.5	12	24.3
2.5	11	4.7	11.5	4.7	20.9	12	25.1
2.5	9.7	4.7	4.7	4.7	19.8	12	32.9
2.5	11.1	4.7	14.1	4.7	16.9	12	37.4
2.5	7.6	4.7	11.2	4.7	16.4	12	30
2.5	14.8	4.7	10.6	4.7	19.8	12	18.2
2.5	3.8	4.7	6.4	4.7	20.4	12	20.3
2.5	11.7	4.7	12.2	4.7	13.5	12	21.7
2.5	16.7	4.7	7.8	4.7	19.7	12	26.2
2.5	10	4.7	9.3	4.7	8.9	12	25.4
2.5	13.4	4.7	11.4	4.7	15.9	12	28.2
2.5	10.5	4.7	10.6	4.7	13.7	12	20.9
2.5	12.2	4.7	14	4.7	5.1	12	21.6
2.5	11.3	4.7	6.5	4.7	16.6	12	23.2
2.5	14	4.7	9	4.7	19	12	30.1
2.5	14.5	4.7	15.3	4.7	16.4	12	19.4
2.5	14.1	4.7	13.7	4.7	17.9	12	27
2.5	15	4.7	15.6	4.7	15.6	12	25.2
2.5	11.8	4.7	12	4.7	19.3	12	27.4
2.5	12.4	4.7	9.8	4.7	12.4	12	21.3
2.5	12.2	4.7	14.8	4.7	12.7	12	26.9
2.5	11.8	4.7	11.9	4.7	10.2	12	25.8
2.5	14.3	4.7	10.9	4.7	5	12	22.7
2.5	11.4	4.7	12.8	4.7	15.1	12	18.3
2.5	10.6	4.7	5.7	4.7	11.5	12	16.8
2.5	12.3	4.7	17	4.7	12.9	12	27.4
2.5	8.2	4.7	13.5	4.7	13.7	12	24.4
2.5	10.7	4.7	7.3	4.7	19.3	12	16.8
2.5	7.2	4.7	7.5	4.7	17.9	12	27.9
2.5	10.8	4.7	17.2	4.7	14.3	12	17.8
2.5	8.7	4.7	16	4.7	9.9	12	21.1
2.5	6.2	4.7	14.8	4.7	18.5	12	24.9

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	9.7	4.7	17	4.7	11	12	14.2
2.5	9.1	4.7	17.8	4.7	11.1	12	11.6
2.5	6	4.7	9.4	4.7	21.3	12	23.3
2.5	9.5	4.7	20	4.7	17.8	12	15.2
2.5	11.4	4.7	11.4	4.7	12.2	12	22
2.5	9.8	4.7	14.6	4.7	18.2	12	31.4
2.5	11.1	4.7	16.2	4.7	20.8	12	27.8
2.5	12.2	4.7	9.4	4.7	8	12	22.2
2.5	10.2	4.7	17.4	4.7	21.7	12	25
2.5	12.9	4.7	17.4	4.7	12.9	12	23.3
2.5	10.3	4.7	10.8	4.7	19.1	12	22.2
2.5	10.4	4.7	14	4.7	16.2	12	24.1
2.5	11.9	4.7	14.6	4.7	17	12	12
2.5	10.1	4.7	13.2	4.7	16.2	12	36
2.5	12.7	4.7	15	4.7	10.9	12	10.9
2.5	10.8	4.7	13.6	4.7	17.4	12	12.9
2.5	12	4.7	15	4.7	16.7	12	18.3
2.5	13.2	4.7	12.1	4.7	18.2	12	19.5
2.5	10.1	4.7	12.4	6	10.1	12	24.3
2.5	11.4	4.7	11.3	6	8.5	12	20.6
2.5	2.8	4.7	8.2	6	9.8	12	26.1
2.5	7.7	4.7	12.7	6	11.1	12	15.7
2.5	9.7	4.7	18	6	14	12	21.2
2.5	7	4.7	17.1	6	12.2	12	9.9
