



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

**PROGRAMA DE POSTGRADO EN SOCIOECONOMÍA, ESTADÍSTICA E
INFORMÁTICA**

ESTADÍSTICA

**UNA METODOLOGÍA PARA EL
ANÁLISIS ESTADÍSTICO DE
RENDIMIENTOS ECONÓMICOS**

AURORA MONTER POZOS

T E S I S

**PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE :**

MAESTRA EN CIENCIAS

MONTECILLO, TEXCOCO, ESTADO DE MÉXICO

2020

La presente tesis titulada: Una metodología para el análisis estadístico de rendimientos económicos realizada por la alumna: Aurora Monter Pozos bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRA EN CIENCIAS
SOCIOECONOMÍA ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO (A) _____
DRA. ELIZABETH GONZÁLEZ ESTRADA

ASESOR (A) _____
DR. JOSÉ A. VILLASEÑOR ALVA

ASESOR (A) _____
DRA. CECILIA RAMÍREZ FIGUEROA

Montecillo, Texcoco, Estado de México, noviembre de 2020

UNA METODOLOGÍA PARA EL ANÁLISIS ESTADÍSTICO DE RENDIMIENTOS ECONÓMICOS

Aurora Monter Pozos, Mtra.
Colegio de Postgraduados, 2020.

RESUMEN

Existe bastante interés en el estudio de variables económicas que sean útiles para realizar estimaciones financieras con la menor incertidumbre posible. El principal problema es encontrar una distribución que provea un buen ajuste, ya que la mayoría de estudios se basan en distribuciones simétricas y específicamente en la distribución normal, pero en general se ha comprobado que la distribución de variables como los rendimientos económicos presentan cierto grado de asimetría. Este trabajo presenta una metodología para el análisis de rendimientos económicos basada en la distribución t-asimétrica. Se proponen dos pruebas de bondad de ajuste para probar dicha distribución y se realiza un estudio de simulación Monte-Carlo basado en el método bootstrap para analizar los criterios de validez de las pruebas propuestas.

Los resultados indican que las distribuciones logística y t-asimétrica son modelos probabilísticos que tienen un buen ajuste para los rendimientos del IPC, los rendimientos de los bonos a 10 años del gobierno mexicano y los rendimientos de las acciones de Wal-Mart.

Palabras clave: rendimientos económicos; distribución t-asimétrica; prueba de Anderson-Darling; método Monte-Carlo; bootstrap paramétrico.

A METHODOLOGY FOR THE STATISTICAL ANALYSIS OF ECONOMIC RETURNS

Aurora Monter Pozos, Mtra.

Colegio de Postgraduados, 2020.

ABSTRACT

There exists considerable interest in the study of economic variables that are useful to conduct financial estimates with the least possible uncertainty. The main issue is to find a distribution that provides a good fit, since most studies are based on symmetric distributions and specifically on the normal distribution, but in general it has been proven that the distribution of variables such as economic returns present a certain degree of asymmetry.

This work introduces a methodology for the analysis of economic returns based on the skew-t distribution. Two goodness-of-fit tests are proposed to test this distribution and a Monte-Carlo simulation study is made based on the parametric bootstrap method to analyze the validity criteria of the proposed tests.

The results indicate that the logistic and skew-t distributions are probabilistic models that have a good fit for the returns of the IPC, the returns of the 10-Year Bond of the Mexican government and the returns of Wal-Mart stocks.

Keywords: economic returns; skew-t distribution; Anderson-Darling test; Monte-Carlo method; parametric bootstrap.

El presente trabajo está dedicado a todas aquellas personas que han sido parte fundamental de mi formación humana y académica.

Al Colegio de Postgraduados y al Consejo Nacional de Ciencia y Tecnología (CONACYT) les agradezco por el apoyo financiero y los recursos que me brindaron para terminar con éxito mis estudios. De igual manera, agradezco al personal del Postgrado en Estadística. Fue un honor formar parte de ésta gran familia.

A MIS PADRES

Por su incanzable labor que día a día han realizado para hacer de mí una persona con principios y valores bien fundamentados, les agradezco por confiar en mí ya que siempre se mantuvieron firmes ante cualquier dificultad. Gracias por enseñarme a luchar por mis objetivos y principalmente a nunca desistir por más difícil que parezca el camino, todo lo que soy es gracias a ustedes.

A MIS HERMANOS

A pesar de tener diferentes metas e ideales, siempre han sido un pilar fundamental en nuestra familia, gracias por el apoyo incondicional y por ser un gran ejemplo de que con esfuerzo y dedicación todo se puede lograr.

A MIS ASESORES DE TESIS

Durante la realización de este trabajo ustedes fueron mi mano derecha; su conocimiento y experiencia han sido un aporte invaluable, no solamente en el desarrollo de esta tesis, sino también en mi formación como estudiante.

AL AMOR DE MI VIDA

Tu presencia ha sido una pieza fundamental en mi vida, gracias por creer en mí y sobre todo por el amor incondicional que me tienes.

CONTENIDO

RESUMEN	iii
ABSTRACT	iv
LISTA DE CUADROS	ix
LISTA DE FIGURAS	xiii
1. INTRODUCCIÓN	1
2. REVISIÓN DE LITERATURA	16
2.1. Distribución Logística	16
2.2. Distribución Laplace	17
2.3. Distribución Laplace asimétrica	19
2.4. Distribución t-Student	20
2.5. Distribución Normal asimétrica	21
2.6. Distribución t-asimétrica	22
2.6.1. Estimación de parámetros	26
2.7. Pruebas de hipótesis	28
2.8. Pruebas de bondad de ajuste	30
2.8.1. Prueba de Anderson-Darling	31
2.8.2. Prueba de Shapiro-Wilk	32
2.9. Método bootstrap paramétrico	32
3. METODOLOGÍA PROPUESTA PARA PROBAR AJUSTE DE LA DISTRIBUCIÓN T-ASIMÉTRICA	34
3.1. Transformación de variables	34
3.2. Prueba de Anderson-Darling para la distribución t-asimétrica	36

3.3.	Prueba de Anderson-Darling con base en datos transformados	37
3.4.	Estudio de Simulación Monte-Carlo	38
4.	APLICACIÓN A LOS RENDIMIENTOS DEL IPC	56
4.1.	Rendimientos semanales	57
4.1.1.	Rendimientos 1995-1996	57
4.1.2.	Rendimientos 2001-2002	62
4.1.3.	Rendimientos 2008-2009	66
4.1.4.	Rendimientos 2019-2020	70
4.2.	Rendimientos mensuales	74
4.2.1.	Rendimientos 1993-1999	74
4.2.2.	Rendimientos 2015-2020	78
4.3.	Rendimientos para períodos largos	81
4.3.1.	Rendimientos mensuales 2000-2020	81
5.	APLICACIÓN A LOS RENDIMIENTOS DE BONOS A 10 AÑOS	85
5.1.	Rendimientos semanales	85
5.1.1.	Rendimientos 2009-2011	86
5.1.2.	Rendimientos 2018-2020	89
6.	OTRAS APLICACIONES	93
6.1.	Rendimientos de las acciones de Wal-Mart México (WALMEX)	93
6.1.1.	Rendimientos diarios 2020	93
6.1.2.	Rendimientos semanales 2019-2020	96
6.1.3.	Rendimientos mensuales 2013-2020	100
7.	DISCUSIÓN DE RESULTADOS Y CONCLUSIONES	104
	ANEXOS	111

LISTA DE CUADROS

2.1. Tipos de errores.	30
3.1. Tamaño de pruebas estimado para diferentes valores de α y $\nu = 3, 5, 8, 12$ con $\alpha^* = 0.05$	40
3.2. Tamaño de pruebas estimado para diferentes valores de α y ν con $\alpha^* = 0.05$	41
3.3. Potencia estimada de las pruebas utilizando $\alpha^* = 0.05$	45
4.1. Parámetros estimados de los rendimientos del IPC 1995-1996 utilizando diferentes distribuciones.	60
4.2. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Enero, 1995 a Diciembre, 1996.	61
4.3. Parámetros estimados de los rendimientos del IPC 2001-2002 utilizando diferentes distribuciones.	64
4.4. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Enero, 2001 a Diciembre, 2002.	65
4.5. Parámetros estimados de los rendimientos del IPC 2008-2009 utilizando diferentes distribuciones.	68
4.6. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.	69
4.7. Parámetros estimados de los rendimientos del IPC 2019-2020 utilizando diferentes distribuciones.	72
4.8. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Abril, 2019 a Agosto, 2020.	73
4.9. Parámetros estimados de los rendimientos mensuales del IPC 1993-1999 utilizando diferentes distribuciones.	76
4.10. Resultados de las diferentes pruebas aplicadas a los rendimientos mensuales del IPC de Julio, 1993 a Enero, 1999.	77

4.11. Parámetros estimados de los rendimientos mensuales del IPC 2015-2020 utilizando diferentes distribuciones.	78
4.12. Resultados de las diferentes pruebas aplicadas a los rendimientos mensuales del IPC de Enero, 2015 a Agosto, 2020.	81
4.13. Parámetros estimados de los rendimientos mensuales del IPC 2000-2020 utilizando diferentes distribuciones.	83
4.14. Resultados de las diferentes pruebas aplicadas a los rendimientos mensuales del IPC de Enero, 2000 a Agosto, 2020.	84
5.1. Parámetros estimados de los rendimientos semanales de los bonos a 10 años 2009-2011 utilizando diferentes distribuciones.	86
5.2. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de los bonos a 10 años de Abril, 2009 a Abril, 2011.	89
5.3. Parámetros estimados de los rendimientos semanales de los bonos a 10 años 2018-2020 utilizando diferentes distribuciones.	91
5.4. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de los bonos a 10 años de Diciembre, 2018 a Septiembre, 2020.	92
6.1. Parámetros estimados de los rendimientos diarios de Wal-Mart de Marzo a Agosto, 2020 utilizando diferentes distribuciones.	95
6.2. Resultados de las diferentes pruebas aplicadas a los rendimientos diarios de Wal-Mart de Marzo a Agosto, 2020.	96
6.3. Parámetros estimados de los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020 utilizando diferentes distribuciones.	97
6.4. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020.	100
6.5. Parámetros estimados de los rendimientos mensuales de Wal-Mart de Septiembre, 2013 a Septiembre, 2020 utilizando diferentes distribuciones.	102
6.6. Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020.	103

LISTA DE FIGURAS

1.1. Histogramas de los rendimientos del IPC 2008-2009.	6
1.2. Gráficas de dispersión de los rendimientos del IPC 2008-2009.	6
1.3. Histogramas de los rendimientos del IPC 2018-2020.	7
1.4. Gráficas de dispersión de los rendimientos del IPC 2018-2020.	7
1.5. Histogramas de los rendimientos de los bonos a 10 años 2009-2010.	8
1.6. Gráficas de dispersión de los rendimientos de los bonos a 10 años 2009-2010.	8
1.7. Histogramas de los rendimientos de Wal-Mart 2008-2010.	9
1.8. Gráficas de dispersión de los rendimientos de Wal-Mart 2008-2010.	9
1.9. Gráficas en el tiempo de los rendimientos semanales.	10
1.10. Gráficas en el tiempo de los rendimientos semanales.	11
1.11. Curvas ajustadas para los rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.	12
1.12. Curvas ajustadas para los rendimientos semanales del IPC de Diciembre, 2018 a Mayo, 2020.	13
1.13. Curvas ajustadas para los rendimientos semanales de los bonos a 10 años de Abril, 2009 a Diciembre, 2010.	14
1.14. Curvas ajustadas para los rendimientos semanales de Wal-Mart de Abril, 2008 a Marzo, 2010.	15
2.1. Densidad logística con parámetros μ y σ	17
2.2. Densidad Laplace con parámetros μ y b	18
2.3. Densidad Laplace asimétrica con parámetros μ , s y λ	19
2.4. Densidad t-Student con ν grados de libertad.	21
2.5. Densidad normal asimétrica para diferentes valores de α	22
2.6. Densidad t-asimétrica con $\xi = 0$, $\omega = 1$ y diferentes valores de los parámetros α y ν	25

2.7. Rango de γ_1 y γ_2 y curvas asociadas para diferentes valores de los grados de libertad con α en la recta real.	26
3.1. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=100, parte I.	46
3.2. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=100, parte II.	47
3.3. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=100, parte III.	48
3.4. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=100, parte IV.	49
3.5. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con n=100, parte I.	50
3.6. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con n=100, parte II.	51
3.7. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con n=100, parte III.	52
3.8. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con n=100, parte IV.	53
4.1. Rendimientos semanales del IPC de Enero, 1995 a Diciembre, 1996.	59
4.2. Curvas ajustadas para los rendimientos semanales del IPC de Enero, 1995 a Agosto, 1996.	60
4.3. Rendimientos semanales del IPC de Enero, 2001 a Diciembre, 2002.	63
4.4. Curvas ajustadas para los rendimientos semanales del IPC de Enero, 2001 a Diciembre, 2002.	64
4.5. Rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.	67
4.6. Curvas ajustadas para los rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.	68
4.7. Rendimientos semanales del IPC de Abril, 2019 a Agosto, 2020.	71
4.8. Curvas ajustadas para los rendimientos semanales del IPC de Abril, 2019 a Agosto, 2020.	73
4.9. Rendimientos mensuales del IPC de Julio, 1993 a Enero, 1999.	75
4.10. Curvas ajustadas para los rendimientos mensuales del IPC de Julio, 1993 a Enero, 1999.	76

4.11. Rendimientos mensuales del IPC de Enero, 2015 a Agosto, 2020.	79
4.12. Curvas ajustadas para los rendimientos mensuales del IPC de Enero, 2015 a Agosto, 2020.	80
4.13. Rendimientos mensuales del IPC de Enero, 2000 a Agosto, 2020.	82
4.14. Curvas ajustadas para los rendimientos mensuales del IPC de Enero, 2000 a Agosto, 2020.	83
5.1. Rendimientos semanales de los bonos a 10 años de Abril, 2009 a Abril, 2011. .	87
5.2. Curvas ajustadas para los rendimientos semanales de los bonos a 10 años de Abril, 2009 a Abril, 2011.	88
5.3. Rendimientos semanales de los bonos a 10 años de Diciembre, 2018 a Septiembre, 2020.	90
5.4. Curvas ajustadas para los rendimientos semanales de los bonos a 10 años de Diciembre, 2018 a Septiembre, 2020.	91
6.1. Rendimientos diarios de las acciones de Wal-Mart de Marzo a Agosto, 2020. .	94
6.2. Curvas ajustadas para los rendimientos diarios de Wal-Mart de Marzo a Agosto, 2020.	95
6.3. Rendimientos semanales de las acciones de Wal-Mart de Marzo, 2019 a Septiembre, 2020.	98
6.4. Curvas ajustadas para los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020.	99
6.5. Rendimientos semanales de las acciones de Wal-Mart de Marzo, 2019 a Septiembre, 2020.	101
6.6. Curvas ajustadas para los rendimientos mensuales de Wal-Mart de Septiembre, 2013 a Septiembre, 2020.	102
7.1. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=500, parte I.	115
7.2. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=500, parte II.	116
7.3. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=500, parte III.	117
7.4. FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con n=500, parte IV.	118

7.5. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte I.	119
7.6. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte II.	120
7.7. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte III.	121
7.8. FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte IV.	122

CAPÍTULO 1

INTRODUCCIÓN

En inferencia estadística paramétrica se parte de la suposición de que los datos constituyen una muestra aleatoria proveniente de una población que tiene una distribución de probabilidades dada, la cual depende de parámetros desconocidos. Las inferencias paramétricas basadas en la muestra aleatoria son válidas sólo si la distribución de probabilidad en la que se basan realmente explica el comportamiento probabilístico de los datos.

La teoría financiera se basa fuertemente en el supuesto de normalidad; sin embargo, existe evidencia empírica la cual indica que variables aleatorias tales como tasas de crecimiento económico, rendimientos de acciones e índices de los distintos mercados de valores del mundo, entre otras, no tienen una distribución normal ([21]), ([37]). Histogramas de frecuencias construidos con base en observaciones de estas variables, indican que las distribuciones de probabilidades de estas cantidades son asimétricas y/o con colas más pesadas que las de la distribución normal. Dichas cantidades son llamadas retornos o rendimientos y se definen como la diferencia del precio final menos el precio inicial dividida entre el precio inicial.

En el área de economía y finanzas se tiene interés en identificar la distribución de probabilidades correcta de los rendimientos económicos ya que en dicha área es de suma importancia hacer predicciones lo más precisas posible. Por lo anterior, diversos autores han tratado de identificar modelos de probabilidad válidos para los rendimientos de los grandes mercados de valores de Europa y Estados Unidos ([21]), ([37]). En esta investigación, se tiene interés en identificar y/o proponer distribuciones de probabilidad, que permitan explicar el comportamiento probabilístico de pérdidas y ganancias de acciones que cotizan en el mercado de valores mexicano.

Definición 1 El S&P/BMV IPC (Índice De Precios y Cotizaciones) pretende medir el rendimiento de las acciones de mayor liquidez listadas en la Bolsa Mexicana de Valores. Su objetivo es proporcionar un índice amplio, representativo, pero al mismo tiempo fácilmente replicable,

que abarque el mercado bursátil mexicano.

El estudio de dicho índice es de interés principalmente en períodos con variaciones económicas importantes, como lo son períodos de nuevos gobiernos, de crisis y sucesos que han impactado la economía mexicana.

La variable de interés es el rendimiento económico, el cual es la ganancia que permite obtener una cierta operación, luego de un cierto período, éste se obtiene de la siguiente forma:

$$r = \frac{\text{precio final} - \text{precio inicial}}{\text{precio inicial}}. \quad (1.1)$$

Los registros del IPC ¹ están medidos con frecuencia diaria, semanal y mensual y están conformados por 7 características: fecha, cierre, apertura, máximo, mínimo, volumen y porcentaje de variación. En este caso sólo se hará uso de la fecha, cierre y apertura, a partir de estas dos últimas se genera la variable rendimientos.

En la Figura 1.1 se observa que la distribución de los rendimientos diarios, semanales y mensuales durante el período del 02/01/2008 al 31/12/2009 presenta una ligera asimetría. Para los rendimientos diarios y semanales dicha asimetría se ve reflejada en la cola derecha del histograma. Además para estos casos se tiene una mayor frecuencia para rendimientos en las clases cercanas al cero teniendo así una distribución con picos altos.

En la Figura 1.2 se puede observar que los rendimientos pertenecientes al período estudiado tuvieron una gran variabilidad. Para el caso de los rendimientos diarios los valores de rendimientos más bajos corresponden a los registros de septiembre a diciembre de 2008. Los rendimientos semanales corresponden a los meses de septiembre y octubre de 2008, y para los rendimientos mensuales se tiene que el valor más bajo se tuvo en el mes de octubre de 2008. Estos períodos en los cuales se tuvo una mayor variabilidad corresponden a los meses en los que la crisis económica del 2008 tuvo un mayor impacto en México.

Es de suma importancia verificar el supuesto de independencia en las observaciones con el fin de poder realizar inferencias acerca de la posible distribución de los rendimientos.

En este caso se utiliza la prueba de Box-Pierce ([6]), ([29]), la cual tiene sus aplicaciones en

¹Fuente:Investing.com

análisis de series temporales y toma como hipótesis nula que las observaciones son independientes en el tiempo. Los resultados se obtuvieron utilizando la función `Box.test` del paquete estadístico R ([42]).

Los valores de probabilidad obtenidos para los rendimientos diarios, semanales y mensuales aplicando la prueba anterior fueron 0.03, 0.07, 0.26 respectivamente, por lo cual a un nivel de significancia de 0.05 se rechaza la hipótesis de independencia para los rendimientos diarios, mientras que para los otros dos casos no se rechaza la hipótesis nula.

Análogamente se analizan los rendimientos diarios, semanales y mensuales durante el período del 01/12/2018 al 10/05/2020 que comprende el inicio del gobierno de Andrés Manuel López Obrador.

La Figura 1.3 muestra que el histograma de frecuencias de los rendimientos diarios presenta una ligera asimetría en la cola izquierda. En la Figura 1.4 se observa que en varias ocasiones los rendimientos tomaron valores negativos, mayormente en los meses de febrero y marzo de 2020 los cuales corresponden al inicio de la pandemia por el COVID-19.

Los valores de probabilidad obtenidos al aplicar la prueba de Box-Pierce para los rendimientos diarios, semanales y mensuales respectivamente fueron 0.03, 0.09, 0.68. A un nivel de significancia de 0.05 se rechaza la hipótesis de independencia para las observaciones pertenecientes a los rendimientos diarios, mientras que para los rendimientos semanales y mensuales no se rechaza, por tanto se pueden considerar independientes en el tiempo.

De forma análoga se realiza un análisis exploratorio considerando los rendimientos diarios, semanales y mensuales de los bonos a 10 años de México del 05/04/2009 al 26/12/2010 y los rendimientos de las acciones de Wal-Mart México (WALMEX) del 04/04/2008 al 07/03/2010. Las Figuras 1.5 y 1.7 muestran los histogramas de frecuencias correspondientes.

Los valores de probabilidad obtenidos al aplicar la prueba de Box-Pierce a los rendimientos diarios, semanales y mensuales de los bonos a 10 años fueron 0.3002, 0.8228 y 0.6006, respectivamente, mientras que los valores de probabilidad obtenidos para los rendimientos diarios, semanales y mensuales de las acciones de Wal-Mart fueron 0.6652, 0.4929 y 0.7424, respectivamente.

En el siguiente análisis se consideran los rendimientos semanales para cada conjunto de datos, debido a que en estos casos se puede considerar que las observaciones son independientes. Se

ajustaron diferentes distribuciones a los rendimientos: la Normal, logística, Cauchy, Laplace, Normal asimétrica, Laplace asimétrica, t-Student y t-asimétrica. Esto con el fin de comparar la distribución empírica de los datos con las distribuciones ajustadas (estimando los respectivos parámetros).

En la Figura 1.11 se observa que las curvas ajustadas con las distribuciones Normal y Normal asimétrica se alejan de la curva empírica en gran parte del centro de la distribución. Con la curva ajustada de la distribución Cauchy, se observa una gran falta de ajuste en las colas de la distribución empírica mientras que las curvas ajustadas con las distribuciones logística y Laplace, tanto simétrica como asimétrica, tienen un buen ajuste en gran parte de la distribución empírica. La distribución t-Student, en su forma simétrica se ajusta de forma adecuada. Debido a la poca asimetría del conjunto de datos se nota el mismo ajuste con la distribución t-asimétrica.

Por otro lado, en la Figura 1.12 se observa que la curva normal tanto en su forma simétrica como asimétrica tiene un mal ajuste en gran parte del centro de la distribución empírica. Además, lo mismo sucede con las curvas ajustadas con la distribución Laplace simétrica y asimétrica. Por otro lado, la distribución Cauchy tiene un mejor ajuste en la parte del centro de la distribución, pero en las colas el ajuste empeora considerablemente. Note que en este caso la distribución logística presenta un buen ajuste y nuevamente en el caso de la distribución t-Student tanto simétrica como asimétrica. El ajuste es muy bueno en el centro y en las colas de la distribución empírica.

Para el caso de los rendimientos semanales de los bonos a 10 años, la Figura 1.13 muestra que el ajuste con las distribuciones normal tanto simétrica como asimétrica tienen un mal ajuste en el centro de la distribución empírica, de forma similar la distribución Cauchy tiene un mal ajuste en parte del centro y las colas de la curva empírica. Por otro lado, las distribuciones logística, Laplace, Laplace asimétrica y t-asimétrica proveen un buen ajuste a los datos.

Para los rendimientos semanales de las acciones de Wal-Mart, la Figura 1.14 muestra que las distribuciones normal, normal asimétrica, Cauchy y Laplace tienen un mal ajuste en gran parte de la curva empírica, contrario a lo que sucede con las distribuciones logística, t-Student y t-asimétrica.

Este análisis exploratorio nos lleva a concluir que en efecto la distribución de los rendimientos semanales no es normal, lo cual amplía el panorama en busca de distribuciones candidatas que mejoren el ajuste de manera general considerando las distribuciones Normal asimétrica,

t-asimétrica, logística y Laplace.