2.5	7.1	4.7	15.2	6	11	12	27.8
2.5	7.4	4.7	16.8	6	7	12	12
2.5	10.5	4.7	16.4	6	7.4	12	24.7
2.5	9.1	4.7	14	6	10.6	12	26.5
2.5	8.4	4.7	15	6	13.2	12	14.7
2.5	10.1	4.7	11.2	6	6.1	12	13.9
2.5	9.2	4.7	9.8	6	12.4	12	25.3
2.5	9.4	4.7	17	6	8.9	12	21.4
2.5	7.4	4.7	17.2	6	12.6	12	19.9
2.5	12	4.7	19	6	13.3	12	19.1
2.5	4.5	4.7	9.4	6	9.8	12	22.3
2.5	11	4.7	15.6	6	10.8	12	24.7
2.5	15.3	4.7	19.2	6	10	12	25.3

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	9.6	4.7	10.7	6	10.2	12	29.7
2.5	12.1	4.7	9.6	6	9.5	12	20.5
2.5	9.5	4.7	11.9	6	10.5	12	20
2.5	10.7	4.7	13.2	6	8.1	12	23
2.5	6.8	4.7	12.2	6	3.3	12	30.2
2.5	10.8	4.7	8.5	6	15.8	12	21.9
2.5	12.2	4.7	5.6	6	9.9	12	34
2.5	13.4	4.7	13.5	6	11.1	12	22.9
2.5	12.5	4.7	14.3	6	11.3	12	20.1
2.5	14	4.7	8.8	6	14.4	12	26.8
2.5	5.3	4.7	9.6	6	6	12	27.2
2.5	12.2	4.7	4.3	6	13.2	12	20.1
2.5	14.3	4.7	8.4	6	10.3	12	19.2
2.5	7.5	4.7	11.4	6	9.5	12	21.9
2.5	6.9	4.7	7.7	6	13.4	12	21.7
2.5	7.2	4.7	3.9	6	7.3	12	24.5
2.5	14.8	4.7	13.7	6	13.9	12	25.3
2.5	10.9	4.7	15.5	6	3.1	12	19.6
2.5	11	4.7	16	6	11.5	12	30.2
2.5	14.2	4.7	3.5	6	12	12	18.4
2.5	16.9	4.7	15.6	6	16.9	12	19.1
2.5	14.2	4.7	12.6	6	9	12	25.5
2.5	18.5	4.7	17.7	6	14.5	12	20
2.5	14.2	4.7	14.4	6	2.5	12	21.3
2.5	11.8	4.7	7.8	6	12.8	12	36.2
2.5	16.2	4.7	13.5	6	13.6	12	27.9
2.5	11	4.7	11	6	12.2	12	31.6
2.5	8.7	4.7	8.4	6	9.1	12	28.4
2.5	7	4.7	14.1	6	11.9	12	22.3
2.5	13	4.7	12.4	6	8.6	12	19
2.5	14.6	4.7	11	6	11.3	12	26
2.5	11.5	4.7	10.9	6	9.2	12	24.3
2.5	16.5	4.7	12.3	6	7.1	12	30.1
2.5	13.3	4.7	13.8	6	14.2	12	35.2
2.5	9.5	4.7	13.5	6	5.7	12	19.4
2.5	10.8	4.7	6.6	6	15.9	12	23.1
2.5	9	4.7	10.4	6	14.2	12	27.2

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	11.2	4.7	17.5	6	11.9	12	19.2
2.5	10	4.7	13	6	12.8	12	22.3
2.5	7.4	4.7	13.4	6	9.2	12	23.7
2.5	7.8	4.7	9.9	6	8	12	28
2.5	12.4	4.7	13	6	8.7	12	13
2.5	9.5	4.7	5.6	6	20.9	12	31.6
2.5	11	4.7	15.5	6	16.6	12	29.5
2.5	8.5	4.7	12.4	6	19.1	12	13.6
2.5	11.5	4.7	9.9	6	13.4	12	26
2.5	11.4	4.7	9.6	6	20.3	12	22.4
2.5	4.6	4.7	5.1	6	15.4	12	26.2
2.5	11.6	4.7	9.6	6	9.4	12	17
2.5	10.7	4.7	10.9	6	18.5	12	16.2
2.5	4.4	4.7	3.7	6	19	12	24
2.5	7.8	4.7	9.8	6	21.7	12	20
2.5	14.4	4.7	8	6	15.8	12	31
2.5	12.7	4.7	17	6	17.9	12	15
2.5	14	4.7	13.7	6	12	12	11.8
2.5	9	4.7	17.8	6	17.7	12	17.3
2.5	9	4.7	18	6	12	12	15.5
2.5	10.3	4.7	11.4	6	17.5	12	22.9
2.5	5.6	4.7	12	6	19.8	12	22
2.5	11.9	4.7	15	6	17.1	12	14.5
2.5	9.8	4.7	17.2	6	24.4	12	21.8
2.5	13.2	4.7	9.8	6	16.1	12	18.2
2.5	12.4	4.7	11.2	6	19.7	12	26.3
2.5	10.5	4.7	15.4	6	21.6	12	21
2.5	13	4.7	9.8	6	9.6	12	11
2.5	9.2	4.7	8.2	6	11.8	12	14.9
2.5	5.1	4.7	22.4	6	11.1	12	25.5
2.5	12	4.7	14.2	6	19.8	12	34
2.5	4.4	4.7	18.6	6	8	12	13
2.5	10.2	4.7	20.5	6	15.5	12	9.3
2.5	12.4	4.7	9.8	6	10.9	12	15
2.5	8.5	4.7	11.8	6	12	12	10
2.5	7.3	4.7	10.4	6	14.9	12	31
2.5	13.1	4.7	8.5	6	5.7	12	9.6

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	5	4.7	12	6	8.8	12	13.4
2.5	3.7	4.7	6.4	6	20.3	12	29.7
2.5	6.4	4.7	18	6	11.9	12	23.7
2.5	11.8	4.7	11.8	6	18.9	12	15.7
2.5	14.8	4.7	19.4	6	13	12	15.5
2.5	15.3	4.7	12.2	6	19.7	12	21.1
2.5	15.4	4.7	17.2	6	10.5	12	36.7
2.5	12.9	4.7	8.5	6	18.2	12	17.5
2.5	12	4.7	16.4	6	6.3	12	29.5
2.5	14.3	4.7	17.2	6	12.7	12	19.5
2.5	14.2	4.7	15.8	6	12.2	12	12.8
2.5	12.5	4.7	16.3	6	5.3	12	19.8
2.5	15.4	4.7	17.2	6	17.7	12	20.8
2.5	15.2	4.7	12.4	6	14.8	12	18.7
2.5	12.5	4.7	13	6	17.8	12	20.7
2.5	11.2	4.7	19.8	6	19.1	12	29.7
2.5	15.9	4.7	15.6	6	17.3	12	17.4
2.5	1.3	4.7	12.4	6	5	12	27.8
2.5	10	4.7	11.6	6	10.1	12	31.6
2.5	13.6	4.7	12.7	6	22.3	12	24.2
2.5	11.8	4.7	11.2	6	5.3	12	19.5
2.5	11.3	4.7	10.3	6	8.1	12	24.1
2.5	9.7	4.7	12.1	6	13.3	12	21.1
2.5	12.7	4.7	6.1	6	5.9	12	21.9
2.5	12.7	4.7	14.6	6	18.8	12	18.7
2.5	13	4.7	13.1	6	11.1	12	19.1
2.5	10.6	4.7	15.9	6	11.5	12	18.5
2.5	11.5	4.7	16.5	6	16.8	12	29.3
2.5	11.4	4.7	16	6	10	12	13.1