Los objetivos principales de esta investigación son:

- 1.- Encontrar una distribución de probabilidad adecuada para modelar probabilísticamente los rendimientos de acciones del mercado de valores mexicano.
- 2.- Proponer una prueba de hipótesis estadística para probar la bondad de ajuste de la distribución de probabilidades obtenida en 1 que tenga propiedades satisfactorias de acuerdo a la teoría de la inferencia estadística.
- 3.- Aplicar la metodología desarrollada en 1 y 2 para los datos históricos, provenientes del mercado de valores de México.

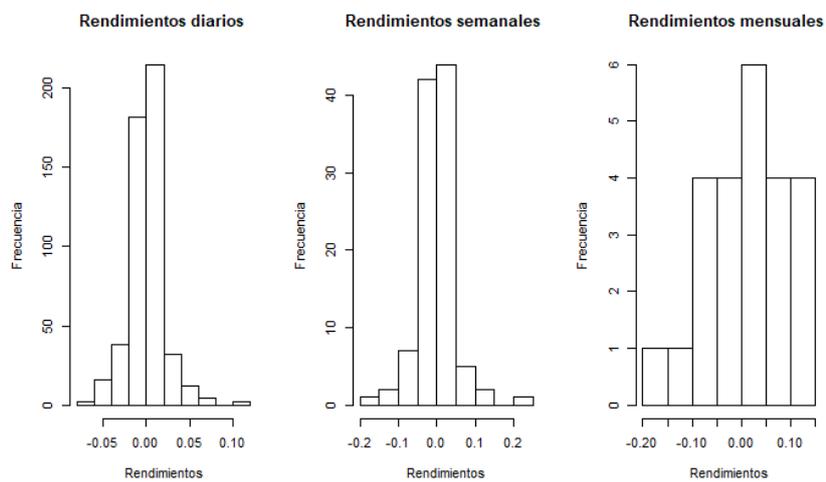


Figura 1.1: Histogramas de los rendimientos del IPC 2008-2009.

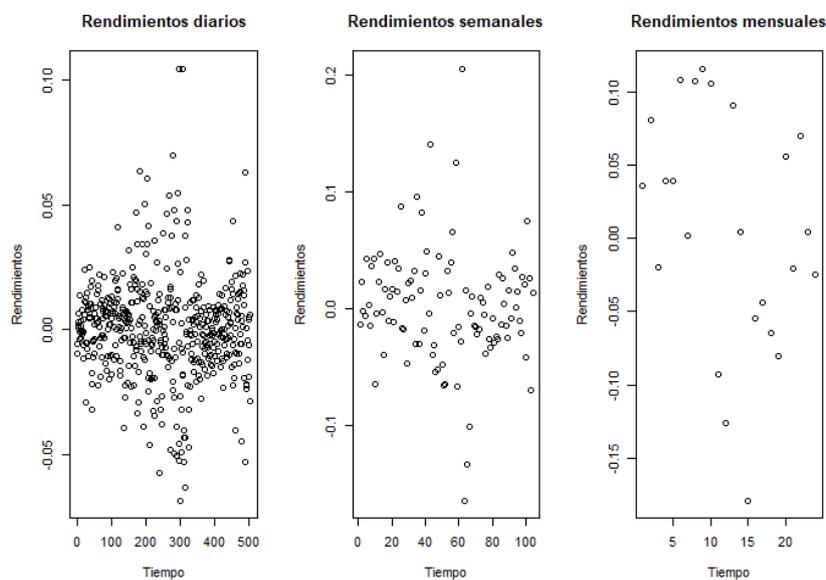


Figura 1.2: Gráficas de dispersión de los rendimientos del IPC 2008-2009.

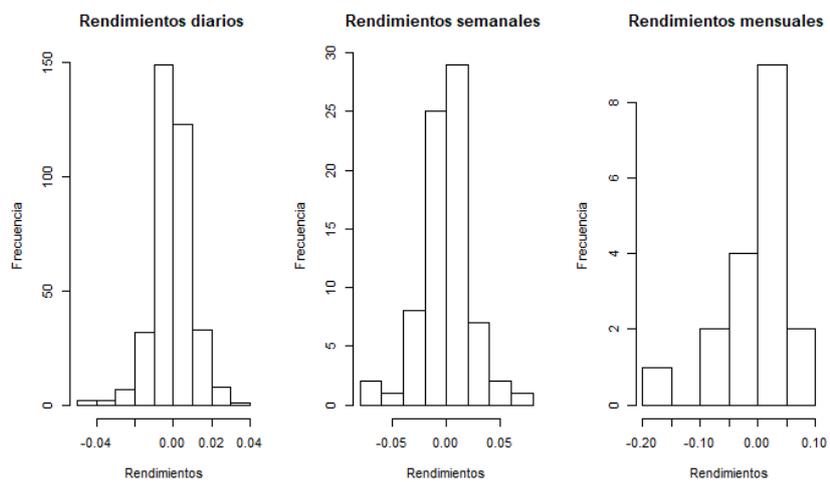


Figura 1.3: Histogramas de los rendimientos del IPC 2018-2020.

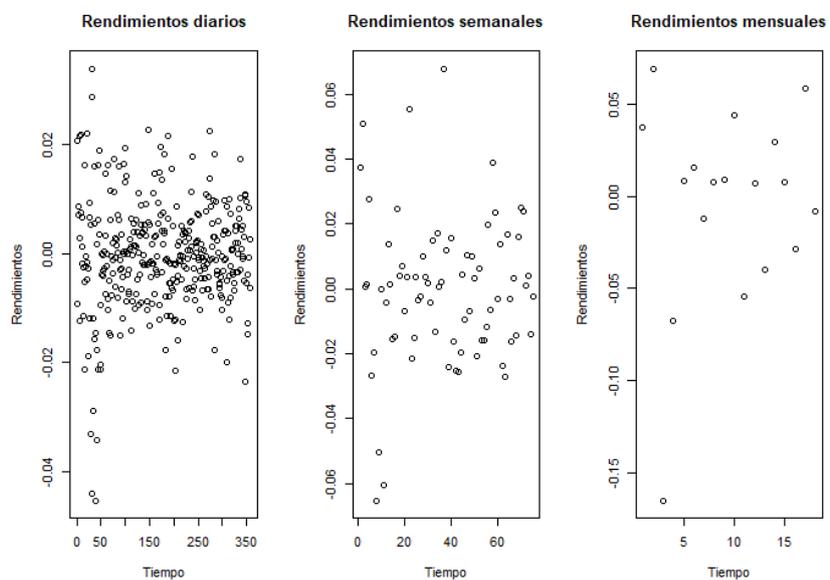


Figura 1.4: Gráficas de dispersión de los rendimientos del IPC 2018-2020.

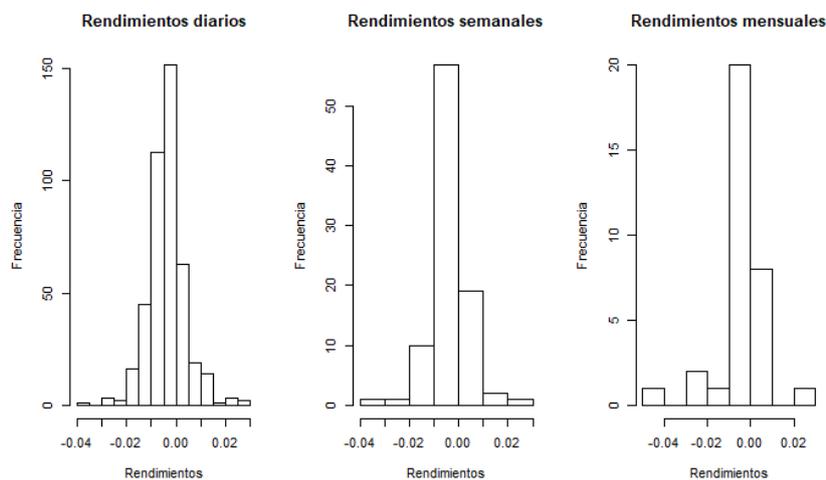


Figura 1.5: Histogramas de los rendimientos de los bonos a 10 años 2009-2010.

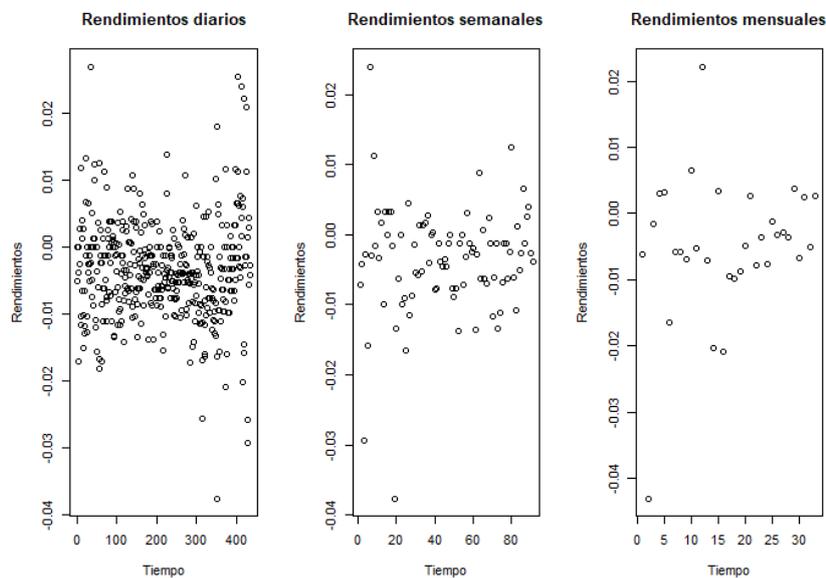


Figura 1.6: Gráficas de dispersión de los rendimientos de los bonos a 10 años 2009-2010.

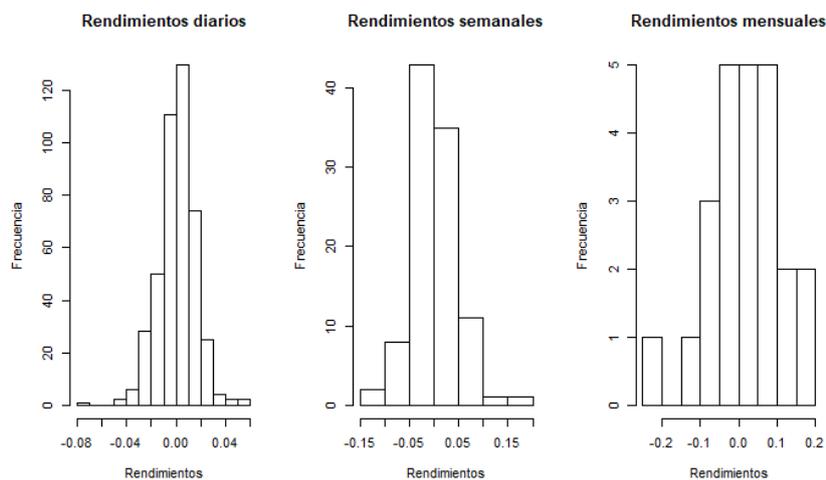


Figura 1.7: Histogramas de los rendimientos de Wal-Mart 2008-2010.

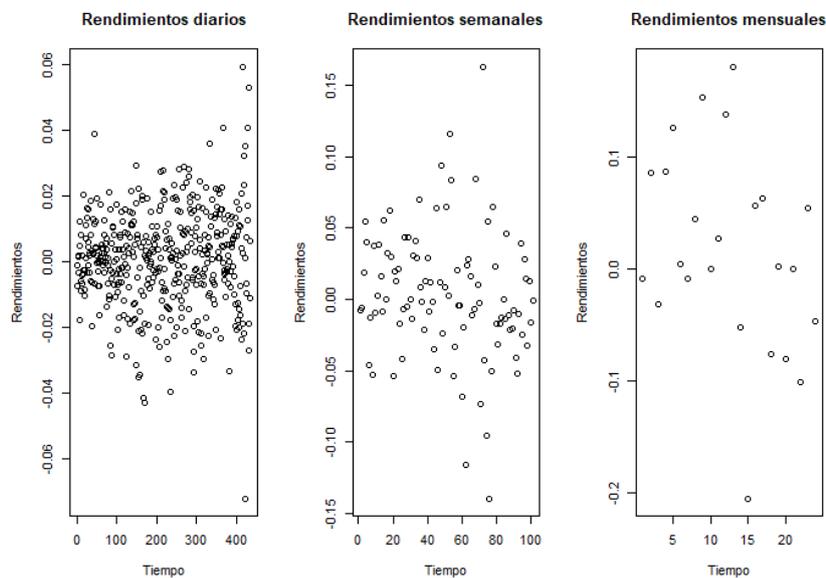
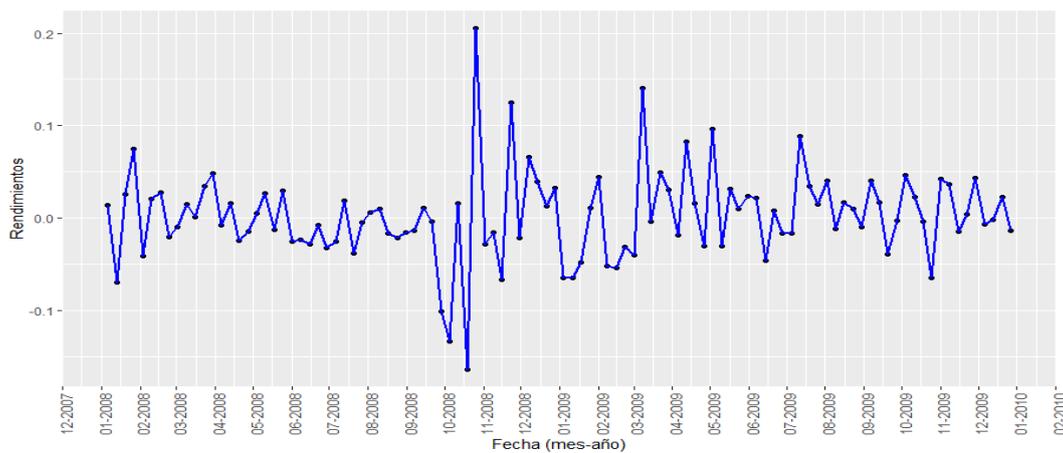
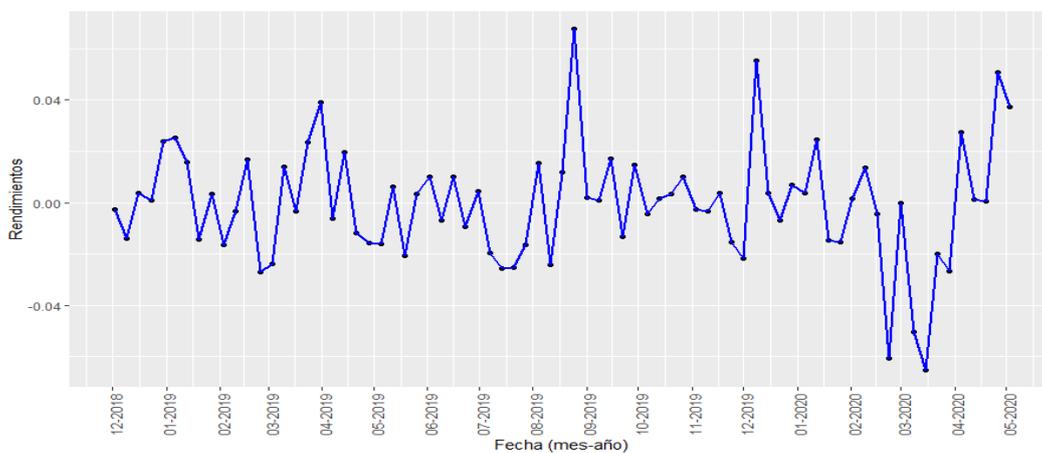


Figura 1.8: Gráficas de dispersión de los rendimientos de Wal-Mart 2008-2010.

Figura 1.9: Gráficas en el tiempo de los rendimientos semanales.

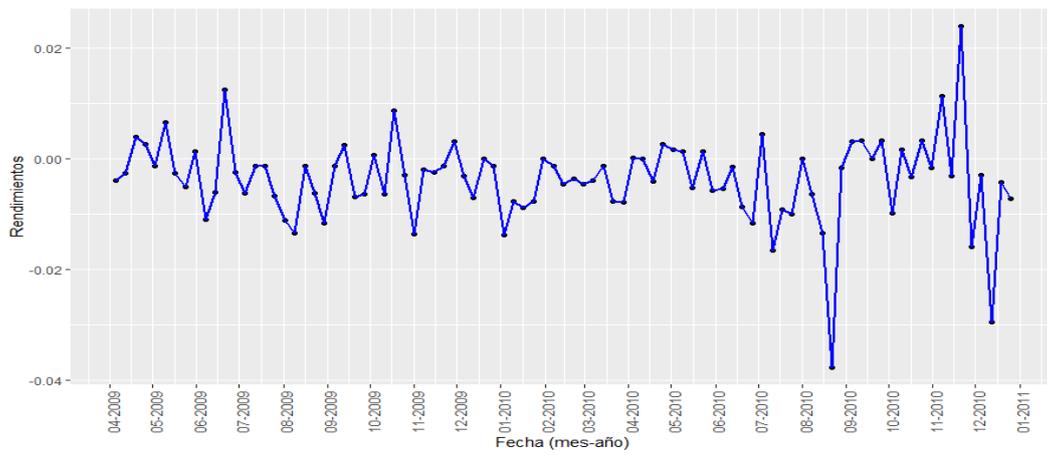


(a) Rendimientos del IPC 2008-2009.

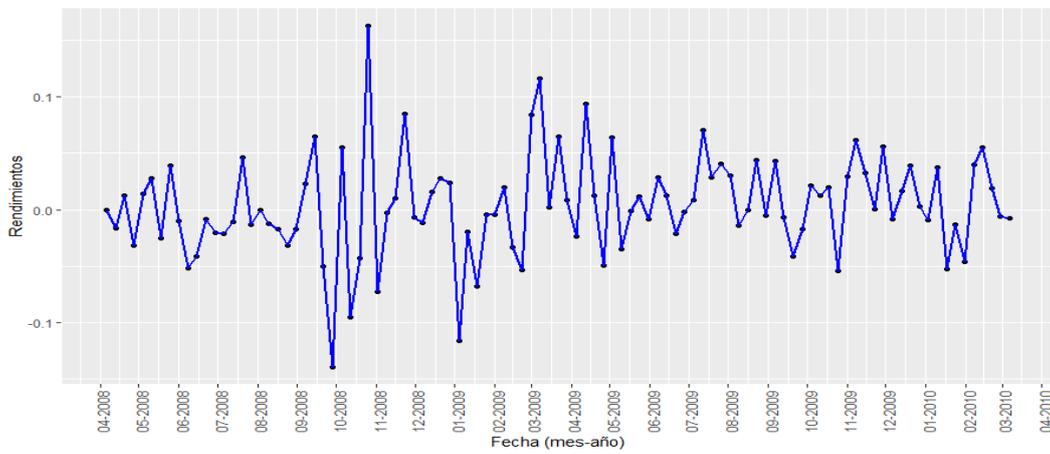


(b) Rendimientos del IPC 2018-2020.

Figura 1.10: Gráficas en el tiempo de los rendimientos semanales.

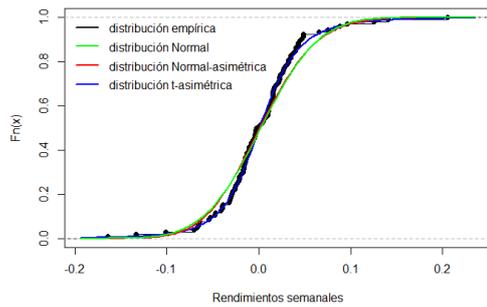


(a) Rendimientos de los bonos a 10 años 2009-2010.

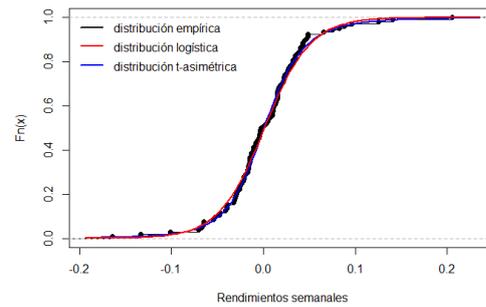


(b) Rendimientos de Wal-Mart 2008-2010.

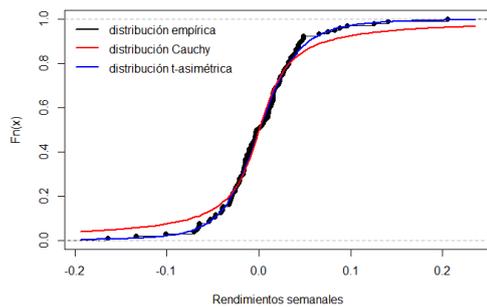
Figura 1.11: Curvas ajustadas para los rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.



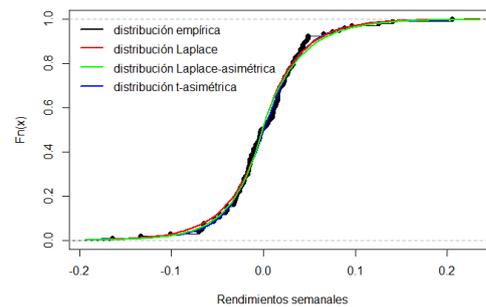
(a)



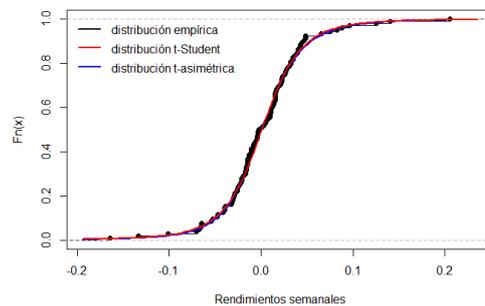
(b)



(c)

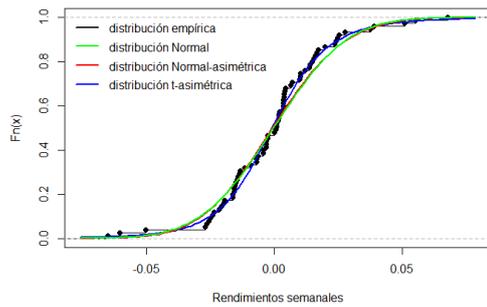


(d)

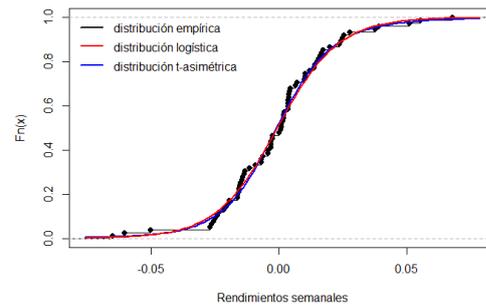


(e)

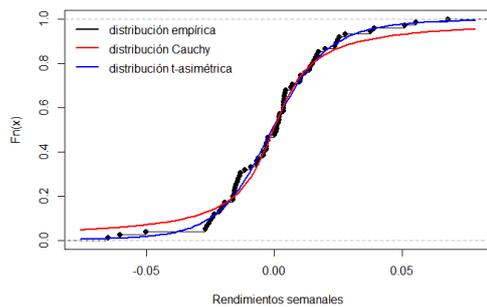
Figura 1.12: Curvas ajustadas para los rendimientos semanales del IPC de Diciembre, 2018 a Mayo, 2020.