2.5	10.3	4.7	15.1	6	9	12	30.4
2.5	10.1	4.7	8.1	6	5.5	12	20
2.5	8.1	4.7	10.4	6	14.5	12	28
2.5	10.5	4.7	13.1	6	8	12	19.3
2.5	12.4	4.7	15.1	6	8.9	12	20.8
2.5	10.7	4.7	14.3	6	6.1	12	17.8
2.5	12.9	4.7	14.2	6	7.2	12	18.2
2.5	12.8	4.7	12.7	6	9.4	12	15.2

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	13.7	4.7	16.5	6	10	12	22.3
2.5	12.1	4.7	13	6	6.5	12	31.2
2.5	13.1	4.7	16.4	6	9.8	12	22.6
2.5	11.6	4.7	12.4	6	9	12	18.4
2.5	12.1	4.7	12.8	6	13.5	12	27.6
2.5	12.7	4.7	18.4	6	6.7	12	22.2
2.5	14.5	4.7	9.6	6	4.5	12	24.7
2.5	14.5	4.7	6.1	6	11.2	12	24
2.5	12.2	4.7	14	6	15.9	12	21.5
2.5	13.5	4.7	8	6	7.2	12	25.9
2.5	14.3	4.7	16.2	6	9.3	12	19.9
2.5	9.3	4.7	9.8	6	18.2	12	24.6
2.5	13.5	4.7	17.8	6	11.8	12	26.7
2.5	10	4.7	20.1	6	8.1	12	25
2.5	12.8	4.7	14.9	6	11.9	12	21.4
2.5	10.6	4.7	13.1	6	16.3	12	22.9
2.5	15.3	4.7	13.9	6	4.1	12	18.1
2.5	11.9	4.7	14.6	6	10.1	12	26.8
2.5	16.3	4.7	12.3	6	17.2	12	25.2
2.5	18.2	4.7	15.2	6	14.3	12	23.6
2.5	11.5	4.7	15.3	6	16.5	12	22.7
2.5	13.3	4.7	7.2	6	10.6	12	19.7
2.5	11.4	4.7	4.8	6	15.3	12	20
2.5	11.5	4.7	17.2	6	6.2	12	22.9
2.5	8.5	4.7	13.6	6	20.7	12	30.1
2.5	12.5	4.7	15.5	6	12.1	12	25.9
2.5	11.9	4.7	2.8	6	10.8	12	22.8
2.5	13.6	4.7	7.4	6	12.5	12	25.5
2.5	10.4	4.7	9.6	6	15.5	12	22.3
2.5	13	4.7	14.9	6	11.3	12	26.7
2.5	11.5	4.7	9.2	6	10.5	12	24.2
2.5	14.8	4.7	14	6	16.4	12	21.5
2.5	14.5	4.7	11.4	6	15.4	12	24.8
2.5	12.5	4.7	15.3	6	13.2	12	26.7
2.5	8	4.7	12	6	11.8	12	21.9
2.5	14.4	4.7	14.2	6	15.5	12	21
2.5	9.2	4.7	15.1	6	16.4	12	24

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	13.2	4.7	13.9	6	13.6	12	25.1
2.5	3.5	4.7	12.8	6	11.2	12	22.9
2.5	15.6	4.7	8.7	6	12.2	12	25.7
2.5	8.2	4.7	13.1	6	12.1	12	27.5
2.5	11.4	4.7	10.8	6	11.1	12	24
2.5	9.3	4.7	13.2	6	12.8	12	27
2.5	11.9	4.7	9.8	6	5.5	12	25.5
2.5	8.4	4.7	12.8	6	8.5	12	35
2.5	12.8	4.7	12.3	6	21.3	12	16.2
2.5	8.9	4.7	8.7	6	20.6	12	25
2.5	13.4	4.7	13.2	6	22.2	12	38