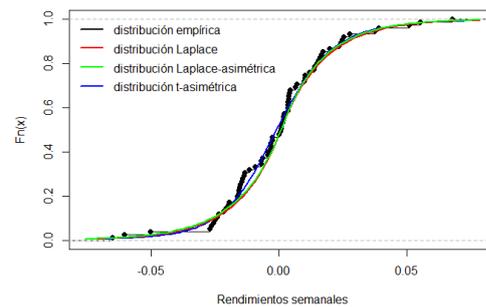
(a)



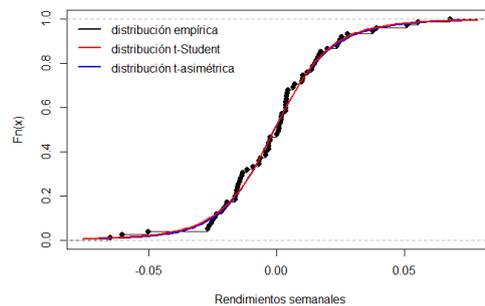
(b)



(c)

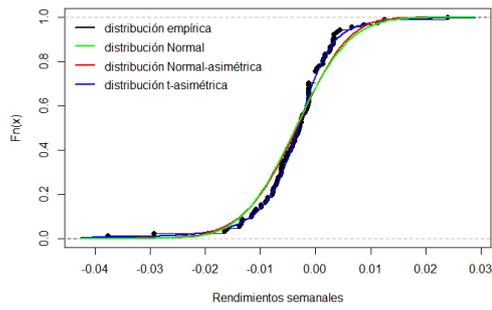


(d)

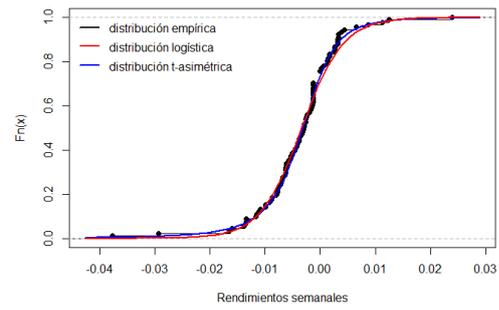


(e)

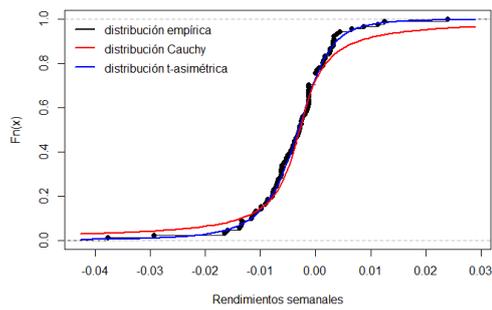
Figura 1.13: Curvas ajustadas para los rendimientos semanales de los bonos a 10 años de Abril, 2009 a Diciembre, 2010.



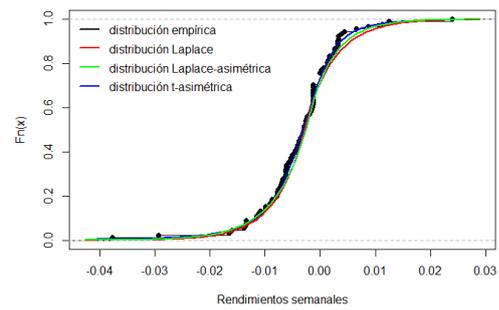
(a)



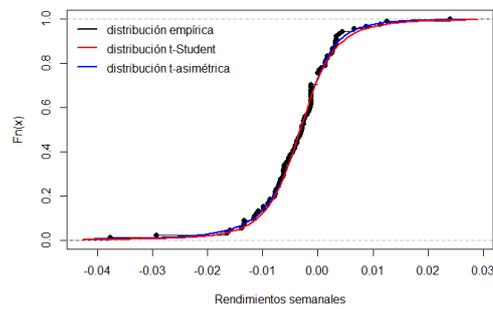
(b)



(c)

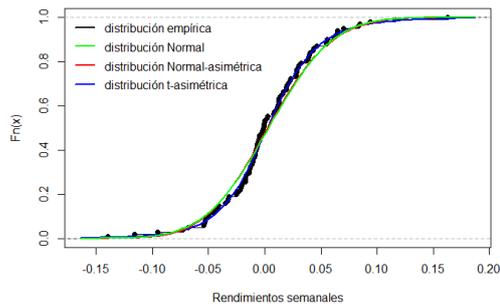


(d)

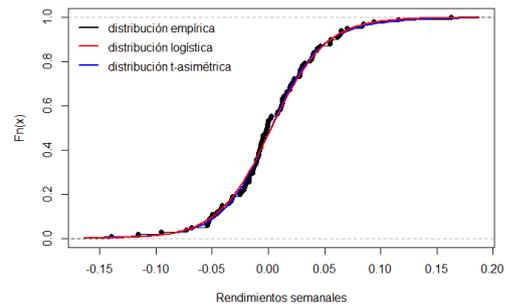


(e)

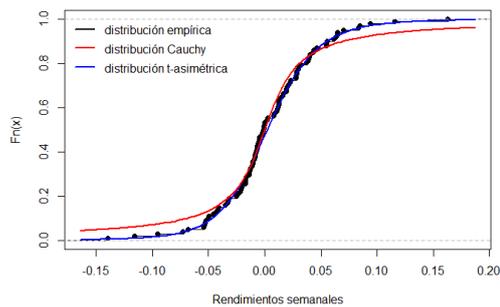
Figura 1.14: Curvas ajustadas para los rendimientos semanales de Wal-Mart de Abril, 2008 a Marzo, 2010.



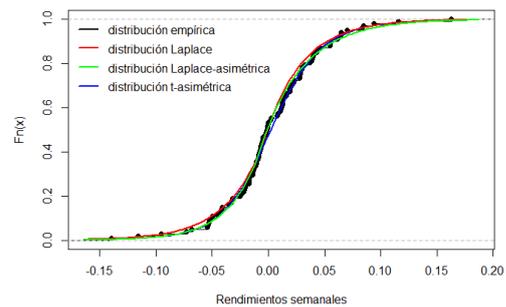
(a)



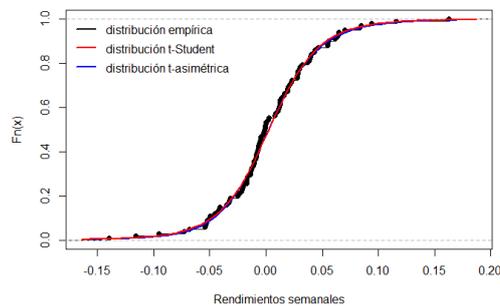
(b)



(c)



(d)



(e)

CAPÍTULO 2

REVISIÓN DE LITERATURA

En este capítulo, se estudian algunos conceptos fundamentales que son de suma importancia en el desarrollo de la teoría estadística como son las pruebas de hipótesis, los estimadores de máxima verosimilitud, la función score y las pruebas de bondad de ajuste entre otros. También se revisan las principales distribuciones en las que se basa el estudio (resultantes del análisis exploratorio que se realizó previamente), la Logística, la Laplace, la t-Student, la Normal asimétrica y la t-asimétrica, Las dos últimas, vistas como extensiones de distribuciones ya conocidas, han sido estudiadas recientemente por presentar diversas características que las hacen más flexibles y útiles en el ámbito de las aplicaciones.

2.1. Distribución Logística

La distribución logística es una distribución continua, se define por sus parámetros de localidad y escala y no tiene parámetro de forma, lo que significa que la función de densidad de probabilidad solo tiene una forma, la curva de la distribución logística es similar a la de la distribución normal, sin embargo, la distribución logística tiene colas más largas. Esta distribución ha sido usada en áreas como la biología, para describir cómo se comportan las especies en entornos competitivos ([43]); en medicina, para el análisis y clasificación de datos médicos ([10]) y en finanzas, para el desarrollo de modelos económicos ([26]).

La función de densidad de probabilidad de una variable aleatoria X con distribución logística con parámetro de localidad μ y de escala σ está dada por:

$$f(x; \mu, \sigma) = \frac{e^{-(x-\mu)/\sigma}}{\sigma(1 + e^{-(x-\mu)/\sigma})^2}, \quad x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0. \quad (2.1)$$

Su función de distribución acumulada es:

$$F(x; \mu, \sigma) = \frac{1}{1 + e^{-(x-\mu)/\sigma}}, \quad x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0. \quad (2.2)$$

Esta distribución se caracteriza por ser simétrica, es decir, la media, la moda y la mediana de esta densidad tienen el mismo valor y en este caso es el parámetro μ .

La Figura 2.1 muestra algunas curvas de la densidad logística con diferentes valores de los parámetros.

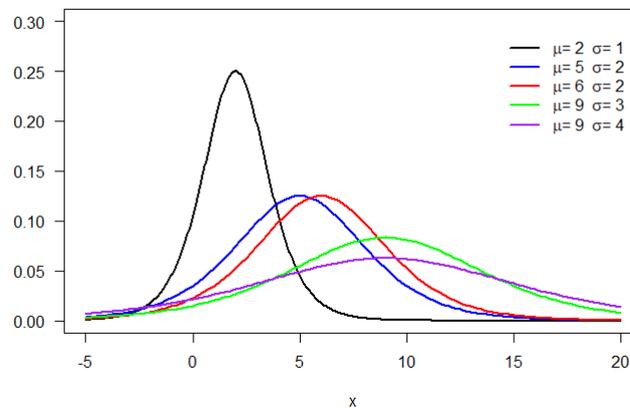


Figura 2.1: Densidad logística con parámetros μ y σ .

2.2. Distribución Laplace

La distribución de Laplace es una densidad de probabilidad continua, llamada así en honor a Pierre-Simon Laplace. Es también conocida como distribución doble exponencial debido a que suele considerarse como dos distribuciones exponenciales, una positiva y una negativa. Esta distribución resulta de la diferencia de dos variables aleatorias exponenciales independientes e idénticamente distribuidas.

En el ámbito aplicado, la distribución Laplace ha sido utilizada en la hidrología para analizar variables aleatorias como los valores máximos de la precipitación, para describir épocas de sequía ([8]) y en general para estudiar datos cuya distribución tenga un pico más alto que la distribución normal.

La función de densidad de probabilidad de una variable X con distribución Laplace con parámetro de localidad μ y de escala b está dada por:

$$f(x; \mu, b) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right), \quad x \in \mathbb{R}, \mu \in \mathbb{R}, b > 0.$$

$$= \frac{1}{2b} \begin{cases} \exp\left(-\frac{\mu-x}{b}\right) & \text{si } x < \mu, \\ \exp\left(-\frac{x-\mu}{b}\right) & \text{si } x \geq \mu. \end{cases} \quad (2.3)$$

La función de distribución acumulada está dada por:

$$F(x; \mu, b) = \begin{cases} \frac{1}{2} \exp\left(-\frac{\mu-x}{b}\right) & \text{si } x < \mu, \\ 1 - \frac{1}{2} \exp\left(-\frac{x-\mu}{b}\right) & \text{si } x \geq \mu. \end{cases} \quad (2.4)$$

La Figura 2.2 muestra algunas curvas de la densidad Laplace considerando diferentes valores de los parámetros.

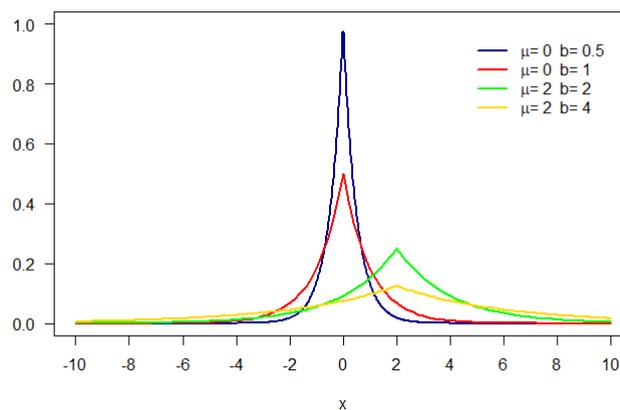


Figura 2.2: Densidad Laplace con parámetros μ y b .

2.3. Distribución Laplace asimétrica

La distribución Laplace descrita en la sección anterior se puede extender agregando un parámetro de asimetría λ el cual regula el sesgo de la distribución.

Esta distribución ha sido usada en áreas como la microbiología, para modelar la dispersión por citometría de flujo, el cual es un método analítico que permite la medición rápida de ciertas características físicas y químicas de células suspendidas en líquido ([23]) y en estadística, para el análisis de modelos de mezclas ([34]).

Se dice que una variable aleatoria X tiene distribución Laplace asimétrica con parámetro de localidad $\mu \in \mathbb{R}$, parámetro de escala $s > 0$ y de asimetría $\lambda > 0$ si su función de densidad de probabilidad está dada por:

$$f(x; \mu, s, \lambda) = \frac{\lambda}{(1 + \lambda^2)s} \begin{cases} \exp\left(\frac{x-\mu}{\lambda s}\right) & \text{si } x < \mu, \\ \exp\left(\frac{-\lambda(x-\mu)}{s}\right) & \text{si } x \geq \mu. \end{cases} \quad (2.5)$$

La media de esta distribución es $\mu + \frac{s(1-\lambda^2)}{\sqrt{2}\lambda}$ y la varianza es $\frac{s^2(1+\lambda^4)}{2\lambda^2}$, la librería `rmutil` ([41]) de R contiene la información completa y las funciones necesarias para el estudio de esta distribución.

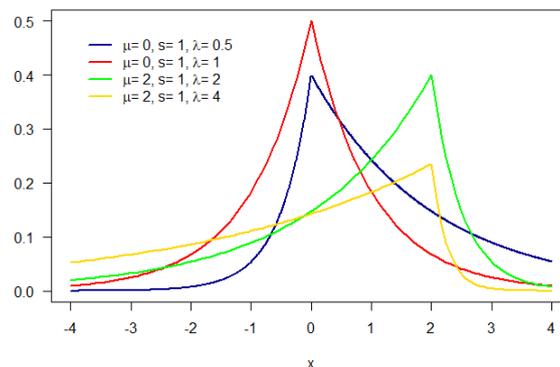


Figura 2.3: Densidad Laplace asimétrica con parámetros μ , s y λ .

2.4. Distribución t-Student

Una densidad utilizada comúnmente para estudiar datos con colas pesadas es la t-Student(ν). Esta distribución fue obtenida en 1908 por William Sealy Gosset ([40]) y se caracteriza por ser una distribución simétrica y, además, para los casos en que el parámetro de forma (ν) es finito, las colas son más pesadas que las de la distribución normal estándar.

Esta distribución es muy utilizada para calcular estimaciones por intervalo de confianza para la media, cuando no se conoce la desviación estándar poblacional para el caso de muestras pequeñas.

Sea Z una variable aleatoria normal estándar y V una variable aleatoria chi-cuadrada con ν grados de libertad, si Z y V son independientes, entonces la variable aleatoria X definida como $X = \frac{Z}{\sqrt{V/\nu}}$, tiene distribución t-Student(ν) ([44]) y su función de densidad de probabilidad está dada por:

$$t(x; \nu) = \frac{\Gamma\left(\frac{1}{2}(\nu + 1)\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}} \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \quad x \in \mathbb{R}, \nu > 0. \quad (2.6)$$

Note que en este caso la distribución no depende de parámetros de localidad ni escala, por lo que está en su forma estándar, para generalizar esta distribución se agrega un parámetro de localidad μ y uno de escala σ . De esta forma, la variable $Y = \mu + \sigma X$ se dice que tiene distribución t-Student no estandarizada y su función de densidad de probabilidad está dada por:

$$f(y; \mu, \sigma, \nu) = \frac{\Gamma\left(\frac{1}{2}(\nu + 1)\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu\sigma}} \left(1 + \frac{1}{\nu} \left(\frac{y - \mu}{\sigma}\right)^2\right)^{-(\nu+1)/2}, \quad (2.7)$$
$$y \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0, \nu > 0.$$

La Figura 2.4 muestra algunas curvas de la densidad t-student estándar para diferentes valores de ν .

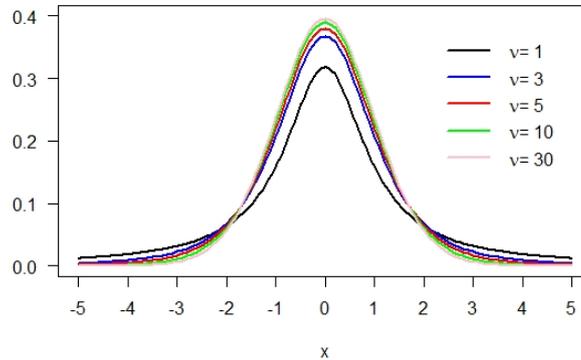


Figura 2.4: Densidad t-Student con ν grados de libertad.

2.5. Distribución Normal asimétrica

En teoría de probabilidad y estadística, la distribución normal asimétrica es una distribución de probabilidad continua la cual generaliza y extiende a la distribución normal mediante el uso de un parámetro de forma adicional que regula el sesgo. Esta distribución fue introducida por Azzalini ([3]) y ha sido utilizada en diferentes ámbitos aplicados como la medicina, para obtener la confiabilidad en el modelo fuerza-estrés ([20]); en estadística, para el estudio de modelos autoregresivos ([28]) y en ingeniería, para modelar tiempos de conmutación en circuitos ([9]).

Sean φ y Φ la función de densidad de probabilidad y la función de distribución acumulada de una variable aleatoria normal estándar respectivamente y sea $\alpha \in \mathbb{R}$, se dice que la variable aleatoria continua Z tiene distribución normal asimétrica con parámetro de asimetría α , denotada como $Z \sim SN(\alpha)$, si su función de densidad de probabilidad está dada por:

$$f_Z(z; \alpha) = 2\varphi(z)\Phi(\alpha z), \quad -\infty < z < \infty, \quad \alpha \in \mathbb{R}. \quad (2.8)$$

Sean $\xi \in \mathbb{R}$ y $\omega \in \mathbb{R}^+$, si $Z \sim SN(\alpha)$ entonces la variable aleatoria $Y = \xi + \omega Z$ se dice que es una variable aleatoria normal asimétrica con parámetros de localidad, escala y asimetría ξ , ω y α respectivamente, en este caso se denota $Y \sim SN(\xi, \omega^2, \alpha)$ y su función de densidad de probabilidad está dada por:

$$f_Y(y; \xi, \omega, \alpha) = \frac{2}{\omega} \varphi\left(\frac{y - \xi}{\omega}\right) \Phi\left(\alpha \frac{y - \xi}{\omega}\right), \quad -\infty < y < \infty. \quad (2.9)$$

Las propiedades generales de esta familia de distribuciones se proporcionan en ([5]).

La Figura 2.5 muestra algunas curvas de la densidad normal asimétrica con parámetro de localización 0 y escala 1 para diferentes valores del parámetro de asimetría α .

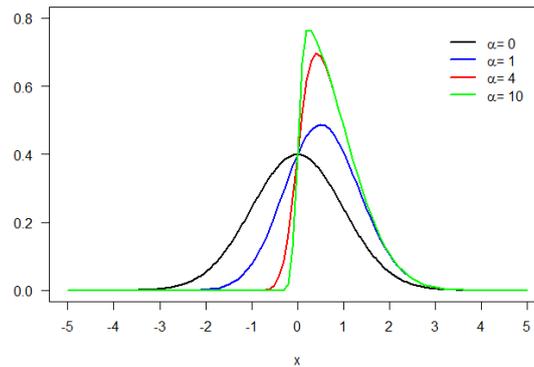


Figura 2.5: Densidad normal asimétrica para diferentes valores de α .

2.6. Distribución t-asimétrica

La distribución t-asimétrica ha sido estudiada recientemente debido a la amplia gama de generalizaciones que se han hecho a partir de distribuciones conocidas ([5]), ([19]). La principal característica que hace llamativa a esta distribución es la flexibilidad para modelar datos asimétricos que presenten colas más pesadas que las de la distribución normal además de que es una familia de distribuciones muy amplia ya que, para ciertos valores de sus parámetros, la distribución Cauchy(0,1), la distribución t-Student y la normal están contenidas dentro de esta familia.

La distribución t-asimétrica ha sido muy utilizada en diferentes aplicaciones en estadística, finanzas, economía, medicina, etc. Por ejemplo, en estadística ha sido usada para el modelado de mezclas de crecimiento ([18]), para el estudio de distribuciones desde un enfoque Bayesiano ([13]), para el estudio de modelos de volatilidad estocástica ([33]) y para el análisis del sesgo de los estimadores de máxima verosimilitud ([31]); en medicina, ha sido utilizada para el estudio de datos sesgados aplicados a la evaluación de las relaciones entre el maltrato infantil y las autolesiones no suicidas ([14]) y en economía, ha sido útil en el estudio y pronóstico de la demanda de turismo en China ([22]) y en investigaciones relacionadas con modelos de precios y volatilidad del mercado ([45]).

Sea $Z = \frac{Z_0}{\sqrt{V}}$, donde $Z_0 \sim SN(0, 1, \alpha)$ y $V \sim \chi_\nu^2/\nu$ son variables aleatorias independientes, entonces la variable aleatoria Z tiene distribución t-asimétrica con parámetro de asimetría α y ν grados de libertad si su función de densidad de probabilidad está dada por:

$$\begin{aligned}
t(x; \alpha, \nu) &= \int_0^\infty 2\varphi(x\sqrt{t})\Phi(\alpha x\sqrt{t})\sqrt{t}h(t) dt \\
&= \frac{2}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}} \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2} \times \\
&\quad \int_0^\infty e^{-u}u^{(\nu+1)/2}\Phi\left(\frac{\alpha x\sqrt{2u}}{\sqrt{x^2 + \nu}}\right) du \\
&= 2t(x; \nu)T\left(\alpha x\sqrt{\frac{\nu+1}{\nu+x^2}}; \nu+1\right), \quad x, \alpha \in \mathbb{R}, \nu > 0.
\end{aligned} \tag{2.10}$$

En este caso $SN(0, 1, \alpha)$ representa la distribución Normal asimétrica con parámetro de localidada 0, de escala 1 y de asimetría α ; $h(*)$ denota la función de densidad de la variable aleatoria V ; $t(x; \nu)$ representa la función de densidad de la t-Student con ν grados de libertad dada en (2.6) y $T(*; \nu+1)$ denota la función de distribución acumulada de la distribución t-Student con $\nu+1$ grados de libertad.

Algunas propiedades de esta densidad son:

Propiedad 1 Si $\alpha = 0$ la ecuación (2.10) se reduce a la densidad de la distribución t-Student con ν grados de libertad.

Propiedad 2 Si $\nu \rightarrow \infty$ la ecuación (2.10) converge a la densidad $SN(0, 1, \alpha)$.

Propiedad 3 La variable aleatoria $Z^2 \sim F(1, \nu)$, donde $F(\nu_1, \nu_2)$ denota la distribución de Snedecor con ν_1 y ν_2 grados de libertad.

La familia anterior se puede extender incluyendo un parámetro de localidada $\xi \in \mathbb{R}$ y uno de escala $\omega \in \mathbb{R}^+$, considerando la variable $Y = \xi + \omega Z$, se dice que Y tiene una distribución t-asimétrica con parámetros $(\xi, \omega^2, \alpha, \nu)$ y se denota como $Y \sim ST(\xi, \omega^2, \alpha, \nu)$.

La función de densidad de la variable $Y \sim ST(\xi, \omega^2, \alpha, \nu)$ se puede obtener de la siguiente forma, utilizando el método de transformación de variables.

Sea $F_Y(y)$ la función de distribución acumulada de la variable Y , entonces:

$$\begin{aligned}
 F_Y(y) &= Pr[Y \leq y] \\
 &= Pr[\xi + \omega Z \leq y] \\
 &= Pr\left[Z \leq \frac{y - \xi}{\omega}\right] \\
 &= F_Z\left(\frac{y - \xi}{\omega}\right)
 \end{aligned}$$

Derivando el resultado anterior respecto a y se sigue que:

$$\begin{aligned}
 f_Y(y) &= \frac{d F_Y(y)}{dy} \\
 &= \frac{d}{dy} F_Z\left(\frac{y - \xi}{\omega}\right) \\
 &= f_Z\left(\frac{y - \xi}{\omega}\right) \frac{d}{dy} \left(\frac{y - \xi}{\omega}\right) \\
 &= f_Z\left(\frac{y - \xi}{\omega}\right) \frac{1}{\omega}
 \end{aligned}$$

Así resulta que la función de densidad de $Y \sim ST(\xi, \omega^2, \alpha, \nu)$ es:

$$f_Y(y) = \omega^{-1} t(z; \alpha, \nu), \quad -\infty < y < \infty, \quad (2.11)$$

donde $z = \omega^{-1}(y - \xi)$ y $t(z; \alpha, \nu)$ representa la función de densidad de la $ST(0, 1, \alpha, \nu)$ es decir de la t-asimétrica estándar.

El procedimiento para la obtención de los momentos de la variable $Y \sim ST(\xi, \omega^2, \alpha, \nu)$ se puede ver en ([5]). El método general se basa en obtener los momentos de la variable $Z = \frac{Z_0}{\sqrt{V}}$ cuya densidad fue definida en la ecuación (2.10) utilizando que $Z_0 \sim SN(0, 1, \alpha)$ y $V \sim \chi_\nu^2/\nu$ son variables independientes, después se aplican propiedades básicas de la esperanza matemática obteniendo los siguientes resultados:

Sean

$$\begin{aligned}
 b_\nu &= \frac{\sqrt{\nu} \Gamma\left(\frac{1}{2}(\nu - 1)\right)}{\sqrt{\pi} \Gamma\left(\frac{1}{2}\nu\right)}, \quad \nu > 1, \\
 \delta &= \frac{\alpha}{\sqrt{1 + \alpha^2}}, \quad \delta \in (-1, 1),
 \end{aligned}$$

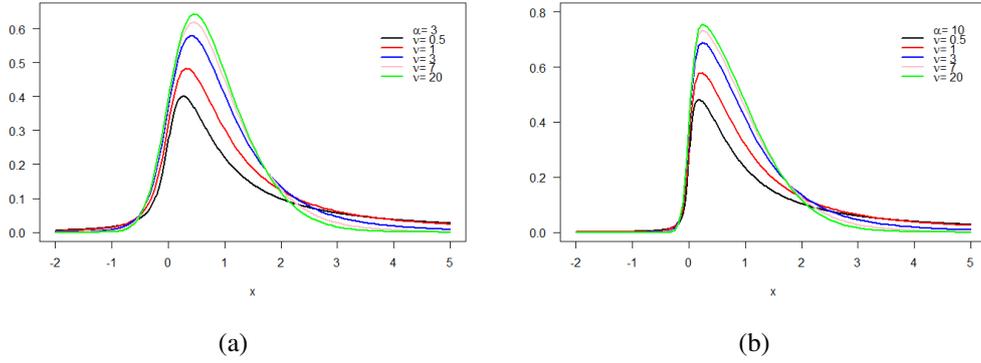


Figura 2.6: Densidad t-asimétrica con $\xi = 0$, $\omega = 1$ y diferentes valores de los parámetros α y ν .

se tiene que

$$\mu = \mathbb{E}\{Y\} = \xi + \omega b_\nu \delta, \text{ si } \nu > 1, \quad (2.12)$$

$$\sigma^2 = \text{var}\{Y\} = \omega^2 \left[\frac{\nu}{\nu - 2} - (b_\nu \delta)^2 \right] = \omega^2 \sigma_z^2, \text{ si } \nu > 2, \quad (2.13)$$

$$\gamma_1 = \frac{b_\nu \delta}{\sigma_z^{3/2}} \left[\frac{\nu(3 - \delta^2)}{\nu - 3} - \frac{3\nu}{\nu - 2} + 2(b_\nu \delta)^2 \right], \text{ si } \nu > 3, \quad (2.14)$$

$$\gamma_2 = \frac{1}{\sigma_z^4} \left[\frac{3\nu^2}{(\nu - 2)(\nu - 4)} - \frac{4(b_\nu \delta^2)\nu(3 - \delta^2)}{\nu - 3} + \frac{6(b_\nu \delta)^2 \nu}{\nu - 2} - 3(b_\nu \delta)^4 \right] - 3, \text{ si } \nu > 4, \quad (2.15)$$

Los coeficientes γ_1 y γ_2 representan el coeficiente de asimetría y el exceso de curtosis, este último sirve para comparar la distribución en estudio con la distribución normal la cual tiene exceso de curtosis 0.

El rango de γ_1 es $(-4, 4)$ cuando $\nu > 4$, para $\nu \leq 3$, γ_1 no existe; sin embargo, al menos una de las colas de la distribución es cada vez más pesada conforme $\nu \rightarrow 0$, dada la conexión con las distribuciones t-Student y Cauchy estándar. Por otro lado, si $\nu \rightarrow 4^+$ entonces el rango de γ_2 es $[0, \infty)$.

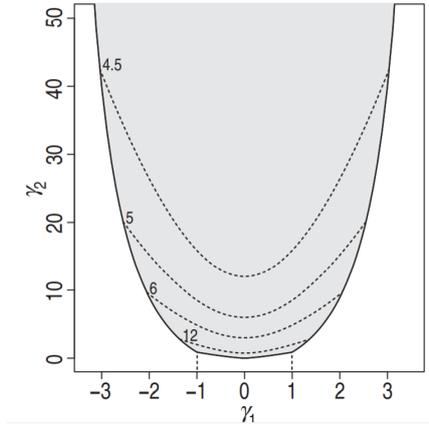


Figura 2.7: Rango de γ_1 y γ_2 y curvas asociadas para diferentes valores de los grados de libertad con α en la recta real.

2.6.1. Estimación de parámetros

Definición 2 Sea $\underline{X} = (X_1, \dots, X_n)$ una muestra de variables aleatorias continuas, independientes e idénticamente distribuidas con función de densidad común $f(x; \theta)$ que depende de un vector de parámetros desconocidos $\theta \in \mathbb{R}^p$, se define la función score $S(\theta; x)$ de la densidad como el vector de primeras derivadas parciales de la función log-verosimilitud con respecto al vector de parámetros θ .

Sea $L(\theta; x)$, la verosimilitud de \underline{X} la cual es una función real con dominio Θ y sea $l(\theta; x) = \ln(L(\theta; x))$, entonces, la función score está dada por:

$$S(\theta; x) = \frac{\partial l(\theta; x)}{\partial \theta}. \quad (2.16)$$

Definición 3 Dada una realización de la muestra digamos $\underline{x} = (x_1, \dots, x_n)$, se define el estimador de máxima verosimilitud de θ como el valor de θ que maximiza la probabilidad de observar \underline{x} bajo el modelo dado y se denota por $\hat{\theta}$, es decir:

$$\begin{aligned} \hat{\theta} &= \operatorname{argmax}_{\theta \in \Theta} (L(\theta; \underline{x})) \\ &= \operatorname{argmax}_{\theta \in \Theta} (l(\theta; \underline{x})). \end{aligned} \quad (2.17)$$

Note que el estimador de máxima verosimilitud de θ se obtiene resolviendo el sistema de

ecuaciones $S(\theta; \underline{x}) = \vec{0}$ con respecto a θ , donde $\vec{0}$ es el vector de ceros de dimensión igual a la de θ .

Sea $Y \sim ST(\xi, \omega^2, \alpha, \nu)$, denotando $\theta^{DP} = (\xi, \omega^2, \alpha, \nu)$, entonces el modelo basado en una observación y tiene la siguiente log-verosimilitud ([5]):

$$l_1(\theta^{DP}; y) = c - \log \omega - \frac{1}{2} \log \nu + \log \Gamma\left(\frac{1}{2}(\nu + 1)\right) - \log \Gamma\left(\frac{1}{2}\nu\right) - \frac{1}{2}(\nu + 1) \log\left(1 + \frac{z^2}{\nu}\right) + \log T(w; \nu + 1), \quad (2.18)$$

donde c es una constante,

$$z = \frac{y - \xi}{\omega}, \quad w = w(z) = \alpha z r, \quad r = r(z, \nu) = \sqrt{\frac{\nu + 1}{\nu + z^2}}.$$

Las componentes de la función score $S(y) = \{S_\xi(y), S_\omega(y), S_\alpha(y), S_\nu(y)\}'$ están dadas por:

$$S_\xi(y) = \frac{\partial l_1}{\partial \xi} = \frac{zr^2}{\omega} - \frac{\alpha \nu r h(w)}{w(\nu + z^2)},$$

$$S_\omega(y) = \frac{\partial l_1}{\partial \omega} = -\frac{1}{\omega} + \frac{(zr)^2}{\omega} - \frac{\nu w h(w)}{w(\nu + z^2)},$$

$$S_\alpha(y) = \frac{\partial l_1}{\partial \alpha} = z r h(w),$$

$$S_\nu(y) = \frac{\partial l_1}{\partial \nu} = \frac{1}{2} \left[\psi\left(\frac{1}{2}\nu + 1\right) - \psi\left(\frac{1}{2}\nu\right) - \frac{2\nu + 1}{\nu(\nu + 1)} - \log\left(1 + \frac{z^2}{\nu}\right) \right] + \frac{1}{2} \left[\frac{(zr)^2}{\nu} + \frac{\alpha z(z^2 - 1)h(w)}{(\nu + z^2)^2 r} + \frac{g(\nu)}{T(w; \nu + 1)} \right],$$

donde

$$\begin{aligned}\psi(x) &= \frac{d(\log \Gamma(x))}{dx}, \\ h(w) &= \frac{t(w; \nu + 1)}{T(w; \nu + 1)}, \\ g(\nu) &= \frac{d T(w(\alpha z r(z, \nu)); \nu + 1)}{d\nu} \\ &= \int_{-\infty}^w \left[\frac{(\nu + 2)x^2}{(\nu + 1)(\nu + 1 + x^2)} - \log \left(1 + \frac{x^2}{\nu + 1} \right) \right] t(x; \nu + 1) dx,\end{aligned}$$

$\psi(x)$ es conocida como la función digamma.

Note que debido a la dificultad algebraica de las componentes de la función score, es necesario utilizar métodos numéricos para resolver el sistema de ecuaciones involucrado, en este caso se utiliza el paquete `sn` ([4]) de R el cual contiene un método de estimación de parámetros. Muchos autores han estudiado la estimación de los parámetros de la distribución t-asimétrica desde diversos enfoques como el Bayesiano ([7]) y han tratado diversos problemas de estimación de dichos parámetros, por ejemplo, el problema de no convergencia del parámetro de forma en ciertos casos en que el tamaño de muestra es moderado ([27]), de igual forma se han propuesto diversas reparametrizaciones basadas en los momentos de la distribución ([2]).

2.7. Pruebas de hipótesis

A continuación se estudian algunos conceptos importantes sobre pruebas de hipótesis, varios de ellos definidos por Mood, Graybill & Boes ([32]).

Definición 4 Una hipótesis estadística es una aseveración o conjetura acerca de la distribución de una o más variables aleatorias. Si la hipótesis estadística especifica completamente la distribución es llamada hipótesis simple, de lo contrario se llama hipótesis compuesta. Comúnmente la hipótesis se denota con la letra H .

Una prueba de hipótesis consta de una hipótesis nula y una alternativa, denotadas como H_0 y H_1 respectivamente. La hipótesis alternativa es la que se establece en base a la evidencia que tenemos y es la que se desea probar.

En general, sea \underline{X} un vector de observaciones cuya distribución conjunta depende de un parámetro desconocido $\theta \in \Omega$, donde Ω representa el espacio paramétrico el cual es conocido.

Considérese la hipótesis nula $H_0 : \theta \in \omega$, $\omega \subset \Omega$ en contra de la hipótesis alternativa $H_1 : \theta \in (\Omega - \omega)$ (es decir $\theta \notin \omega$), donde ω es conocido.

Sea \mathcal{X} el conjunto de todas las realizaciones posibles de \underline{X} .

Definición 5 Cualquier partición de \mathcal{X} en dos subconjuntos, X_R y X_A tales que $\mathcal{X} = X_R \cup X_A$ y además $X_R \cap X_A = \emptyset$ define una prueba donde para una realización \underline{x} si $\underline{x} \in X_R$ entonces se rechaza H_0 , de otro modo si $\underline{x} \in X_A$, no se rechaza H_0 . Note que $X_A = X_R^c$.

Al conjunto X_R se le llama región de rechazo y al conjunto X_A región de no rechazo.

Note que cualquier función $\phi : \mathcal{X} \rightarrow \{0, 1\}$, en este caso $\phi(\underline{x})$ sólo puede tomar uno de dos valores ya sea que $\phi(\underline{x}) = 0$ ó $\phi(\underline{x}) = 1$, por tanto se pueden definir $X_R(\phi)$ y $X_A(\phi)$ como:

$$X_R(\phi) = \{\underline{x} \in \mathcal{X} : \phi(\underline{x}) = 1\},$$

$$X_A(\phi) = \{\underline{x} \in \mathcal{X} : \phi(\underline{x}) = 0\}.$$

Es decir, cada función $\phi : \mathcal{X} \rightarrow \{0, 1\}$ tiene asociada una prueba y la regla de decisión asociada a una función de prueba ϕ es tal que se rechaza H_0 si $\phi(\underline{x}) = 1$ y no se rechaza H_0 si $\phi(\underline{x}) = 0$.

Definición 6 Como las conclusiones a las que lleguemos en una prueba de hipótesis se basan en una muestra aleatoria, hay posibilidades de que nos equivoquemos en la decisión que se toma, en la literatura se definen dos tipos de errores los cuales ocurren de manera aleatoria:

Error tipo I: Rechazar H_0 cuando H_0 es verdadera.

Error tipo II: No rechazar H_0 cuando H_0 es falsa.

Se denotará el Error tipo I como $E.I$ y el Error tipo II como $E.II$.

Es deseable encontrar la prueba ϕ^* tal que $P(E.I, usando \phi^*)$ y $P(E.II, usando \phi^*)$ sean mínimas respecto a cualquier otra prueba ϕ .

Definición 7 La probabilidad de cometer el Error Tipo I se conoce como nivel de significancia, se denotará como α^* y es el tamaño de la región de rechazo, se dice que una prueba ϕ es de nivel α^* si $P(E.I, usando \phi) \leq \alpha^*$.

Definición 8 La función de potencia de una prueba ϕ , con parámetro hipotetizado θ denotada como $\beta_\phi(\theta)$ es la probabilidad de rechazar la hipótesis nula usando ϕ cuando el valor verdadero del parámetro es θ , es decir $\beta_\phi(\theta) = Pr(\text{Rechazar } H_0 \text{ usando } \phi \mid \theta)$.

	H_0 verdadera	H_0 falsa
Rechazar H_0	Error tipo I	Decisión correcta
No Rechazar H_0	Decisión correcta	Error tipo II

Cuadro 2.1: Tipos de errores.

Note que

$$Pr(E.I \text{ usando } \phi) = \beta_\phi(\theta), \theta \in \omega$$

$$Pr(E.II \text{ usando } \phi) = 1 - \beta_\phi(\theta), \theta \notin \omega.$$

Definición 9 El valor de probabilidad también llamado valor-p de una prueba de hipótesis sirve para cuantificar la incertidumbre o el error en una determinada afirmación, se define como la $Pr(\text{Rechazar } H_0 \text{ dado que } H_0 \text{ es verdadera})$, la regla de decisión es rechazar H_0 si el valor-p < α y tomar la decisión de no rechazar en caso contrario, en este caso α es el nivel de significancia de la prueba.

2.8. Pruebas de bondad de ajuste

En general las pruebas de bondad de ajuste permiten verificar que la población de la cual proviene una muestra tiene una distribución especificada o supuesta. El propósito de este tipo de pruebas es averiguar si existen diferencias estadísticamente significativas entre la distribución observada y la distribución esperada o supuesta.

Sea X variable aleatoria poblacional y $f_0(x)$ la distribución (o densidad) de probabilidad especificada o supuesta para X , se desea probar la siguiente hipótesis:

$$H_0 : X \sim f_0(x) \text{ en contraste con la hipótesis alternativa: } H_1 : X \not\sim f_0(x).$$

Existen diversas pruebas de bondad de ajuste, entre las más utilizadas están la prueba de Ji-cuadrada, la de Kolmogorov-Smirnov y la de Anderson-Darling, todas consideradas pruebas no paramétricas debido a que son pruebas libres de distribución, es decir no se utilizan para una única distribución específica y debido a esto la distribución del estadístico de prueba depende de la distribución objetivo.

2.8.1. Prueba de Anderson-Darling

T. W. Anderson & D. A. Darling ([1]) propusieron un estadístico que tiene como objetivo medir qué tan bien siguen los datos una distribución en particular, es decir, se usa para comprobar si una muestra de datos proviene de una población con una distribución especificada. Por lo general, mientras mejor se ajuste la distribución a los datos, menor será el valor tomado por el estadístico.

Las hipótesis para la prueba de Anderson-Darling son:

H_0 : Los datos siguen la distribución especificada.

H_1 : Los datos no siguen la distribución especificada.

Sean X_1, \dots, X_n una muestra aleatoria de $f(x; \theta)$; se define el estadístico de Anderson-Darling como sigue:

$$A^2 = n \int_{-\infty}^{\infty} \frac{[F_n(x) - \hat{F}_X(x)]^2}{\hat{F}_X(x) [1 - \hat{F}_X(x)]} d\hat{F}_X(x), \quad (2.19)$$

donde $F_n(\cdot) : \mathbb{R} \rightarrow [0, 1]$ es la función de distribución empírica de las X_i 's, la cual en general se define como:

$$F_n(x) = \begin{cases} 0 & \text{para } x < x_{(1)}, \\ \frac{i}{n} & \text{para } x_{(i)} \leq x < x_{(i+1)}, \\ 1 & \text{para } x \geq x_{(n)}, \end{cases} \quad (2.20)$$

donde $x_{(1)} = \min_i \{x_i\} \leq \dots \leq x_{(n)} = \max_i \{x_i\}$, es decir, son los X_i 's ordenados en forma creciente. $F_X(x)$ es la función de distribución supuesta que depende del vector de parámetros θ , $\hat{F}_X(x)$ es la distribución obtenida una vez que se han estimado los parámetros desconocidos.

La fórmula para calcular el estadístico A^2 dada una realización de X_1, \dots, X_n es:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left[\log \hat{F}_X(x_{(i)}) + \log(1 - \hat{F}_X(x_{(n+1-i)})) \right], \quad (2.21)$$

La hipótesis $H_0 : X_1, \dots, X_n \sim F_X(x)$ es rechazada con un tamaño de prueba $\alpha \in [0, 1]$ si $A^2 > k'_\alpha$ donde k'_α es el valor crítico, debido a que la distribución de A^2 depende en general de parámetros desconocidos, el valor crítico se obtiene a partir de métodos como el bootstrap paramétrico, el cual se describe más adelante.

2.8.2. Prueba de Shapiro-Wilk

La prueba de Shapiro-Wilk fue propuesta por Samuel Shapiro y Martin Wilk en 1965 ([39]) y se usa para contrastar la normalidad de un conjunto de datos por medio del uso de un estadístico W , el juego de hipótesis en esta prueba dada una muestra aleatoria X_1, \dots, X_n es:
 $H_0 : X_1, \dots, X_n$ sigue una distribución Normal vs $H_1 : X_1, \dots, X_n$ no sigue una distribución Normal.

El estadístico W se define como sigue:

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.22)$$

donde $x_{(i)}$ son los valores ordenados de la muestra, es decir $x_{(1)} = \min_i \{X_i\} \leq \dots \leq x_{(n)} = \max_i \{X_i\}$ y las a_i son constantes generadas a partir de las medias, varianzas y covarianzas de las estadísticas de orden de una muestra de tamaño n de una distribución normal estándar (ver Pearson y Hartley 1972, tabla 15 [36]).

La prueba rechaza la hipótesis nula para valores pequeños de W , los valores críticos fueron obtenidos mediante simulaciones Monte Carlo y fueron reproducidos por Pearson y Hartley en 1972 ([36]).

2.9. Método bootstrap paramétrico

El método Bootstrap fue propuesto por Bradley Efron ([12]) en 1979 como un método que reemplaza las técnicas complejas de análisis estadístico convencional por cálculos intensivos sobre las muestras generadas por simulación Monte-Carlo.

En ciertos casos no es posible calcular un valor de probabilidad o las constantes críticas de una prueba de hipótesis con exactitud debido a que no se conoce la distribución de probabilidad del estadístico de prueba. Cuando se cuenta con alguna aproximación, suele suceder que bajo ciertas circunstancias dicha aproximación podría no ser del todo buena. En estos casos, una alternativa para el cálculo del valor de probabilidad puede ser el bootstrap paramétrico.

Supongamos que sabemos que la función de distribución poblacional pertenece a cierta familia paramétrica, es decir, $F = F_\theta$, donde θ es el vector de parámetros p -dimensional. En ese caso se estiman los parámetros desconocidos, por ejemplo, mediante el método de máxima verosimilitud obteniendo $\hat{\theta}$. A partir de dicha estimación se generan muestras de $F_{\hat{\theta}}$ mediante el método Monte-Carlo y se estima la distribución del estadístico de interés R .

El método general es:

- 1) Calcular $\hat{\theta}$ a partir de la muestra X_1, \dots, X_n , la cual tiene distribución F_θ .
- 2) Generar una muestra bootstrap X_1^*, \dots, X_n^* de la distribución $F_{\hat{\theta}}$
- 3) Calcular el valor del estadístico con la muestra generada en 2) obteniendo R^* .
- 4) Repetir B veces el paso 2) y 3) para obtener B réplicas del estadístico R , es decir, $R^{*(1)}, \dots, R^{*(B)}$.
- 5) Utilizar las réplicas bootstrap para aproximar la distribución de R .
- 6) Ordenar los valores obtenidos en el paso 4) en forma ascendente.
- 7) Calcular el cuantil muestral del $100(1 - \alpha)\%$ el cual es la estadística de orden $B(1 - \alpha)$ de los valores obtenidos en el paso 6) y tomarlo como una aproximación de k_α , el cuantil poblacional del $100(1 - \alpha)\%$.

NOTAS:

En el caso de la prueba de Anderson-Darling se rechaza H_0 con un tamaño de prueba $\alpha \in (0, 1)$ si $A^2 > k_\alpha$ donde A^2 es el estadístico calculado con la muestra dada en 1) mediante la fórmula (2.21); en caso contrario, tomar la decisión de no rechazar H_0 .

A partir de la distribución estimada en (5) se pueden calcular probabilidades como el valor- p enunciado en la Definición 9, para el caso en que la prueba es de cola superior el valor- p se estima como $\frac{\sum_i (R^{*(i)} > R^*)}{B}$.

CAPÍTULO 3

METODOLOGÍA PROPUESTA PARA PROBAR AJUSTE DE LA DISTRIBUCIÓN T-ASIMÉTRICA

En el análisis exploratorio de los conjuntos de datos presentados en la Introducción , se observa gráficamente que la distribución t-asimétrica provee un ajuste razonable a los rendimientos del IPC. Debido a que en la literatura no existen pruebas formales para probar la hipótesis de que una muestra aleatoria sigue una distribución t-asimétrica, en este capítulo se presentan dos metodologías para resolver este problema. Se presentan algunos resultados teóricos fundamentales, se analizan los criterios de confiabilidad y validez.

3.1. Transformación de variables

Considere el siguiente resultado.

Teorema 3.1.1 Sea X una variable aleatoria tal que $X \sim ST(0, 1, \alpha, \nu)$ entonces la variable aleatoria Y definida como

$$Y = \begin{cases} X & \text{con probabilidad } 1/2, \\ -X & \text{con probabilidad } 1/2, \end{cases} \quad (3.1)$$

tiene distribución t-Student con ν grados de libertad.

DEMOSTRACIÓN 1 Por definición la función de densidad de una mezcla Z de k variables aleatorias X_1, \dots, X_k con densidad $f_j(x)$, $j = 1, 2, \dots, k$ y probabilidad p_j respectivamente está dada por:

$$f_Z(z) = \sum_{j=1}^k p_j f_j(z), \quad (3.2)$$

luego para obtener la densidad de Y es necesario conocer la distribución de la variable $-X$ la cual se obtiene a continuación.

Utilizando la definición de función de distribución de una variable aleatoria se tiene que

$$\begin{aligned} F_{-X}(z) &= Pr[-X \leq z] \\ &= Pr[X > -z] \\ &= 1 - Pr[X \leq -z] \\ &= 1 - F_X(-z). \end{aligned}$$

Por otro lado

$$\begin{aligned} f_{-X}(z) &= \frac{d}{dz} F_{-X}(z) \\ &= \frac{d}{dz} (1 - F_X(-z)) \\ &= f_X(-z). \end{aligned}$$

Por tanto resulta

$$\begin{aligned} f_{-X}(-z) &= f_X(-z) \\ &= 2t(-z; \nu)T \left(\alpha(-z) \sqrt{\frac{\nu+1}{\nu+z^2}}; \nu+1 \right). \end{aligned}$$

Utilizando la simetría de la función de densidad de la t-Student resulta que $t(z; \nu) = t(-z; \nu)$.