2.5	8.2	4.7	9.8	6	12.5	12	13
2.5	6.5	4.7	12.3	6	17.3	12	18
2.5	11	4.7	9.4	6	18.6	12	32
2.5	9.7	4.7	11	6	20.2	12	20.3
2.5	11.6	4.7	9.8	6	17.3	12	24.3
2.5	11.6	4.7	13.6	6	21.5	12	26
2.5	12.3	4.7	13.2	6	24	12	33
2.5	11	4.7	12.8	6	11.2	12	26.6
2.5	11.5	4.7	14.6	6	22.5	12	25
2.5	11.6	4.7	13.2	6	19.8	12	22.2
2.5	15.2	4.7	12.9	6	15.6	12	17
2.5	6.9	4.7	10.8	6	17.3	12	25.3
2.5	14.5	4.7	15.2	6	17.7	12	30
2.5	13.4	4.7	13.6	6	17.9	12	25.2
2.5	14.1	4.7	12.8	6	18.7	12	21.1
2.5	8.9	4.7	14.2	6	14.2	12	26
2.5	15.4	4.7	13.6	6	17	12	38
2.5	10.6	4.7	11.8	6	18	12	23.3
2.5	7.6	4.7	12.4	6	13.1	12	31
2.5	14.2	4.7	13.6	6	17	12	17
2.5	8.6	4.7	9.5	6	20	12	27
2.5	17.8	4.7	12.8	6	19.9	12	16.4
2.5	10.9	4.7	13.1	6	20.7	12	18
2.5	7.9	4.7	20.5	6	8.1	12	28
2.5	15.7	4.7	13	6	21	12	24.5
2.5	12.2	4.7	8.7	6	13.8	12	13.2

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
2.5	9.6	4.7	22	6	17.5	12	22
2.5	17.3	4.7	11.4	6	17.6	12	15
2.5	5.6	4.7	11.7	6	22.1	12	26.2
2.5	6.3	4.7	11.1	6	23.2	12	24
2.5	12.8	4.7	11.6	6	18.4	12	36
2.5	16.2	4.7	12	6	6.8	12	13
2.5	15.9	4.7	12.1	6	20	12	31.1
2.5	4	4.7	12.3	6	18.7	12	23.3
2.5	11	4.7	9	6	10.2	12	8.2
2.5	6.1	4.7	15	6	21.7	12	30.3
2.5	10	4.7	13.8	6	16.1	12	16
2.5	17.8	4.7	12	6	19.7	12	27
2.5	14.7	4.7	10.6	6	12	12	24
2.5	6	4.7	7	6	20.3	12	13
2.5	17	4.7	10.2	6	13	12	25.4
2.5	14.6	4.7	10.8	6	21.9	12	21.2
2.5	13.5	4.7	13.4	6	15.3	12	28.2
2.5	11.8	4.7	8.6	6	20.7	12	18.6
2.5	7.6	4.7	12.2	6	9.1	12	25.2
2.5	14.4	4.7	13.1	6	17.2	12	19.6
2.5	8.2	4.7	14.2	6	15.4	12	17
2.5	12.1	4.7	12.2	6	13	12	26
2.5	17.5	4.7	14	6	21.9	12	22.7
2.5	13.2	4.7	11.1	6	15.5	12	21.5
2.5	15.5	4.7	13.1	6	22	12	18.5
2.5	15.2	4.7	12.8	6	25.7	12	19.5
2.5	9.4	4.7	10.2	6	17.1	12	20.8
2.5	11.2	4.7	8.3	6	15.4	12	22
2.5	12.6	4.7	12.1	6	22.2	12	20.6
2.5	17.8	4.7	11.9	6	18.7	12	28
2.5	11	4.7	13	6	27.1	12	25
2.5	15.8	4.7	10.8	6	25.2	12	27.2
4.7	12.8	4.7	12.4	6	26.1	12	16.3