Por tanto

$$f_{-X}(-z) = 2t(z; \nu)T \left((-\alpha)z \sqrt{\frac{\nu+1}{\nu+z^2}}; \nu+1 \right). \quad (3.3)$$

De la ecuación anterior resulta que la variable $-X$ definida en (3.1) tiene distribución t-asimétrica estándar con parámetro de asimetría $-\alpha$ y parámetro de forma ν , es decir, $-X \sim ST(0, 1, -\alpha, \nu)$.

Por tanto, la función de densidad de la variable aleatoria Y está dada por:

$$\begin{aligned}
 f_Y(y) &= \frac{1}{2}f_X(y) + \frac{1}{2}f_{-X}(y) \\
 &= \frac{1}{2} \left[2 t(y; \nu) T \left(\alpha y \sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right) + 2 t(y; \nu) T \left((-\alpha)y \sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right) \right] \\
 &= t(y; \nu) \left[T \left(\alpha y \sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right) + T \left((-\alpha)y \sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right) \right]
 \end{aligned}$$

Usando propiedades de las distribuciones simétricas, sea T la función de distribución de una variable aleatoria t-Student centrada con ν grados de libertad note que $T(-a; \nu) = 1 - T(a; \nu)$ con $a \in \mathbb{R}$ entonces resulta

$$\begin{aligned}
 f_Y(y) &= t(y; \nu) \left[T \left(\alpha y \sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right) + 1 - T \left(\alpha y \sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right) \right] \\
 &= t(y; \nu).
 \end{aligned} \tag{3.4}$$

A partir de lo anterior se puede utilizar el Teorema para transformar variables t-asimétricas (estandarizadas) en variables t-simétricas lo cual resulta útil para reducir las estimaciones de parámetros, ya que con dicha transformación se elimina el parámetro de asimetría.

Una vez conocida la distribución de la variable Y se desea encontrar una prueba de bondad de ajuste tal que preserve el tamaño nominal de prueba y que sea potente para realizar inferencias acerca de la distribución de los datos de interés.

3.2. Prueba de Anderson-Darling para la distribución t-asimétrica

En trabajos recientes se ha estudiado esta prueba utilizando como herramienta principal el método de simulación Monte-Carlo y técnicas de bootstrap paramétrico utilizadas para estimar los cuantiles del estadístico de prueba ([30]).

Sea X_1, \dots, X_n una muestra aleatoria de la variable de interés, el objetivo es aplicar la prueba de Anderson-Darling a los datos para probar el siguiente juego de hipótesis:

$$H_0 : X_1, \dots, X_n \sim ST(\xi, \omega^2, \alpha, \nu) \quad \text{vs} \quad H_1 : X_1, \dots, X_n \not\sim ST(\xi, \omega^2, \alpha, \nu),$$

donde el vector de parámetros $(\xi, \omega^2, \alpha, \nu)$ es desconocido.

El algoritmo general es el siguiente:

- 1) Calcular los estimadores de $(\xi, \omega^2, \alpha, \nu)$ dados por $(\hat{\xi}, \hat{\omega}^2, \hat{\alpha}, \hat{\nu})$ respectivamente.
- 2) Calcular el valor del estadístico de Anderson-Darling A^2 usando la ecuación (2.21) y la muestra X_1, \dots, X_n .
- 3) Generar una muestra bootstrap X_1^*, \dots, X_n^* de la distribución $ST(\hat{\xi}, \hat{\omega}^2, \hat{\alpha}, \hat{\nu})$.
- 4) Calcular el valor del estadístico de Anderson-Darling A^{2*} usando la ecuación (2.21) y la muestra generada en 3).
- 5) Repetir B veces el paso 3) y 4) para obtener B valores del estadístico, digamos $A^{2*(1)}, \dots, A^{2*(B)}$.
- 6) Ordenar los valores obtenidos en el paso 5) en forma ascendente.
- 7) Calcular el cuantil muestral del $100(1 - \alpha)\%$, el cual es la estadística de orden $B(1 - \alpha)$ de los valores obtenidos en el paso 6) y tomarlo como una aproximación de k'_α , el cuantil poblacional del $100(1 - \alpha)\%$.
- 8) Tomar la decisión de rechazar H_0 con un tamaño de prueba $\alpha \in (0, 1)$ si $A^2 > k'_\alpha$ en caso contrario tomar la decisión de no rechazar H_0 .

3.3. Prueba de Anderson-Darling con base en datos transformados

Sea X_1, \dots, X_n una muestra aleatoria con distribución t asimétrica, el objetivo es investigar las propiedades de tamaño y potencia de la prueba de Anderson-Darling utilizando las observaciones que resultan al aplicar la transformación definida en (3.1). Sean Y_1, \dots, Y_n dichas observaciones, entonces se desea probar el siguiente juego de hipótesis:

$$H'_0 : Y_1, \dots, Y_n \sim t - Student(\mu, \sigma, \nu) \quad \text{vs} \quad H'_1 : Y_1, \dots, Y_n \not\sim t - Student(\mu, \sigma, \nu),$$

donde el vector de parámetros (μ, σ, ν) es desconocido.

El algoritmo general es el siguiente.

- 1) Utilizando X_1, \dots, X_n calcular los estimadores de $(\xi, \omega^2, \alpha, \nu)$ dados por $(\hat{\xi}, \hat{\omega}^2, \hat{\alpha}, \hat{\nu})$ respectivamente.
- 2) Estandarizar la muestra en 1) haciendo $Z_i = \frac{X_i - \hat{\xi}}{\hat{\omega}}$, $i = 1, 2, \dots, n$. De esta forma $Z_1, \dots, Z_n \sim ST(0, 1, \alpha, \nu)$.
- 3) Aplicar la transformación (3.1) a cada Z_i , bajo las condiciones del Teorema (3.1). Las nuevas observaciones $Y_1, \dots, Y_n \sim t - Student(\nu)$.

- 4) Usando Y_1, \dots, Y_n estimar los parámetros de localización, escala y grados de libertad de la *t-Student* dados por $(\hat{\mu}, \hat{\sigma}, \hat{\nu})$.
- 5) Calcular el valor del estadístico de Anderson-Darling A^2 usando la ecuación (2.21) y la muestra dada en 4).
- 6) Generar una muestra bootstrap Z_1^*, \dots, Z_n^* de la distribución $t - Student(\hat{\mu}, \hat{\sigma}, \hat{\nu})$.
- 7) Calcular el valor del estadístico de Anderson-Darling A^{2*} usando la ecuación (2.21) y la muestra en 6). Note que en este caso $\hat{F}_X(x)$ es la función de distribución acumulada correspondiente a la $t - Student(\hat{\mu}, \hat{\sigma}, \hat{\nu})$.
- 8) Repetir B veces los pasos 6) y 7) para obtener B valores de A^{2*} .
- 9) Ordenar los valores obtenidos en el paso 7) en forma ascendente.
- 10) Calcular el cuantil muestral del $100(1 - \alpha)\%$, el cual es la estadística de orden $B(1 - \alpha)$ de los valores obtenidos en el paso 9) y tomarlo como una aproximación de k'_α , el cuantil poblacional del $100(1 - \alpha)\%$.
- 11) Tomar la decisión de rechazar H'_0 con un tamaño de prueba $\alpha \in (0, 1)$ si $A^2 > k'_\alpha$ en caso contrario tomar la decisión de no rechazar H'_0 .

3.4. Estudio de Simulación Monte-Carlo

A continuación se presentan algunos cálculos computacionales obtenidos para comprobar si en efecto la prueba de Anderson-Darling en sus dos formas definidas en las Secciones 3.2 y 3.3 preserva el tamaño nominal de prueba y además para estudiar la potencia de dichas pruebas para diferentes distribuciones alternativas.

Se denotaron las dos pruebas estudiadas como W y W^* , la primera denota la prueba de Anderson-Darling utilizando la distribución *t*-asimétrica y la segunda utilizando la distribución *t-Student* basada en la transformación definida en la ecuación (3.1).

Las estimaciones del tamaño de prueba se obtuvieron utilizando 200 muestras Monte-Carlo y 200 bootstrap para la prueba W mientras que para la prueba W^* se utilizaron 500 muestras Monte-Carlo y 200 bootstrap. Se consideraron muestras de tamaños $n=50, 100, 500$ y un tamaño nominal de prueba $\alpha^* = 0.05$. Para estimar la potencia de las pruebas para la prueba W se utilizaron 200 muestras Monte-Carlo y 200 bootstrap mientras que para la prueba W^* se tomaron 1000 muestras Monte-Carlo y 200 muestras Bootstrap. El número de muestras Monte-Carlo y

bootstrap se consideró en base al criterio de tiempo computacional debido a que los algoritmos son computacionalmente muy demandantes.

Los cálculos se realizaron utilizando el paquete estadístico R ([42]). La estimaciones de los parámetros se realizaron usando el paquete sn ([4]) de R.

Para la estimación del tamaño de prueba se utilizaron diferentes valores para los parámetros $\alpha \in [-20, 20]$ y $\nu > 0$, dichos parámetros corresponden a la asimetría y los grados de libertad respectivamente. Los valores considerados para los parámetros de localización y escala fueron 0 y 1 debido a que las dos pruebas son invariantes ante cambios de localización y escala. Cuando dichos parámetros se estiman utilizando estimadores equivariantes de ubicación y escala (como lo son los estimadores de máxima verosimilitud), las distribuciones muestrales de las estadísticas de la función de distribución empírica no dependen de los valores reales de esos parámetros ([11]).

α	ν	n=50		n=100		n=500	
		W	W*	W	W*	W	W*
20	3	0.03	0.04	0.03	0.04	0.08	0.04
15	3	0.05	0.05	0.02	0.05	0.04	0.06
10	3	0.01	0.03	0.06	0.05	0.05	0.04
5	3	0.01	0.05	0.02	0.05	0.04	0.05
2	3	0.03	0.04	0.02	0.04	0.02	0.05
0	3	0.02	0.04	0.05	0.04	0.04	0.05
-2	3	0	0.05	0.08	0.04	0.07	0.03
-5	3	0.02	0.04	0.02	0.05	0.07	0.05
-10	3	0.02	0.03	0.01	0.05	0.05	0.03
-15	3	0.01	0.05	0.01	0	0.06	0.03
-20	3	0.02	0.04	0.02	0.04	0.03	0.04
20	5	0.03	0.04	0.03	0.04	0.06	0.05
15	5	0.01	0.04	0.02	0.05	0.07	0.05
10	5	0.01	0.05	0.02	0.03	0.05	0.03
5	5	0.01	0.03	0.03	0.04	0.07	0.03
2	5	0	0.04	0.04	0.03	0.07	0.04
0	5	0.02	0.04	0.04	0.04	0.09	0.05
-2	5	0.01	0.04	0.05	0.02	0.06	0.03
-5	5	0.01	0.03	0.04	0.04	0.04	0.02
-10	5	0.01	0.02	0.02	0.02	0.06	0.05
-15	5	0.03	0.05	0.02	0.04	0.10	0.04
-20	5	0.03	0.04	0.03	0.03	0.03	0.05
20	8	0.05	0.04	0.01	0.04	0.04	0.05
15	8	0.01	0.03	0.02	0.02	0.07	0.06
10	8	0.01	0.02	0.04	0.04	0.04	0.05
5	8	0	0.04	0.04	0.04	0.04	0.04
2	8	0.01	0.04	0.04	0.04	0.08	0.03
0	8	0.05	0.05	0.05	0.05	0.03	0.05
-2	8	0.02	0.04	0.05	0.03	0.05	0.04
-5	8	0.02	0.02	0.01	0.03	0.03	0.03
-10	8	0.01	0.03	0.01	0.04	0.05	0.03
-15	8	0.02	0.05	0.05	0.04	0.05	0.04
-20	8	0.02	0.05	0.02	0.04	0.05	0.05
20	12	0.02	0.03	0.04	0.03	0.05	0.04
15	12	0.02	0.05	0.02	0.03	0.02	0.05
10	12	0.02	0.05	0.03	0.04	0.03	0.04
5	12	0.01	0.03	0.03	0.06	0.05	0.02
2	12	0.02	0.02	0.03	0.04	0.05	0.06
0	12	0.04	0.03	0.06	0.05	0.06	0.04
-2	12	0.02	0.04	0.02	0.03	0.05	0.06
-5	12	0.02	0.03	0.02	0.03	0.06	0.02
-10	12	0.03	0.04	0.03	0.03	0.05	0.05
-15	12	0.03	0.03	0.01	0.03	0.06	0.05
-20	12	0.02	0.03	0.04	0.04	0.06	0.04

Cuadro 3.1: Tamaño de pruebas estimado para diferentes valores de α y $\nu = 3, 5, 8, 12$ con $\alpha^* = 0.05$.

α	ν	n=50		n=100		n=500	
		W	W^*	W	W^*	W	W^*
2	1	0.01	0.05	0	0.05	0	0.06
2	4	0.02	0.03	0.03	0.04	0.05	0.04
2	6	0.05	0.03	0.03	0.04	0.07	0.04
2	10	0.02	0.04	0.03	0.03	0.04	0.04
2	15	0.05	0.04	0.01	0.04	0.06	0.03
-20	1	0.02	0.04	0.06	0.05	0.06	0.06
-2	1	0.02	0.03	0.06	0.05	0	0.06
-10	4	0.01	0.05	0	0.03	0.06	0.05
-5	6	0.02	0.03	0.01	0.05	0.05	0.04
0	1	0.02	0.05	0.03	0.04	0	0.04
0	7	0.03	0.07	0.04	0.04	0.07	0.05
3	7	0	0.04	0.03	0.03	0.05	0.04
2	8	0.02	0.03	0.03	0.04	0.05	0.05
5	15	0.01	0.04	0.03	0.04	0.06	0.04
10	30	0.02	0.02	0.02	0.05	0.03	0.04

Cuadro 3.2: Tamaño de pruebas estimado para diferentes valores de α y ν con $\alpha^* = 0.05$.

En el Cuadro 3.1 se observa que el tamaño nominal de la prueba se preserva en ambas pruebas cuando las muestras son de tamaño 50, por otro lado para muestras de tamaño 100 se observa que el tamaño estimado es mayor al nominal en algunos casos en que $\nu = 12$, para el caso en que el tamaño de muestra es 500 se observa que la prueba W no controla el error tipo I, mientras que la prueba W^* no controla dicho error en algunos casos en que los parámetros ν y α son muy grandes.

En el Cuadro 3.2 se observa que la prueba W^* preserva el tamaño nominal para los casos en que las muestras son de tamaño 50 y 100, excepto en el caso en el que $\alpha = 0$ y $\nu = 7$, mientras que la prueba W no tiene control del error tipo I en algunos casos en que $\nu = 1$ es decir en el caso en que la distribución tiene colas muy pesadas. Note que cuando el tamaño de muestra es igual a 500 ambas pruebas parecen no controlar el error tipo I, pero la prueba W presenta un número mayor de casos en los que el error es mayor.

Es importante tomar en cuenta que los resultados están sujetos a la variabilidad de las simulaciones y al número de muestras Monte-Carlo y bootstrap consideradas.

La potencia de las pruebas W y W^* fue estimada utilizando las siguientes distribuciones alternativas:

1) $N(0, 1)$: Normal estándar

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-x^2}{2}\right\}, \quad x \in \mathbb{R}.$$

2) $\text{Cauchy}(0, 1)$: Cauchy estándar

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}.$$

3) $\text{Logística}(0, 1)$: Logística estándar

$$f_X(x) = \frac{\exp\{-x\}}{(1 + \exp\{-x\})^2}, \quad x \in \mathbb{R}.$$

4) $\text{Beta}(a, b)$

$$f_X(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)}, \quad a > 0, b > 0, x \in [0, 1].$$

5) $\text{Uniforme}(0, 1)$

$$f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

6) $\text{Laplace}(0, 1)$: Laplace estándar

$$f_X(x) = \frac{1}{2} \exp\{-|x|\}, \quad x \in \mathbb{R}.$$

7) $\text{Gamma}(a, b)$

$$f_X(x) = \begin{cases} \frac{x^{a-1} \exp\left\{\frac{-x}{b}\right\}}{\Gamma(a)b^a} & x > 0, \\ 0 & x \leq 0, \end{cases}$$

$a > 0, b > 0.$

Nota: Cuando $a=1$ y $b=1$ se obtiene la densidad de la distribución exponencial estándar $\text{Exp}(1)$. Cuando $a=4/2$ y $b=2$ se obtiene la densidad de la distribución chi-cuadrada con 4 grados de libertad, denotada como χ_4^2 .

8) Gumbel(0, 1)

$$f_X(x) = \frac{\exp\{-\exp(-x)\}}{\exp\{x\}}, \quad x \in \mathbb{R}.$$

9) sLaplace(0, 1, λ): Laplace asimétrica estándar con parámetro de asimetría $\lambda \in \mathbb{R}$.

$$f_X(x) = \begin{cases} \frac{\lambda e^{-\lambda x}}{1+\lambda^2} & x \geq 0, \\ \frac{\lambda e^{x/\lambda}}{1+\lambda^2} & x < 0. \end{cases}$$

10) Log-Normal(μ, σ): Lognormal con parámetro de localidad $\mu \in \mathbb{R}$ y de escala $\sigma > 0$.

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right\}, \quad x > 0.$$

11) Weibull(λ, k): Weibull con parámetro de escala $\lambda > 0$ y parámetro de forma $k > 0$.

$$f_X(x) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^k\right\} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

12) ParetoG($\lambda, 1$): Pareto Generalizada con parámetro de forma λ .

$$f_X(x) = (1 + \lambda x)^{-(1/\lambda+1)},$$

donde $x \geq 0$ para $\lambda \geq 0$ y $0 \leq x \leq -1$ para $\lambda < 0$.

13) DParetoG(α): Doble Pareto generalizada

Se utilizó el siguiente resultado:

Sean X y Y variables aleatorias independientes con distribución ParetoG($\alpha, 1$) entonces la

variable $Z = X - Y$ tiene distribución Doble Pareto generalizada con parámetro α .

Note que en la lista de distribuciones se incluyeron la distribución Normal(0, 1) y la Cauchy(0, 1) pero debe recordarse que estas pertenecen a la familia de distribuciones t-asimétrica bajo ciertos parámetros.

El Cuadro 3.3 presenta las potencias estimadas de las pruebas para diferentes valores de los parámetros de las distribuciones alternativas. Las distribuciones alternativas se clasificaron como simétricas o asimétricas y se ordenaron de acuerdo a los valores de los coeficientes de asimetría γ_1 y de exceso de curtosis γ_2 .

Se puede ver que las dos pruebas tienen alta potencia para los casos en que las distribuciones alternativas tienen como dominio algún subconjunto de los números reales por ejemplo la Beta(0.5, 0.5) y la Uniforme(0, 1). Note que para el caso en el que $n=500$ la prueba W tiene mayor potencia en la mayoría de alternativas y en específico cuando la distribución es Laplace tanto simétrica como asimétrica o una distribución relacionada a éstas, como lo es la exponencial.

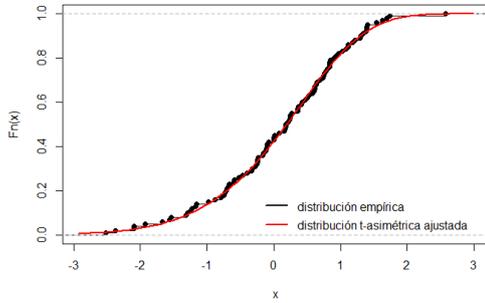
Para comparar los ajustes con las diferentes distribuciones alternativas utilizadas en las dos pruebas W y W^* se generaron las siguientes gráficas presentadas en las Figuras 3.1-3.8. Se tomaron muestras de tamaño $n=100$ de la distribución alternativa y se obtuvo su respectiva curva empírica y la curva de la distribución t-asimétrica (en el caso de la prueba W) estimando los respectivos parámetros. De forma análoga se realizó con la curva t-Student que resulta al aplicar la transformación que define la prueba W^* .

Alternativa	γ_1	γ_2	n=50		n=100		n=500	
			W	W*	W	W*	W	W*
Distribuciones simétricas								
Beta(0.5, 0.5)	0	-1.5	0.93	0.60	1	0.99	1	1
N(0, 1)	0	0	0.04	0.04	0.07	0.04	0.06	0.03
Cauchy(0, 1)	*	*	0.03	0.05	0.02	0.07	0.01	0.04
Logística(0, 1)	0	1.2	0.02	0.04	0.04	0.04	0.06	0.05
Uniforme(0, 1)	0	1.8	0.28	0.18	0.97	0.78	1	1
Laplace(0, 1)	0	3	0.05	0.05	0.14	0.07	0.62	0.33
Distribuciones asimétricas								
sLaplace(0, 1, 3)	-1.96	5.8	0.09	0.29	0.95	0.80	1	1
sLaplace(0, 1, 1.5)	-1.4	4.34	0.05	0.06	0.29	0.07	0.99	0.29
Beta(1, 3)	0.75	0.10	0.12	0.04	0.18	0.08	0.39	0.21
Gamma(4, 1)	1	1.5	0.01	0.03	0.05	0.04	0.14	0.04
Gumbel(0, 1)	1.14	2.4	0.01	0.04	0.05	0.05	0.08	0.04
sLaplace(0, 1, 0.7)	1.27	4.12	0.06	0.05	0.15	0.07	0.95	0.21
χ_4^2	1.41	3	0	0.04	0.01	0.04	0.16	0.08
Log-normal(0, 0.5)	1.75	5.9	0.03	0.03	0.05	0.03	0.10	0.05
sLaplace(0, 1, 0.5)	1.8	5.33	0.29	0.09	0.81	0.17	1	1
Exp(1)	2	6	0.14	0.06	0.24	0.08	0.80	0.39
Gamma(0.5, 1)	2.83	12	0.04	0.13	0.29	0.25	0	1
Weibull(0.75, 1)	3.1	16.03	0.36	0.10	0.56	0.17	0.11	0.97
ParetoG(0.15, 1)	3.5	30.5	0.16	0.06	0.25	0.09	0.80	0.28
Log-normal(0, 1.5)	33.47	10075.25	0.02	0.07	0.03	0.07	0	0.18
ParetoG(0.5, 1)	*	*	0.07	0.07	0.08	0.07	0.01	0.15
DParetoG(0.4)	*	*	0.02	0.06	0.10	0.08	0.13	0.17
DParetoG(2.5)	*	*	0.34	0.11	0.45	0.20	0.73	0.41

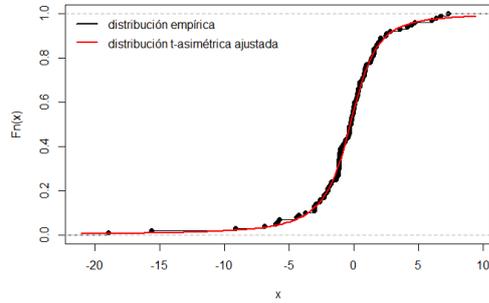
* Coeficientes no definidos

Cuadro 3.3: Potencia estimada de las pruebas utilizando $\alpha^* = 0.05$.

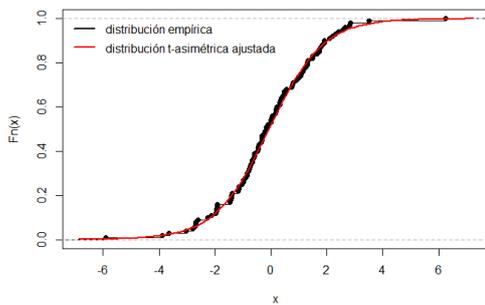
Figura 3.1: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=100$, parte I.



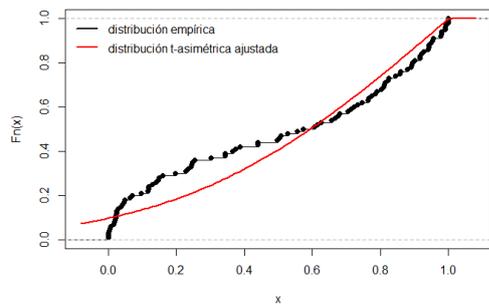
(a) Normal(0, 1)



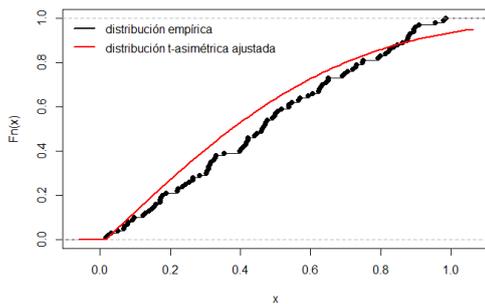
(b) Cauchy(0, 1)



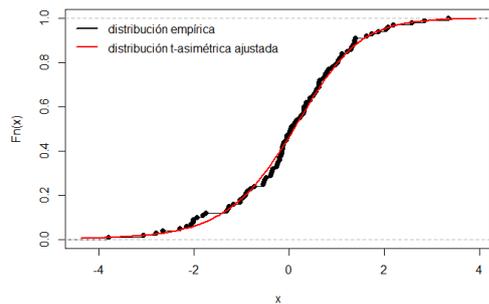
(c) Logística(0, 1)



(d) Beta(0.5, 0.5)

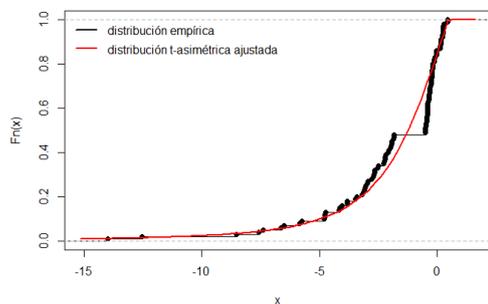


(e) Uniforme(0, 1)

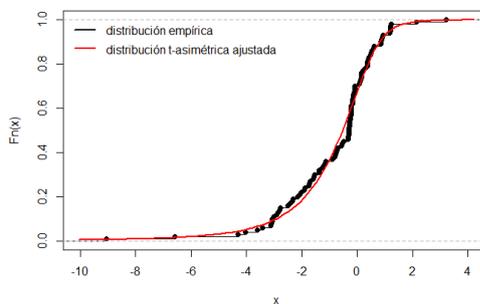


(f) Laplace(0, 1)

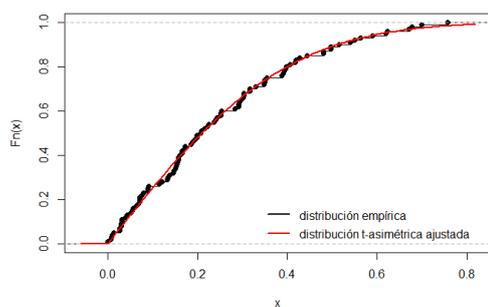
Figura 3.2: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=100$, parte II.



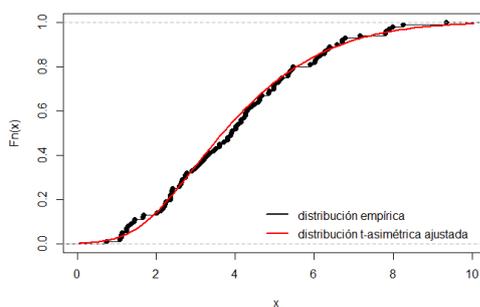
(a) sLaplace(0 ,1, 3)



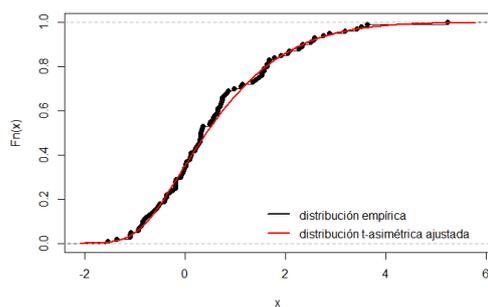
(b) sLaplace(0 ,1, 1.5)



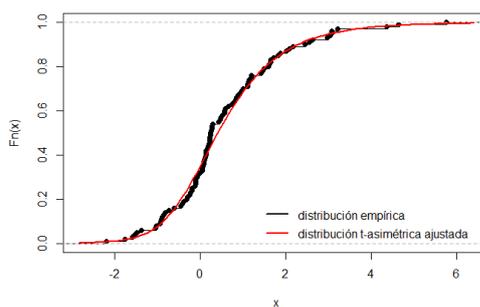
(c) Beta(1, 3)



(d) Gamma(4, 1)

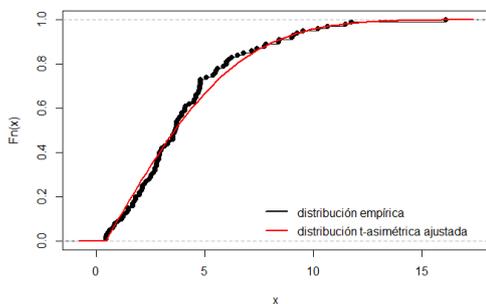


(e) Gumbel(0, 1)

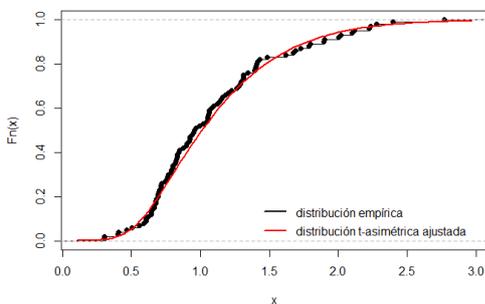


(f) sLaplace(0, 1, 0.7)

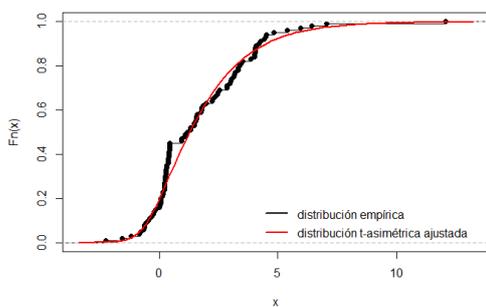
Figura 3.3: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=100$, parte III.



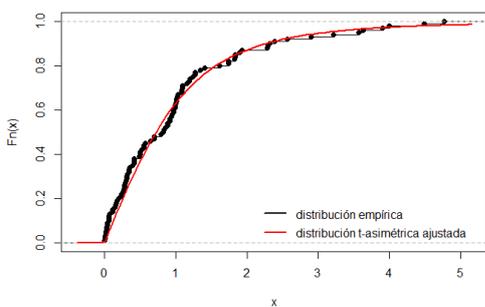
(a) χ_4^2



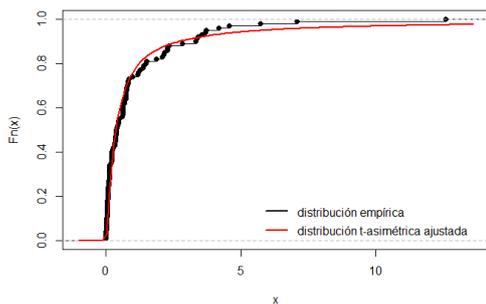
(b) Log-Normal(0, 0.5)



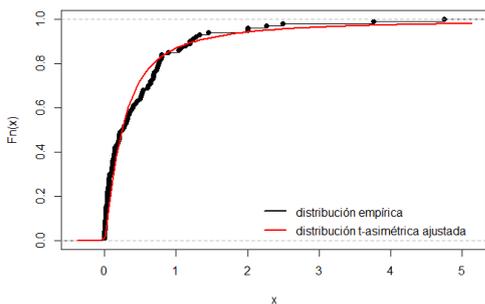
(c) sLaplace(0, 1, 0.5)



(d) Exp(1)

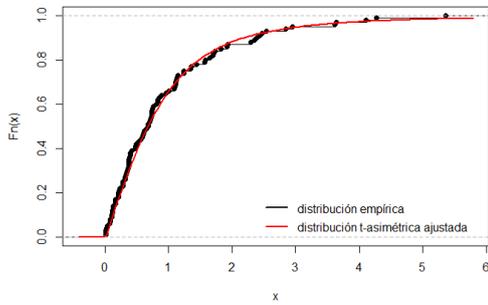


(e) Weibull(0.75, 1)

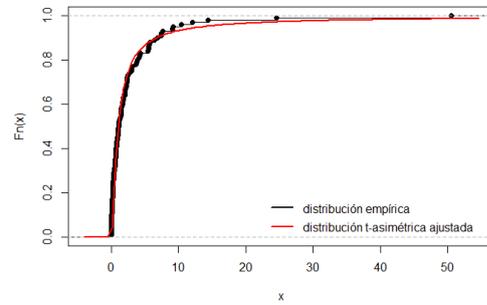


(f) Gamma(0.5, 1)

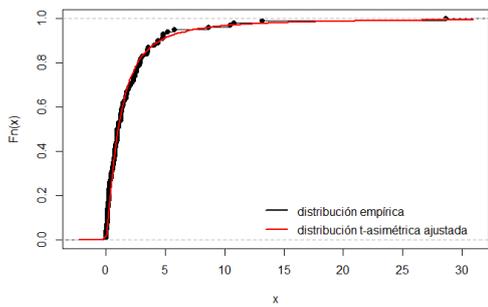
Figura 3.4: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=100$, parte IV.



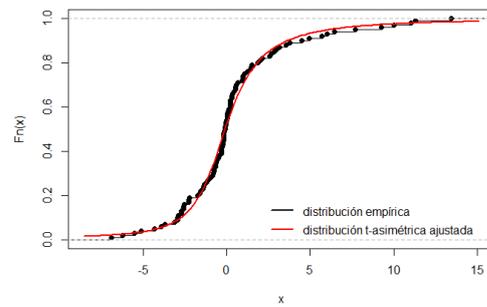
(a) ParetoG(0.15, 1)



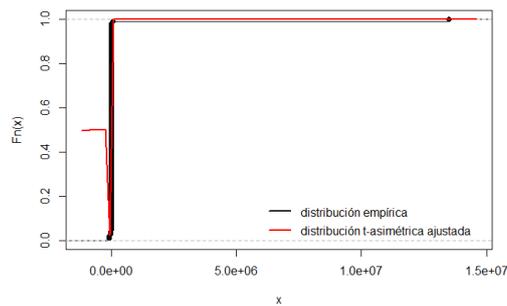
(b) Log-Normal(0, 1.5)



(c) ParetoG(0.5, 1)

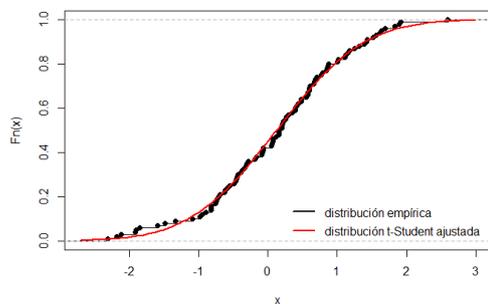


(d) DParetoG(0.4, 1)

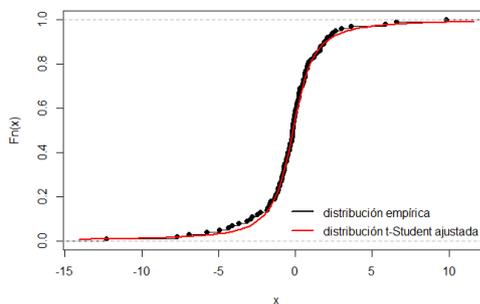


(e) DParetoG(2.5, 1)

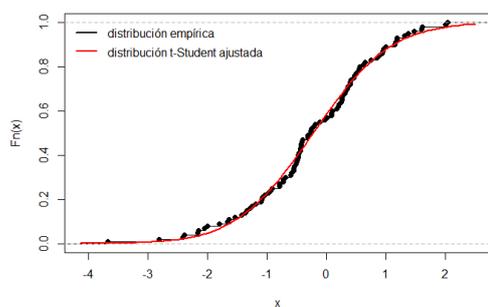
Figura 3.5: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=100$, parte I.



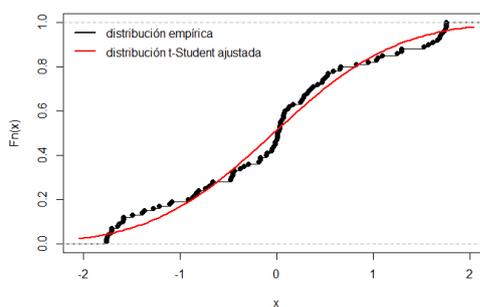
(a) Normal(0, 1)



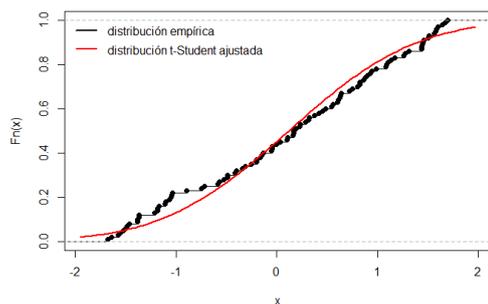
(b) Cauchy(0, 1)



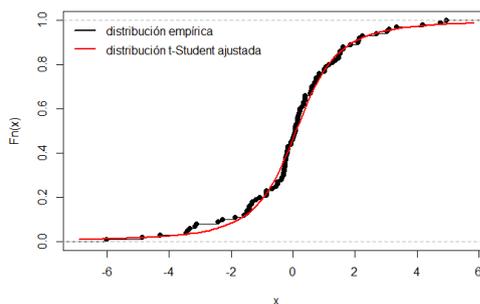
(c) Logística(0, 1)



(d) Beta(0.5, 0.5)

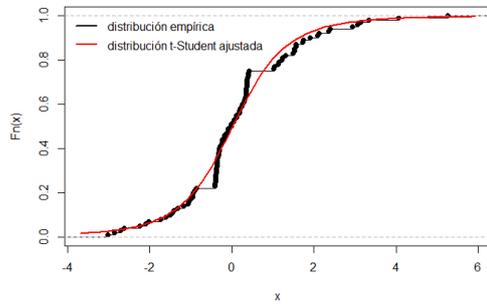


(e) Uniforme(0, 1)

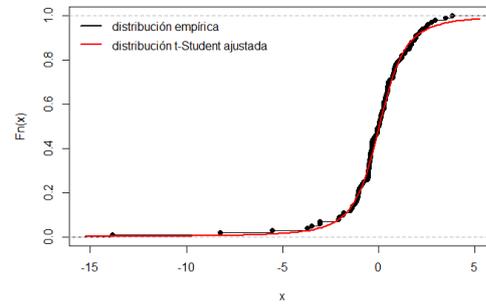


(f) Laplace(0, 1)

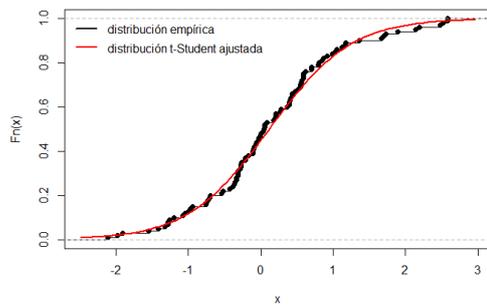
Figura 3.6: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=100$, parte II.



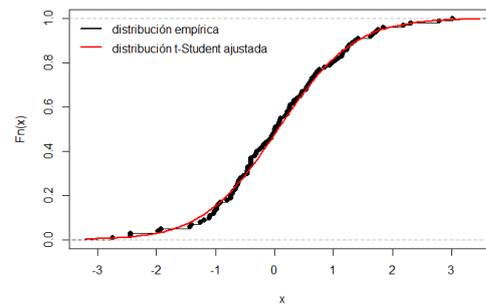
(a) sLaplace(0 ,1, 3)



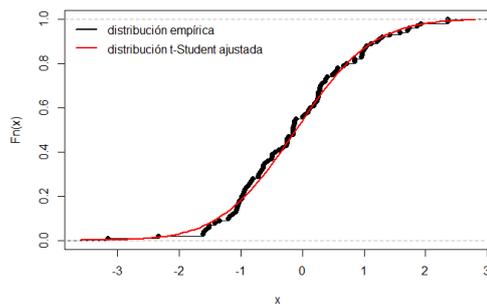
(b) sLaplace(0 ,1, 1.5)



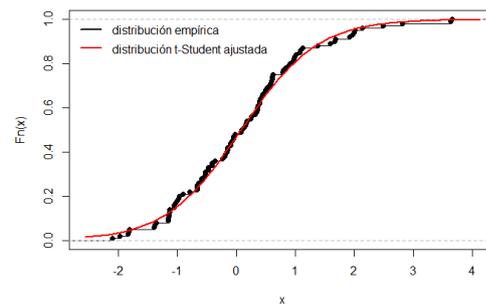
(c) Beta(1, 3)



(d) Gamma(4, 1)

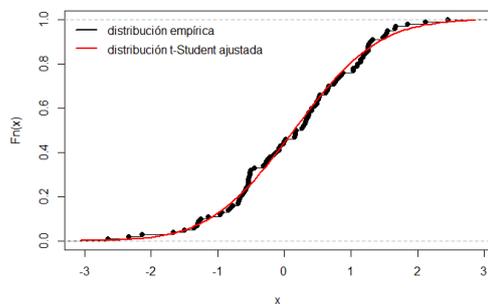


(e) Gumbel(0, 1)

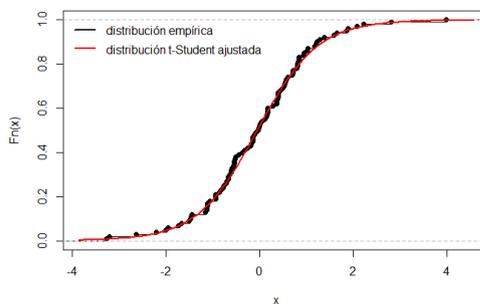


(f) sLaplace(0, 1, 0.7)

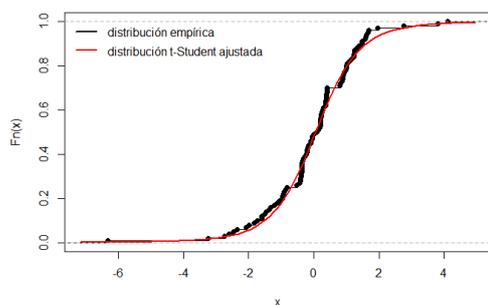
Figura 3.7: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=100$, parte III.



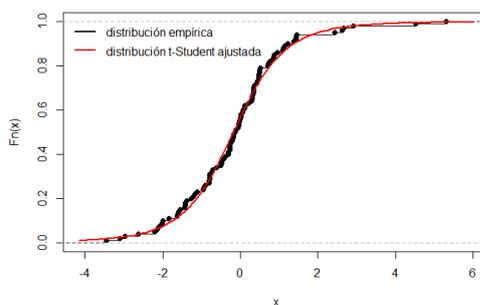
(a) χ_4^2



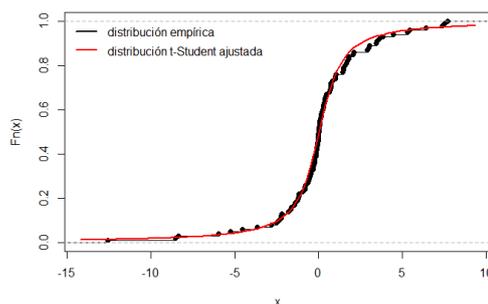
(b) Log-Normal(0, 0.5)



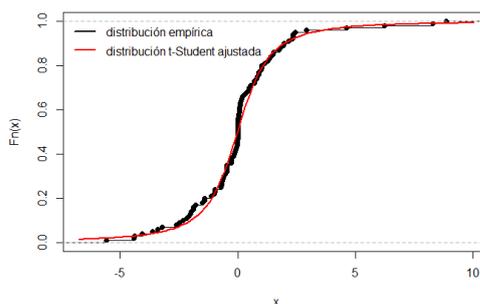
(c) sLaplace(0, 1, 0.5)



(d) Exp(1)

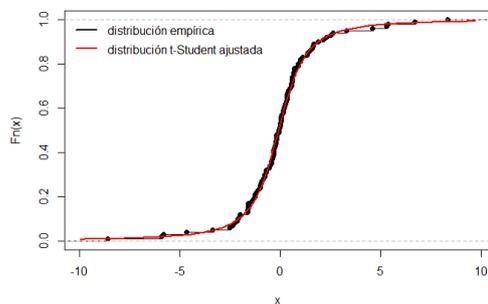


(e) Weibull(0.75, 1)

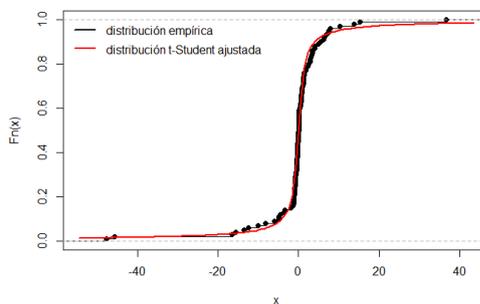


(f) Gamma(0.5, 1)

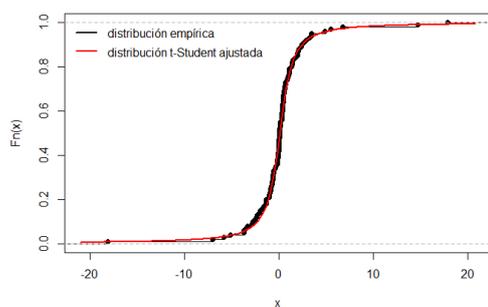
Figura 3.8: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=100$, parte IV.



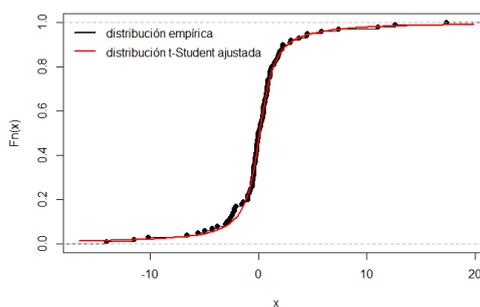
(a) ParetoG(0.15, 1)



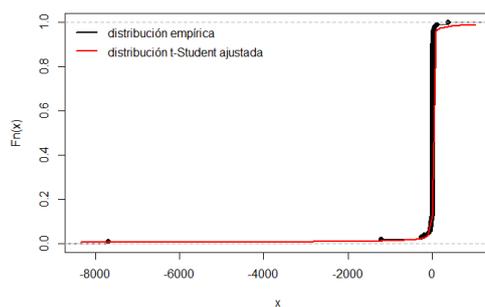
(b) Log-Normal(0, 1.5)



(c) ParetoG(0.5, 1)



(d) DParetoG(0.4, 1)



(e) DParetoG(2.5, 1)

En las Figuras 3.1(d) y 3.1(e) se observa que en el caso de las distribuciones Beta(0.5,0.5) y la Uniforme(0,1). Las curvas ajustadas con la distribución t-asimétrica se alejan bastante de la curva empírica y como se puede ver en el Cuadro 3.3 la potencia en estos casos es 1 y 0.97, respectivamente. Lo cual indica que la prueba puede detectar de manera eficaz muestras de estas distribuciones y rechazar la correspondiente hipótesis nula.

Por otro lado, la Figura 3.2(a) muestra que para la distribución Laplace asimétrica con parámetros (0,1,3), la curva ajustada difiere en gran parte de la curva empírica, lo cual es notable ya que se obtuvo una potencia de 0.95. Para las otras distribuciones incluidas en esta figura. Se puede notar que la curva ajustada no difiere en gran medida de la curva empírica y por tanto, como se puede ver en el Cuadro 3.3 en estos casos la potencia es baja.

También en la Figura 3.3(c), se observa que la curva ajustada con la distribución t-asimétrica no coincide en gran parte con la curva empírica generada con la distribución Laplace-asimétrica con parámetros (0,1,0.5). En este caso en la prueba W tiene una potencia alta de 0.81, lo mismo sucede con las distribuciones de las figuras 3.3(e) y 3.4(e) aunque para estos casos la potencia de la prueba W es menor 0.56 y 0.45 respectivamente.

Para las distribuciones alternativas utilizadas en la Figura 3.5 se observa que la curva t-Student ajustada a la muestra de la distribución Beta(0.5,0.5) y Uniforme(0,1) no tiene un buen ajuste a la curva empírica. Note que en estos casos en el Cuadro 3.3 se obtuvo una potencia de 0.99 y 0.78 respectivamente lo cual indica que contra estas distribuciones alternativas la prueba es eficaz. En la Figura 3.6(a) se observa que la curva ajustada difiere de la curva empírica en mayor parte en el centro de la curva, para este caso la potencia obtenida fue de 0.80. Mientras que para las otras distribuciones la curva ajustada es muy cercana a la curva empírica, en esos casos la potencia de la prueba W^* es baja. Note que en las figuras posteriores 3.7 y 3.8 las curvas ajustadas y las curvas empíricas no presentan grandes diferencias y para esos casos el Cuadro 3.3 presenta potencias bajas.