4.7	11.6	4.7	7.6	6	22.2	12	25.5
4.7	15.1	4.7	9.3	6	20.9	12	26
4.7	11.9	4.7	9	6	24	12	34.3
4.7	17.5	4.7	14.1	6	9.8	12	29.7

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
4.7	19.5	4.7	12.8	6	15.9	12	29.4
4.7	17.2	4.7	13.4	6	22.7	12	23
4.7	17.2	4.7	13.2	6	17	12	18.1
4.7	14.1	4.7	15	6	15.4	12	30.9
4.7	16.4	4.7	9.7	6	16.1	12	29.7
4.7	7	4.7	12.5	6	13.8	12	25.5
4.7	23.9	4.7	6.1	6	10.5	12	20
4.7	18.2	4.7	10.7	6	12.2	12	38
4.7	16.5	4.7	7	6	22.7	12	19
4.7	8.1	4.7	13.2	6	21.7	12	16
4.7	15.6	4.7	11.1	6	14	12	22
4.7	18.1	4.7	12.2	6	18	12	26
4.7	19.1	4.7	13	6	23.7	12	28.7
4.7	4.4	4.7	9	6	19.3	12	30.2
4.7	8.8	4.7	12.5	6	13.6	12	29
4.7	16.3	4.7	13.2	6	21.5	12	18
4.7	19	4.7	15.1	6	16.5	12	24.9
4.7	19.9	4.7	8.7	6	22	12	25
4.7	17.8	4.7	10.6	6	19.1	12	27
4.7	18.4	4.7	6.1	6	10.1	12	22.8
4.7	10.1	4.7	11.2	6	18.1	12	27.7
4.7	18.2	4.7	12.3	6	21	12	21.5
4.7	19.1	4.7	11.1	6	18.4	12	19
4.7	16.3	4.7	14.2	6	20.2	12	21.8
4.7	19.3	4.7	13.6	6	22.9	12	17.7
4.7	16	4.7	12.7	6	6.8	12	16.5
4.7	5.8	4.7	13.7	6	18.2	12	31
4.7	14.2	4.7	13.2	6	13.8	12	19.4
4.7	19	4.7	14	6	15	12	19
4.7	12.6	4.7	11.3	6	18.3	12	32
4.7	10.5	4.7	12.2	6	18.6	12	14
4.7	13.6	4.7	15.7	6	22.5	12	22.1
4.7	17.3	4.7	19	6	15.3	12	26.5
4.7	16.1	4.7	6	6	15.5	12	24.3
4.7	5.3	4.7	18.8	6	18.6	12	25
4.7	20.4	4.7	13.4	6	12	12	17
4.7	16.9	4.7	9.7	6	9.6	12	31

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
4.7	5.4	4.7	15.4	6	19.5	12	35
4.7	9.9	4.7	15	6	22.4	12	18.5
4.7	11.3	4.7	10.4	6	17.2	12	30
4.7	12.3	4.7	4.5	6	13.9	12	21.5
4.7	17.8	4.7	17.5	6	12.4	12	19
4.7	18.3	4.7	14.9	6	11.9	12	25
4.7	18.5	4.7	15.1	6	8.5	12	23.6
4.7	10.3	4.7	12	6	11	12	11.4
4.7	9.7	4.7	7.3	6	15.4	12	19.5
4.7	10	4.7	18.2	6	12.4	12	21
4.7	11.2	4.7	16.5	6	13.9	12	24.2
4.7	3.1	4.7	13.6	6	20	12	18.5
4.7	14.7	4.7	9	6	21.2	12	28
4.7	15.4	4.7	14.2	6	17.5	12	23.7
4.7	15.6	4.7	17.8	6	15.2	12	13
4.7	7.3	4.7	16	6	13.1	12	25.2
4.7	12.7	4.7	9.6	6	9.8	12	22.5