Por otra parte, en el Cuadro 3.3 se presentan algunos casos en que la potencia disminuye cuando el tamaño de muestra aumenta a 500, un ejemplo de esto fueron los resultados obtenidos al aplicar la prueba W^* a las muestras de la distribución Gumbel(0,1), ya que para $n=50$ la potencia fue de 0.04, para $n=100$ fue de 0.05 y para $n=500$ fue de 0.04. Debido a que las muestras tomadas son aleatorias se analizaron algunas y se notó que para ciertas muestras cuando $n=500$, las estimaciones de los grados de libertad no son finitas. Por lo cual, se infiere que los casos en los que sucede esto es debido a errores en las estimaciones o debido a la variabilidad de

estas. Lo mismo sucede con algunos casos en los que se calcula la potencia de la prueba W . Por ejemplo, para muestras de la distribución $\text{Gamma}(0.5,1)$ con $n=50$ la potencia fue de 0.04, con $n=100$ de 0.29 y con $n=500$ fue de 0, analizando dichas muestras se observa que la estimación del parámetro de asimetría de la t - asimétrica es muy grande y las estimaciones de los grados de libertad son cercanas a 1. De igual forma sucede para muestras de la $\text{Weibull}(0.75,1)$, $\text{lognormal}(0,1.5)$ y la $\text{Pareto Generalizada}(0.5,1)$.

Para comparar estas muestras se realizaron algunas figuras con las curvas estimadas para cada distribución tomando muestras de tamaño 500 para las dos pruebas W y W^* (ver Anexos).

CAPÍTULO 4

APLICACIÓN A LOS RENDIMIENTOS DEL IPC

En este capítulo se consideran diferentes conjuntos de datos de rendimientos del IPC correspondientes a períodos importantes en la economía mexicana. En primera instancia, se realizó un análisis gráfico para después realizar un análisis estadístico utilizando algunos indicadores como el coeficiente de asimetría de Fisher y el coeficiente de curtosis de Pearson.

Para verificar el supuesto de independencia entre las observaciones se aplicó la prueba de Box-Pierce ([6]) utilizando la función `Box.test` de R. Para estimar los coeficientes de asimetría y curtosis, se utilizaron las funciones *skewness* y *kurtosis* del paquete `moments` de R ([25]). Note que en el caso del estimador de curtosis dicha función considera como referencia la curtosis de la distribución normal que es 3, pero para fines prácticos se tomará el exceso de curtosis, el cual se obtiene restando 3 al estimador anterior esto con el objeto de generar un coeficiente que valga 0 para la distribución Normal.

Posteriormente, para cada conjunto de datos se ajustaron las siguientes distribuciones, denotadas como sigue: t-asimétrica $ST(\xi, \omega, \alpha, \nu)$, Normal $N(\mu, \sigma)$, Normal asimétrica $SN(\xi, \omega, \alpha)$, logística $Logis(\mu, \sigma)$ y la Laplace $Lap(\mu, b)$. Para estimar los parámetros de la distribución t-asimétrica se utilizó la función *st.mpl* del paquete `sn` de R ([4]). Para estimar los parámetros de las distribuciones Normal y logística se utilizó la función *fitdistr* del paquete `MASS` ([38]) de R, y para los parámetros de la distribución Normal asimétrica se utilizó la función *sn.mpl* del paquete `sn`. Este ajuste se llevo a cabo con el fin de realizar una comparación entre la curva empírica y la curva ajustada, para poder determinar la posible distribución de los rendimientos económicos.

En última instancia, se aplicaron diferentes pruebas de bondad de ajuste para validar el análisis. Para probar la hipótesis de normalidad, se aplicó la prueba de Shapiro-Wilk descrita en la Subsección 2.8.2 y programada en la función *shapiro.test* de R. Para probar Normal asimétrica

se utilizó la prueba de Anderson-Darling estudiada por González-Estrada y Cosmes-Martínez ([15]). En este caso, la prueba utiliza el método bootstrap para estimar la distribución del estadístico.

Para probar distribución logística se utilizó la prueba de Anderson-Darling programada en la función *ad.test* del paquete *goftest* de R ([24]) y para la distribución Laplace se utilizó la prueba de Anderson-Darling propuesta por González-Estrada y Villaseñor A. ([16]) programada en la función *laplace.test* del paquete *goft* de R ([17]). Se aplicaron las pruebas propuestas W y W^* para probar la hipótesis de distribución t-asimétrica. Para estas dos últimas y para la prueba para normal asimétrica se utilizaron 500 muestras bootstrap para estimar la distribución del estadístico y calcular el valor de probabilidad.

Para fines prácticos las pruebas realizadas se denotaron de la siguiente forma:

BP: Box-Pierce para probar independencia en el tiempo.

SW: Shapiro-Wilk para normalidad.

AD_N: Anderson-Darling para Normal asimétrica.

AD_L: Anderson-Darling para logística.

AD_{LP}: Anderson-Darling para Laplace.

W: Anderson-Darling para t-asimétrica.

*W**: Anderson-Darling para t-Student.

4.1. Rendimientos semanales

4.1.1. Rendimientos 1995-1996

La crisis económica de 1994 fue una crisis iniciada en México, el mercado de cambios estuvo sujeto a múltiples perturbaciones provocadas por la falta de reservas internacionales, causando la devaluación del peso mexicano durante los primeros días de la presidencia de Ernesto Zedillo, a principios de 1995. La administración de Zedillo decidió establecer el sistema de libre flotación del peso, el cual llegaría a 7.20 pesos por dólar en tan sólo una semana. Cuando el dólar dejó de ser controlado por el gobierno el peso perdió la mitad de su valor, hecho que ocasionó que las deudas en dólares no pudieran ser pagadas ¹.

El conjunto de datos utilizado consta de 100 observaciones, el período que se consideró fue del

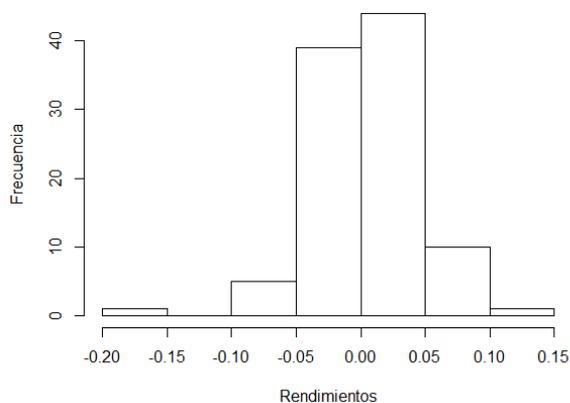
¹Fuente: Informe anual 1994 Banxico.

15/01/1995 al 08/12/1996, se calcularon los rendimientos semanales utilizando la ecuación (1.1). En la Figura 4.1 se presenta un histograma de frecuencias y la gráfica en el tiempo de los rendimientos. La prueba de Box-Pierce dio un valor de probabilidad de 0.1458 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis obtenidos fueron -0.1497 y 1.6702, respectivamente.

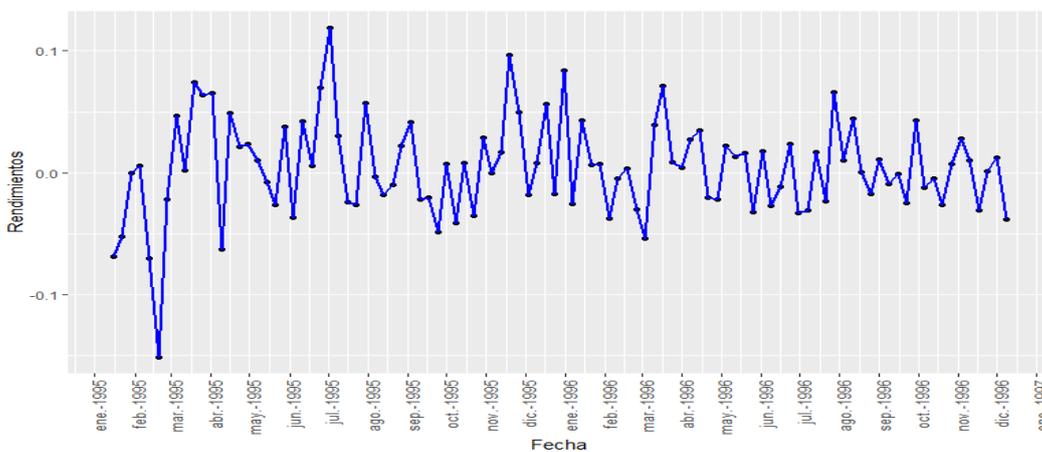
En la Figura 4.1 (a) se puede notar que el histograma de frecuencias presenta una ligera asimetría negativa, la cual provoca un alargamiento en la cola izquierda del histograma. Por otro lado, la curtosis resultó ser mayor a 0, lo cual indica que gran parte de las observaciones están concentradas alrededor de la media que tiene un valor de 0.004, en este caso se dice que la distribución es leptocúrtica.

En la Figura 4.1(b) se observa que los rendimientos en el período analizado tuvieron grandes variaciones en especial en las observaciones pertenecientes a los meses de Enero a Julio de 1995. El rendimiento máximo fue de 0.1188937, obtenido el 02/07/1995, mientras que el mínimo fue de -0.1510978, obtenido el 19/02/1995.

Figura 4.1: Rendimientos semanales del IPC de Enero, 1995 a Diciembre, 1996.



(a) Histograma de frecuencias.



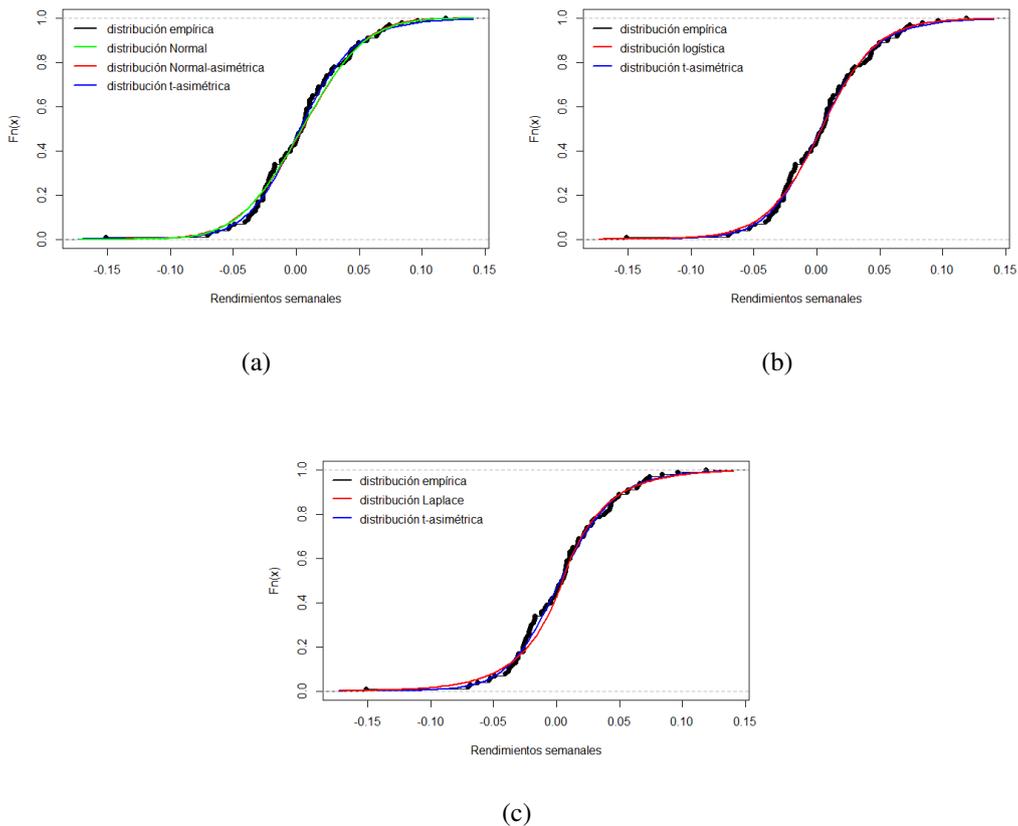
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0131, 0.0362, 0.6505, 6.2571)
$N(\mu, \sigma)$	(0.004, 0.0396)
$SN(\xi, \omega, \alpha)$	(0.0274, 0.0458, -0.8152)
$Logis(\mu, \sigma)$	(0.0036, 0.0217)
$Lap(\mu, b)$	(0.0047, 0.03)

Cuadro 4.1: Parámetros estimados de los rendimientos del IPC 1995-1996 utilizando diferentes distribuciones.

Figura 4.2: Curvas ajustadas para los rendimientos semanales del IPC de Enero, 1995 a Agosto, 1996.



La Figura 4.2 muestra la curva de la distribución empírica, la curva ajustada con la distribución t-asimétrica y la curva ajustada de las distribuciones estudiadas anteriormente; las estimaciones

de los parámetros correspondientes se pueden ver en el Cuadro 4.2.

Note que en la Figura 4.2(a) la curva justada Normal simétrica coincide con la normal asimétrica debido a la poca asimetría de los datos. En este caso, la curva ajustada con la distribución t-asimétrica parece ajustar mejor en el centro de la distribución empírica pero la diferencia con las otras curvas es mínima. Note además que en este caso los grados de libertad estimados son relativamente grandes. Debido a la Propiedad (2)(pág 23), el parecido con la curva de la distribución normal asimétrica se incrementa.

La Figura 4.2(b) muestra la comparación con la curva ajustada de la distribución logística, se puede notar que tanto la curva anterior como la generada con la distribución t-asimétrica tienen un buen ajuste y ligeramente esta última curva tiene un mejor ajuste en el centro de la distribución empírica. La Figura 4.2(c) muestra la comparación con la curva ajustada de la distribución Laplace. Note que la curva de la t-asimétrica tiene un mejor ajuste en gran parte del centro y la cola izquierda de la distribución empírica.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	2.1153	0.9742	0.5595	0.36133	1.0537	0.2693	0.2579
Valor-p	0.1458	0.04686	0.052	0.8858	0.105	0.36	0.622
Rechazar H_0	No	Sí	No	No	No	No	No

Cuadro 4.2: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Enero, 1995 a Diciembre, 1996.

El cuadro anterior muestra que los rendimientos considerados no tienen una distribución normal ya que se rechazó la hipótesis de normalidad usando la prueba de Shapiro-Wilk, pero debe notarse que la hipótesis de normal asimétrica no se rechazó aunque el valor de probabilidad obtenido fue cercano a 0.05. Por otra parte, las pruebas para la distribución Logística, Laplace y t-asimétrica no rechazaron las hipótesis correspondientes y la prueba con valor de probabilidad más alto fue la de Anderson-Darling para la distribución logística con un valor de 0.8858, lo cual indica que no se puede rechazar la hipótesis de que los datos siguen una distribución logística.

De manera similar sucede con los resultados de las pruebas W y W^* .

4.1.2. Rendimientos 2001-2002

En 2001 México resintió los efectos de la desaceleración económica global, particularmente de los Estados Unidos. Las exportaciones, la producción y el empleo se contrajeron. El Producto Interno Bruto sufrió una reducción de 0.3 por ciento. El debilitamiento de la actividad económica se reflejó en una pérdida significativa de empleos formales, especialmente en aquellos sectores más estrechamente ligados a la actividad exportadora, la economía mexicana se vio afectada por un entorno internacional particularmente complejo que se reflejó en los mercados financieros y propició caídas en los principales índices bursátiles. Así, en los países industriales prevaleció un ambiente poco propicio para la reactivación de la demanda ².

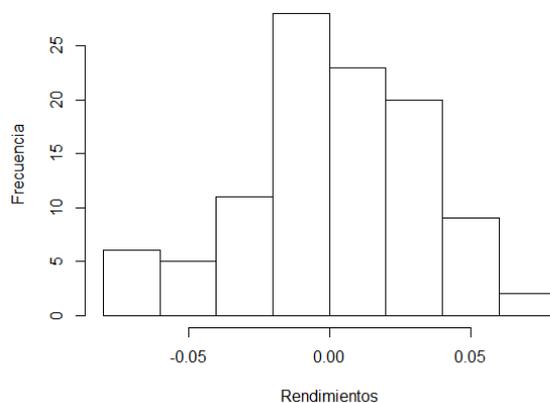
El conjunto de datos consta de 104 observaciones, el período considerado fue del 07/01/2001 al 29/12/2002. En este caso, con la prueba de Box-Pierce se obtuvo un valor-p de 0.2974, por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis fueron -0.3114 y 0.0120, respectivamente.

El histograma de frecuencias correspondiente a los datos se muestra en la Figura 4.3(a). Se puede ver que dicho histograma presenta una ligera asimetría y una curtosis pequeña lo cual indica que los datos están ligeramente concentrados hacia la media que tiene un valor de 0.002. Esto provoca que el histograma presente una forma achatada.

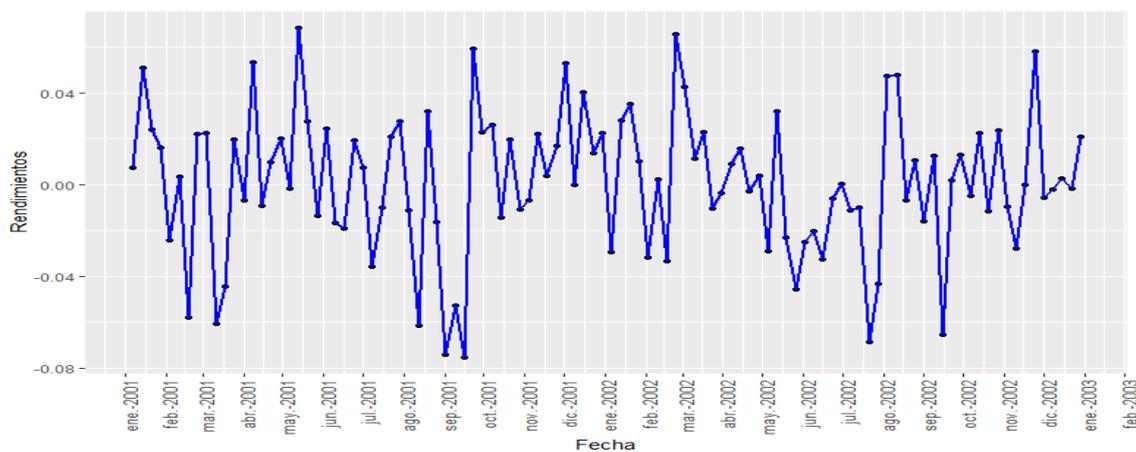
En la Figura 4.3(b) se observa que los rendimientos tuvieron grandes variaciones tanto positivas como negativas, de Septiembre a Octubre de 2001 se pueden ver dos valores negativos uno de -0.0741 correspondiente al 2/09/2001 y el otro de -0.0752 correspondiente al 16/09/2001. Este último valor fue consecuencia de la desaceleración económica a nivel mundial resultado de los atentados sufridos el 11 de Septiembre de 2001.

²Fuente: Informe anual 2001 Banxico.

Figura 4.3: Rendimientos semanales del IPC de Enero, 2001 a Diciembre, 2002.



(a) Histograma de frecuencias.



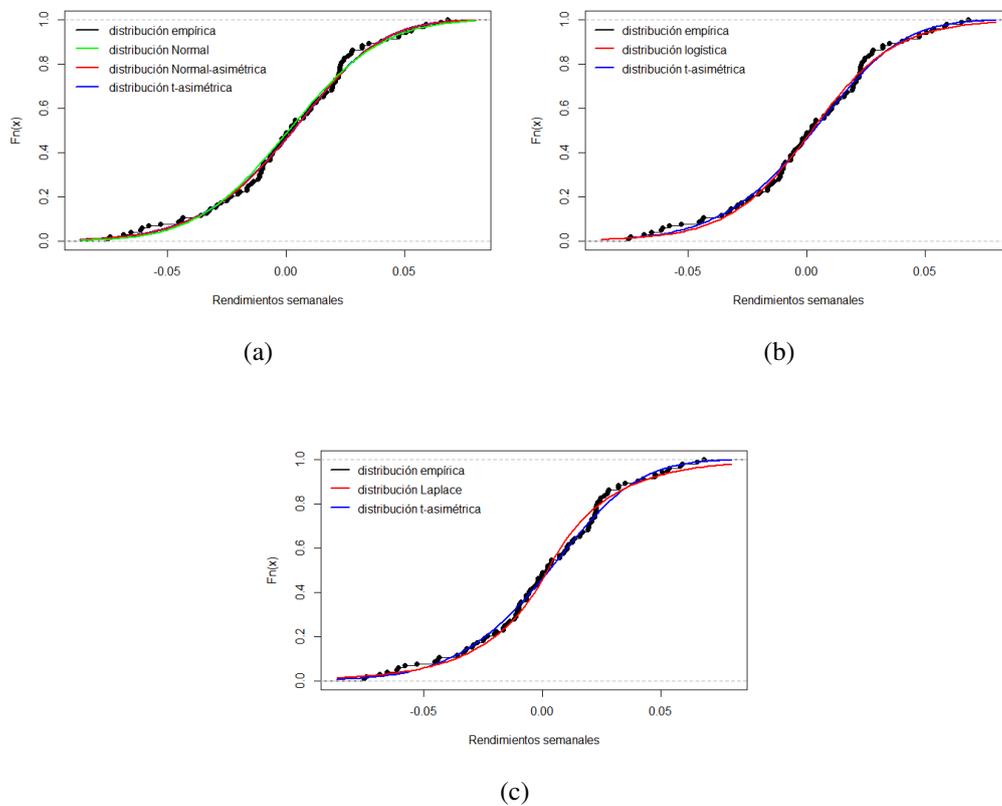
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(0.0294, 0.0419, -1.5929, 277680.2)
$N(\mu, \sigma)$	(0.0011, 0.0309)
$SN(\xi, \omega, \alpha)$	(0.0294, 0.0419, -1.5928)
$Logis(\mu, \sigma)$	(0.0021, 0.0174)
$Lap(\mu, b)$	(0.0022, 0.0241)

Cuadro 4.3: Parámetros estimados de los rendimientos del IPC 2001-2002 utilizando diferentes distribuciones.

Figura 4.4: Curvas ajustadas para los rendimientos semanales del IPC de Enero, 2001 a Diciembre, 2002.



La Figura 4.4(a) muestra la curva de la distribución empírica y las curvas ajustadas con la distribución Normal, Normal asimétrica y t-asimétrica. Note que el ajuste con dichas distribuciones es muy similar debido a la poca asimetría de los datos, en especial la curva de la normal asimétrica y la t-asimétrica no se diferencian ya que los grados de libertad estimados para esta última fueron muy grandes (ver Cuadro 4.3) y por la Propiedad 2(pág 23) la distribución t-asimétrica converge a la normal asimétrica.

La Figura 4.4(b) muestra la comparación de la curva empírica con la curva ajustada con la distribución logística, se puede notar que tanto la curva anterior como la generada con la distribución t-asimétrica tienen un buen ajuste y esta última curva tiene ligeramente un mejor ajuste en el centro y en las colas de la distribución empírica. Por otro lado, la Figura 4.4(c) muestra la comparación de la curva empírica con la curva ajustada con la distribución Laplace. Note que la curva de la distribución t-asimétrica tiene un mejor ajuste en gran parte del centro y en las colas de la distribución empírica.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	1.086	0.9811	0.3616	0.3780	1.2908	0.3616	0.2507
Valor-p	0.2974	0.1462	0.282	0.87	0.0532	0.2	0.648
Rechazar H_0	No	No	No	No	No	No	No

Cuadro 4.4: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Enero, 2001 a Diciembre, 2002.

El cuadro anterior muestra que no se rechazó la hipótesis de normalidad de la prueba de Shapiro-Wilk y la prueba de Anderson-Darling para la normal asimétrica, aunque el valor de probabilidad obtenido en ambos casos fue relativamente bajo, similar a lo que sucede con las pruebas de Anderson-Darling para la Laplace y la t-asimétrica. Note que el valor de probabilidad más alto fue el obtenido con la prueba de Anderson-Darling para la distribución logística.