4.7	12.6	4.7	18.3	6	19.3	12	26
4.7	7.8	4.7	9.8	6	21.2	12	23
4.7	12.3	4.7	16.2	6	20	12	16.5
4.7	9.5	4.7	11.4	6	15.2	12	20.5
4.7	4.4	4.7	13	6	13.1	12	22.7
4.7	14.3	4.7	8.8	6	12.9	12	14.5
4.7	18	4.7	13.4	6	13.1	12	22.7
4.7	8.9	4.7	19.8	6	14.2	12	31
4.7	12.4	4.7	12.2	6	14.8	12	23
4.7	7.2	4.7	5.6	6	24.5	12	24.7
4.7	18.8	4.7	17.2	6	24.8	12	16.7
4.7	13.6	4.7	13.2	6	22	12	22.7
4.7	8.4	4.7	13.7	6	16.3	12	21.7
4.7	9.1	4.7	17.3	6	20.2	12	29.1
4.7	15.9	4.7	13.2	6	26.3	12	12.6
4.7	15.2	4.7	12.9	6	17	12	28
4.7	18.1	4.7	10.9	6	21.1	12	27.7
4.7	18.8	4.7	18.2	6	16.8	12	13.2
4.7	6.7	4.7	16.2	6	21.5	12	15.1
4.7	14.1	4.7	13	6	18.6	12	21.4

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
4.7	14.6	4.7	20	6	18.3	12	26
4.7	16.1	4.7	13.6	6	26.8	12	20
4.7	15.4	4.7	13.8	6	19.6	12	12
4.7	15	4.7	11.9	6	20.6	12	20
4.7	7.6	4.7	17.4	6	14.7	12	21
4.7	19.9	4.7	14.2	6	4	12	26
4.7	15.2	4.7	3.5	6	13.9	12	25
4.7	15.5	4.7	3.9	6	9.6	12	19.1
4.7	13.1	4.7	17.1	6	13.2	12	18.2
4.7	14.3	4.7	11.4	6	14.5	12	36
4.7	10.5	4.7	13.2	6	14.4	12	22.1
4.7	14	4.7	17.8	6	17.6	12	13
4.7	11.3	4.7	13.6	6	8.2	12	17.4
4.7	18.9	4.7	8.8	6	17.1	12	24.2
4.7	6.1	4.7	13.9	6	11.7	12	22.6
4.7	21.2	4.7	14.5	6	16.4	12	9.3
4.7	26.5	4.7	18.8	6	16.4	12	27.5
4.7	12.9	4.7	16.6	6	11.3	12	26
4.7	14.2	4.7	20	6	19.1	12	13.6
4.7	18.6	4.7	20.1	6	13.7	12	18.1
4.7	19.1	4.7	4.9	6	15.7	12	34
4.7	15.8	4.7	6.8	6	8.6	12	18.8
4.7	9.7	4.7	21.4	6	17.9	12	14
4.7	24.3	4.7	18.4	6	17.9	12	19.3
4.7	10	4.7	14	6	8.4	12	22.3
4.7	15.7	4.7	15.4	6	12.7	12	16.2
4.7	21.9	4.7	19.3	6	16.6	12	24.7
4.7	6.7	4.7	20.8	6	14.6	12	13.8
4.7	15.9	4.7	12	6	13.5	12	22
4.7	4.1	4.7	18.7	6	13.3	12	15.5
4.7	19.5	4.7	11.1	6	17.8	12	14.3
4.7	19.6	4.7	4.8	6	13.9	12	19.6
4.7	17.3	4.7	18.8	6	13.2	12	23
4.7	3.4	4.7	19.4	6	9.5	12	22
4.7	16	4.7	16.9	6	12.5	12	17.7
4.7	14.3	4.7	17.4	6	15.1	12	28.4
4.7	14.7	4.7	13	6	17.5	12	14.8

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
4.7	11	4.7	6.2	6	12.2	12	25.2
4.7	15.8	4.7	4.3	6	7	12	26.5
4.7	12.9	4.7	15	6	19.1	12	24
4.7	21.5	4.7	20.2	6	6.5	12	21.7
4.7	18.3	4.7	6.8	6	9.4	12	28.6
4.7	10.2	4.7	20.8	6	14	12	29.9
4.7	11.5	4.7	18.8	6	13.1	12	21.8
4.7	22.3	4.7	14.5	6	8.1	12	8.4
4.7	16.3	4.7	4.7	6	15.1	12	26
4.7	8.7	4.7	16.5	6	15.6	12	29.6
4.7	22	4.7	13.8	6	9.4	12	17.8
4.7	21	4.7	3.3	6	22.4	12	29.5
4.7	15.3	4.7	19.8	6	21.8	12	22
4.7	9.9	4.7	8.4	6	14.4	12	28.8
4.7	17.7	4.7	11.9	6	18.9	12	9
4.7	13.8	4.7	13.6	6	21.3	12	17.5
4.7	13	4.7	15.5	6	16.3	12	15.6
4.7	2.9	4.7	17	6	11.6	12	20.6
4.7	7.4	4.7	7.2	6	18.7	12	22.1
4.7	19	4.7	11.2	6	15.6	12	34
4.7	11.5	4.7	21.2	6	14.7	12	37
4.7	15.5	4.7	15.9	6	12.2	12	19.4
4.7	16.4	4.7	19	6	16.2	12	22
4.7	8.7	4.7	8.5	6	12.2	12	18.6
4.7	12.9	4.7	3.4	6	4.3	12	15.6
4.7	14.9	4.7	17.6	6	12.5	12	23.7
4.7	10.7	4.7	17.5	6	12.1	12	27
4.7	16.6	4.7	18.2	6	14.5	12	22.5
4.7	10.6	4.7	14.6	6	12.8	12	30
4.7	9.6	4.7	13.7	6	10.7	12	11
4.7	13.9	4.7	19.4	6	9.5	12	22.3
4.7	16	4.7	18.5	6	16.2	12	10.1
4.7	15.4	4.7	11.3	6	15.6	12	25.5
4.7	14	4.7	9.1	6	12.3	12	23
4.7	5.4	4.7	12.8	6	16	12	14
4.7	15.4	4.7	13.9	6	13.2	12	23.7
4.7	19.1	4.7	18.9	6	18	12	9

Datos utilizados en la Sección Aplicación

Edad	Diámetro	Edad	Diámetro	Edad	Diámetro	Edad	Diámetro
4.7	9.5	4.7	15	6	17.8	12	16
4.7	15.5	4.7	7.7	6	10.5	12	23
4.7	15.3	4.7	16.9	6	14.7	12	23
4.7	14.9	4.7	17.6	6	9.5	12	26
4.7	14.9	4.7	9.5	6	14.6	12	19
4.7	8.6	4.7	6.9	6	9.2	12	37
4.7	10.6	4.7	12	6	11	12	16.4
4.7	16.1	4.7	20.4	6	14.6	12	28.4
4.7	4.3	4.7	15.5	6	12.8	12	13.5
4.7	11.2	4.7	4	6	20.9	12	27
4.7	9.7	4.7	16.4	6	13.8	12	16
4.7	15.9	4.7	2.9	6	15.2	12	26
4.7	14.2	4.7	3.1	6	9.5	12	28.4
4.7	10.7	4.7	10	6	14	12	35.5
4.7	13.2	4.7	13.5	6	11.8	12	18.6
4.7	14.9	4.7	8.6	6	9.7	12	39.7
4.7	9.1	4.7	18.3	6	12.8	12	15.6
4.7	18.7	4.7	15.8	6	11		
4.7	18.3	4.7	14.5	6	13.6		
4.7	6.4	4.7	8.2	6	15.9		