4.1.3. Rendimientos 2008-2009

A lo largo de 2008, el peso mexicano se vio sometido a sucesivas devaluaciones que le llevaron a perder alrededor del 50 % de su valor frente al dólar estadounidense. Esto ocurrió en el marco de la crisis financiera mundial, que entre otras cosas implicó la quiebra de numerosas instituciones bancarias, las más importantes ocurrieron en Estados Unidos. En México los inversionistas comenzaron a adquirir grandes cantidades de dólares. Al aumentar la demanda de la divisa estadounidense, el peso se devaluó. En diciembre de 2008 la depreciación del peso mexicano alcanzó 26.70 %, con una cotización de 14 pesos por dólar, además según el INEGI, la tasa de desempleo alcanzó un nivel de 5 % enero de 2009 ([35]).

La turbulencia en los mercados financieros internacionales afectó significativamente el desempeño de la economía global. A pesar de la implementación de medidas de estímulo monetario y fiscal en un elevado número de países, la tasa de crecimiento del PIB mundial disminuyó de 5.2 por ciento en 2007 a 3.2 por ciento en 2008.

Al cierre de 2008, la inflación general anual resultó de 6.53 por ciento, ubicándose 2.77 puntos porcentuales por arriba de la cifra registrada en diciembre de 2007. La tendencia ascendente que exhibió esta variable fue resultado de varios eventos que afectaron a la oferta de bienes y servicios en el país. Al respecto, destacan los aumentos significativos que se presentaron durante el primer semestre en los precios internacionales de las materias primas alimenticias, metálicas y energéticas. Estas alzas tuvieron un impacto determinante en la estructura de costos de producción en prácticamente todos los sectores de la economía ³.

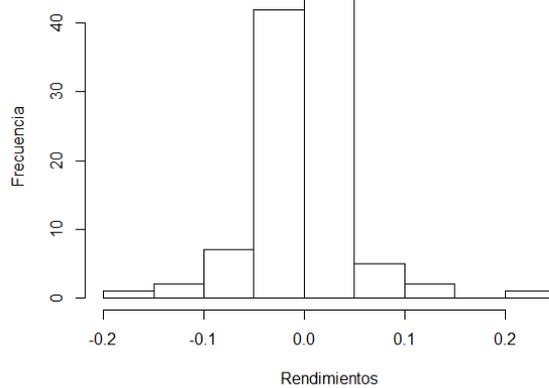
El conjunto de datos analizado consta de 104 observaciones, el período considerado fue del 06/01/2008 al 27/12/2009, en este caso con la prueba de Box-Pierce se obtuvo un valor de probabilidad de 0.068 por lo cual no se rechazó la hipótesis de independencia. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron 0.4364 y 3.5761, respectivamente.

En la Figura 4.5(a) se puede notar que el histograma de frecuencias para este período presenta una asimetría positiva la cual provoca un alargamiento en la cola derecha para valores superiores a la media, que tiene un valor de 0.0023. Por otro lado, la curtosis resultó ser grande, lo cual indica que los datos están muy concentrados hacia la media causando que la forma del histograma presente un pico más alto comparado con el de la distribución Normal.

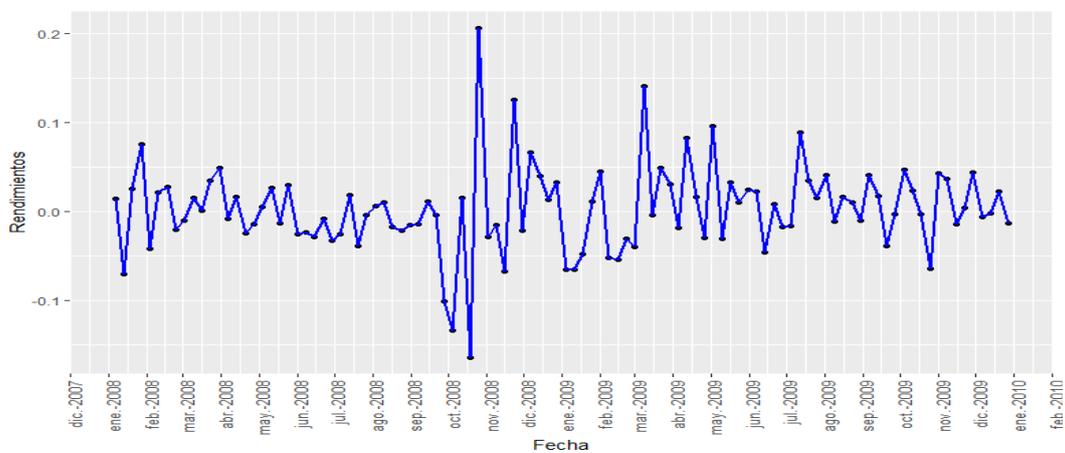
³Fuente: Informe anual 2008 Banxico.

La Figura 4.5(b) muestra que las mayores variaciones en los rendimientos para este período se tuvieron durante los meses de octubre a diciembre de 2008; dichas variaciones son producto de la crisis económica que se vivió en México.

Figura 4.5: Rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.



(a) Histograma de frecuencias.



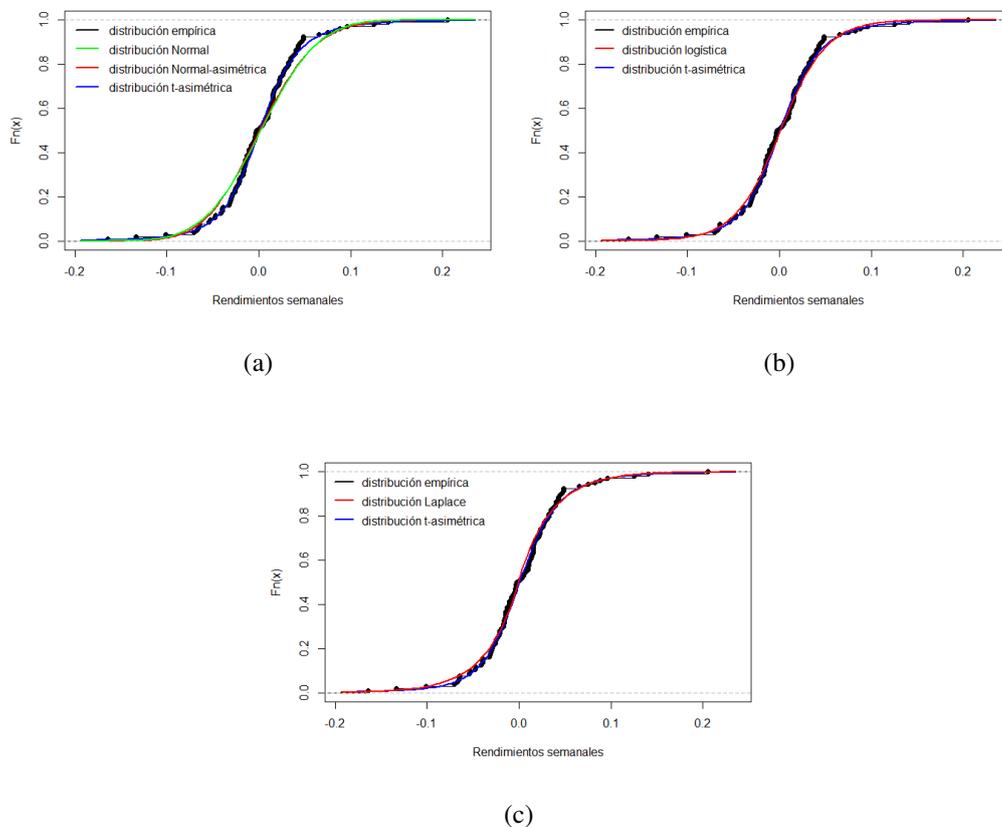
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0047, 0.0322, 0.2063, 3.1754)
$N(\mu, \sigma)$	(0.0023, 0.0490)
$SN(\xi, \omega, \alpha)$	(-0.0347, 0.0614, 1.1852)
$Logis(\mu, \sigma)$	(0.0013, 0.0248)
$Lap(\mu, b)$	(-0.0006, 0.0346)

Cuadro 4.5: Parámetros estimados de los rendimientos del IPC 2008-2009 utilizando diferentes distribuciones.

Figura 4.6: Curvas ajustadas para los rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.



La Figura 4.6(a) muestra la curva de la distribución empírica y la curva ajustada con las distribuciones Normal, Normal asimétrica y t-asimétrica, se puede ver que el ajuste con la distribución Normal tanto simétrica como asimétrica es muy similar debido a la poca asimetría que presenta la estimación de este parámetro, pero debe notarse que el ajuste con la distribución t-asimétrica es mejor en gran parte del centro de la distribución.

La Figura 4.6(b) muestra las curvas ajustadas con la distribución logística y con la t-asimétrica, en este caso el ajuste es muy parecido con ambas distribuciones, pero en gran parte del centro de la curva empírica esta última distribución presenta un mejor ajuste. Por otra parte, en la Figura 4.6(c) se puede notar que el ajuste con la distribución t-asimétrica es mejor en parte del centro de la distribución empírica, esto comparado con la curva ajustada con la distribución Laplace, la cual se aleja considerablemente de la cola izquierda de la curva empírica.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	3.3302	0.9366	1.4081	0.4298	1.6155	0.1965	0.4001
Valor-p	0.0680	8.7e-5	0	0.8183	0.0214	0.618	0.228
Rechazar H_0	No	Sí	Sí	No	Sí	No	No

Cuadro 4.6: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Enero, 2008 a Diciembre, 2009.

En el cuadro anterior se puede ver que se rechazan las hipótesis nulas correspondientes para las pruebas de la distribución normal y la normal asimétrica, con un valor de probabilidad cercano a 0 para la prueba de Shapiro-Wilk y con un valor de 0 para la prueba de Anderson-Darling (AD_N). De forma similar se rechaza la hipótesis nula de la prueba de Anderson-Darling para la distribución Laplace con un valor de probabilidad de 0.0214. La prueba de Anderson-Darling para la distribución logística dio un valor de probabilidad de 0.8183, el cual resultó ser el más alto seguido del obtenido mediante la prueba W , es decir, se puede considerar que en este caso la distribución logística y la t-asimétrica son modelos que ajustan de mejor forma los datos.

4.1.4. Rendimientos 2019-2020

En el tercer trimestre de 2019 la actividad económica de México mostró un estancamiento, luego de registrar ligeras caídas consecutivas entre el cuarto trimestre de 2018 y el segundo de 2019. Dicho comportamiento reflejó una pérdida de dinamismo de la demanda agregada, en un contexto en el que prevaleció la incertidumbre asociada tanto a factores externos, como internos. En particular, se continuó observando debilidad en el consumo y un desempeño desfavorable en la inversión fija bruta, al tiempo que las exportaciones manufactureras presentaron una desaceleración respecto del trimestre previo ⁴.

Según la Secretaría del Trabajo federal del 13 de marzo al 6 de abril de 2020, se perdieron 347000 empleos formales y según informó el director general de Citibanamex, Manuel Romo. El banco estima que la tasa desempleo en México a fin de 2020 será del 7% (Infobae), es decir, alrededor de 1 millón 750 mil nuevos desempleados; igualmente, las asociaciones empresariales Canacindra y Concanaco Servitur estiman pérdidas de 2 millones de empleos.

La economía nacional ha resentido los efectos de la pandemia desde finales de febrero por el abrupto freno económico de China y Estados Unidos, que impactó el precio del dólar que se devaluó hasta un máximo de 30%, ya que costaba alrededor de los 19 pesos en febrero y llegó a costar hasta 26 pesos por dólar el 6 de abril de 2020.

Los sectores más dañados han sido, el manufacturero por la ralentización de las cadenas productivas, materias primas como el petróleo que han tenido una reducción de la demanda por el encierro a nivel internacional, y el freno a la actividad económica ha afectado de forma severa en el sector servicios y turismo, complicando la vida de micro, pequeñas y medianas empresas ⁵.

El conjunto de datos que se consideró para esta sección se amplió con algunas observaciones recientes, el período considerado fue del 21/04/2019 al 16/08/2020 con 70 observaciones, esto debido a que la economía mexicana se vio amenazada por diversas causas como son la crisis sanitaria del coronavirus Covid-19, la caída en las bolsas de valores y del precio del petróleo a nivel global, el alza en el precio del dólar, etc.

Se aplicó la prueba de Box-Pierce obteniendo un valor de probabilidad de 0.9289 por lo cual no se rechaza la hipótesis de que los datos son independientes. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron 0.3403 y 1.3927,

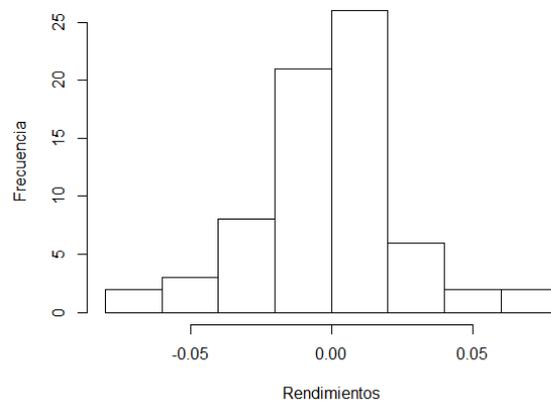
⁴Fuente: Informes trimestrales 2019 Banxico.

⁵Fuente: Infobae.com

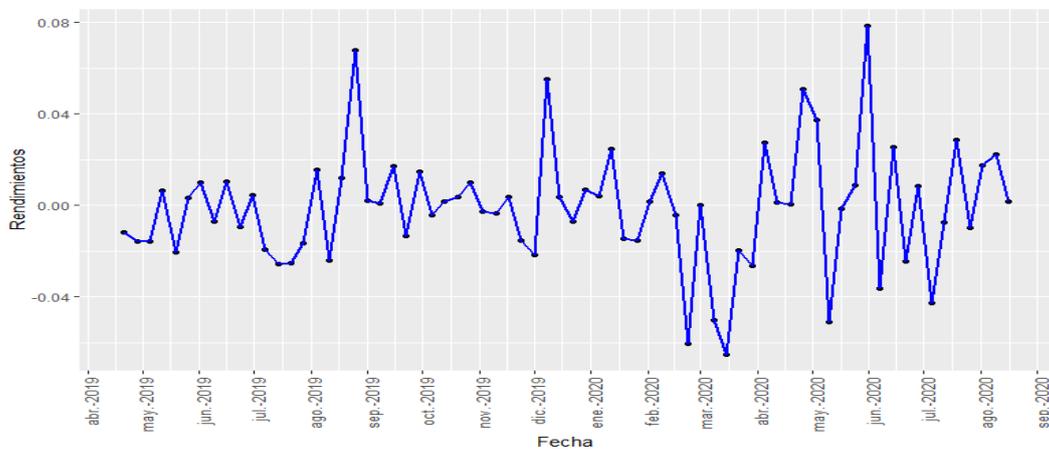
respectivamente.

La Figura 4.7 a) y b) muestra el histograma de frecuencias y la gráfica en el tiempo correspondiente.

Figura 4.7: Rendimientos semanales del IPC de Abril, 2019 a Agosto, 2020.



(a) Histograma de frecuencias.



(b) Gráfica en el tiempo.

La Figura 4.7(a) muestra que el histograma de frecuencias presenta ligera asimetría positiva y debido a que el valor del exceso de curtosis resultó mayor a 0 se puede ver que los datos están

concentrados alrededor de la media que tiene un valor de -0.00125 provocando una distribución con un pico más alto que el de la distribución Normal.

Por otra parte la Figura 4.7(b) muestra la variación en el tiempo de los rendimientos; se puede ver que a partir de agosto de 2019 la variación aumentó de manera significativa tanto de forma positiva como negativa y de febrero a mayo de 2020 se obtuvieron los rendimientos más bajos. Este período coincide con el inicio de la pandemia por el Covid-19.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

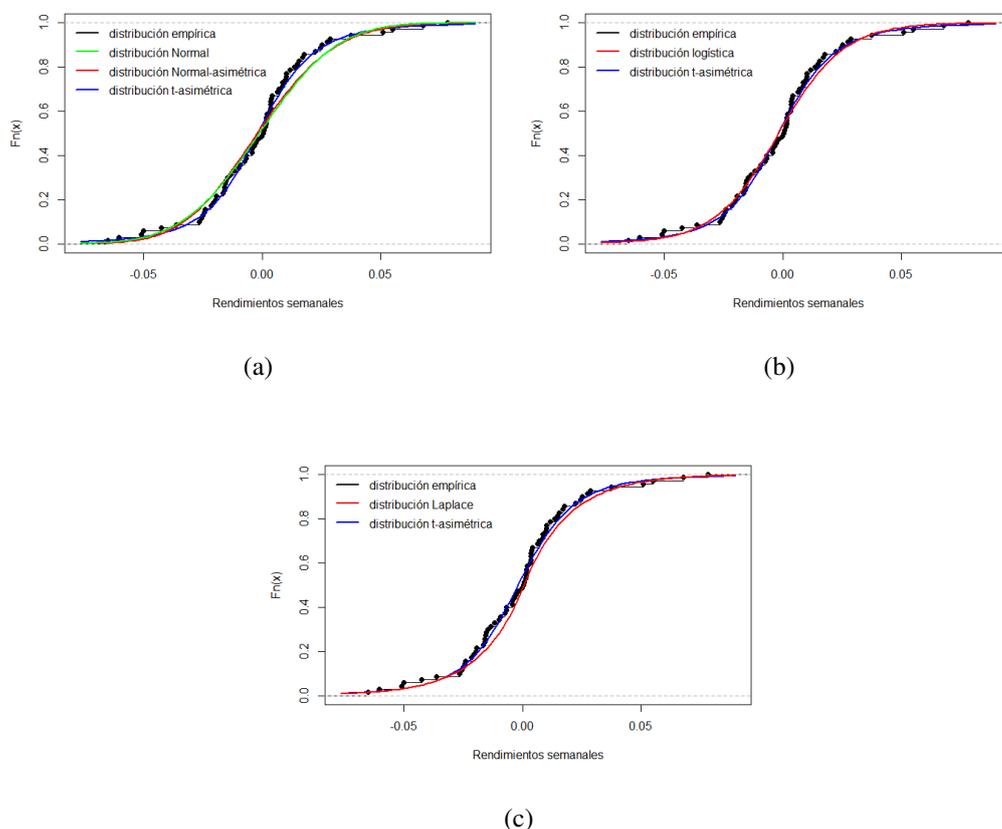
Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(0.0243, 0.0375, -0.9833, 4.90202)
$N(\mu, \sigma)$	(-0.0011, 0.0415)
$SN(\xi, \omega, \alpha)$	(0.0414, 0.0594, -2.2260)
$Logis(\mu, \sigma)$	(0.0014, 0.0218)
$Lap(\mu, b)$	(0.0021, 0.0299)

Cuadro 4.7: Parámetros estimados de los rendimientos del IPC 2019-2020 utilizando diferentes distribuciones.

La Figura 4.8(a) muestra la curva de la distribución empírica y la curva ajustada con las distribuciones Normal, Normal asimétrica y t-asimétrica, se puede ver que el ajuste con la distribución Normal tanto simétrica como asimétrica es muy similar, pero la curva ajustada con la t-asimétrica tiene un mejor ajuste en el centro de la distribución.

En la Figura 4.8(b) se puede ver que tanto la curva logística como la t-asimétrica tienen un buen ajuste en casi toda la curva de la distribución empírica. Por otra parte, en la Figura 4.8(c) se puede ver que la distribución Laplace no es un buen modelo para ajustar los datos pues en gran parte del centro de la distribución empírica el ajuste es mejor con la curva t-asimétrica.

Figura 4.8: Curvas ajustadas para los rendimientos semanales del IPC de Abril, 2019 a Agosto, 2020.



El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	0.0079	0.96	0.8788	0.4226	0.3365	0.2299	0.1574
Valor-p	0.9289	0.0251	0.0711	0.8255	0.7619	0.45	0.934
Rechazar H_0	No	Sí	No	No	No	No	No

Cuadro 4.8: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales del IPC de Abril, 2019 a Agosto, 2020.

El Cuadro 4.8 muestra que se rechaza la hipótesis de normalidad con un valor de probabilidad

de 0.0251, pero no se rechaza la hipótesis nula de la prueba de Anderson-Darling para la normal asimétrica, aunque debe notarse que el valor de probabilidad obtenido en este caso es relativamente pequeño (0.0711), contrario al valor de probabilidad obtenido para las demás pruebas de bondad de ajuste.

4.2. Rendimientos mensuales

4.2.1. Rendimientos 1993-1999

El siguiente análisis se realizó considerando los rendimientos mensuales durante el período que comprende del 01/07/1993 al 01/01/1999 con un total de 67 observaciones.

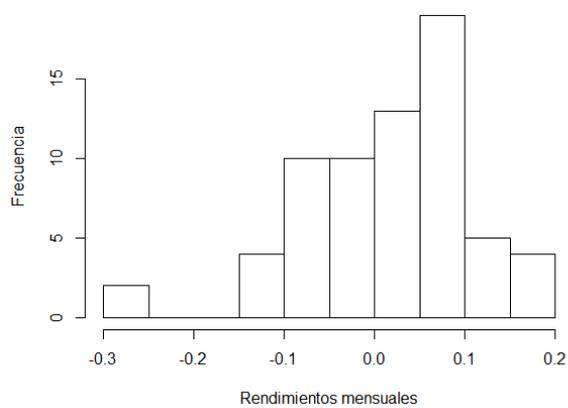
En este caso con la prueba de Box-Pierce se obtuvo un valor de probabilidad de 0.469 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron -0.7069 y 1.0291, respectivamente.

La Figura 4.9 muestra el histograma de frecuencias de los rendimientos mensuales en el período considerado y su respectiva gráfica en el tiempo.

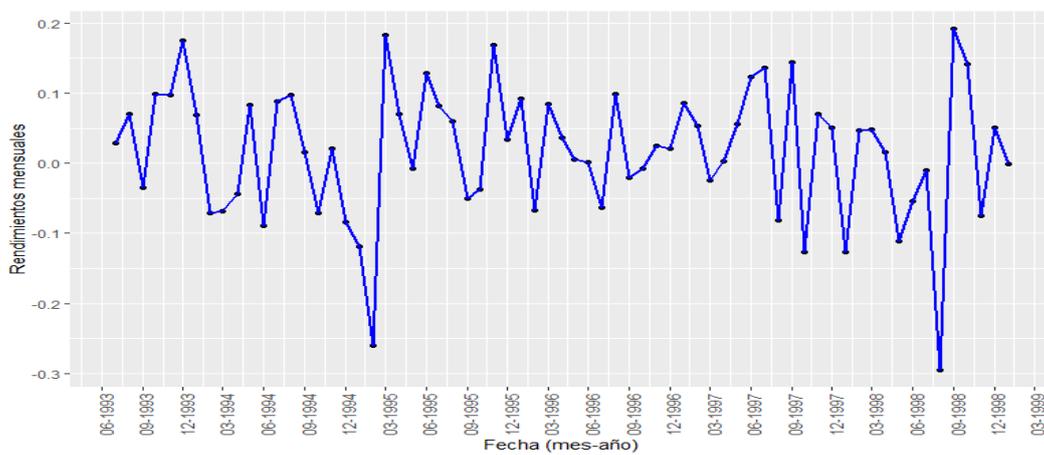
Se puede ver que el histograma presenta asimetría negativa, lo cual provoca un alargamiento en la cola izquierda para valores inferiores a la media que tiene un valor de 0.01729. Por otro lado, la Figura 4.9(b) muestra que las mayores variaciones de los rendimientos se obtuvieron durante febrero y marzo de 1995 y agosto y septiembre de 1998, esto debido a que durante 1995, la economía mexicana sufrió la crisis más severa ocurrida desde la década de los años treinta. La interrupción repentina de los flujos de capital del exterior hacia México a finales de 1994 e inicios de 1995, sumada a la consecuente devaluación de la moneda nacional, impusieron a la economía del país un ajuste doloroso pero inevitable ⁶.

⁶Fuente: Informe Anual 1995 Banxico.

Figura 4.9: Rendimientos mensuales del IPC de Julio, 1993 a Enero, 1999.



(a) Histograma de frecuencias.



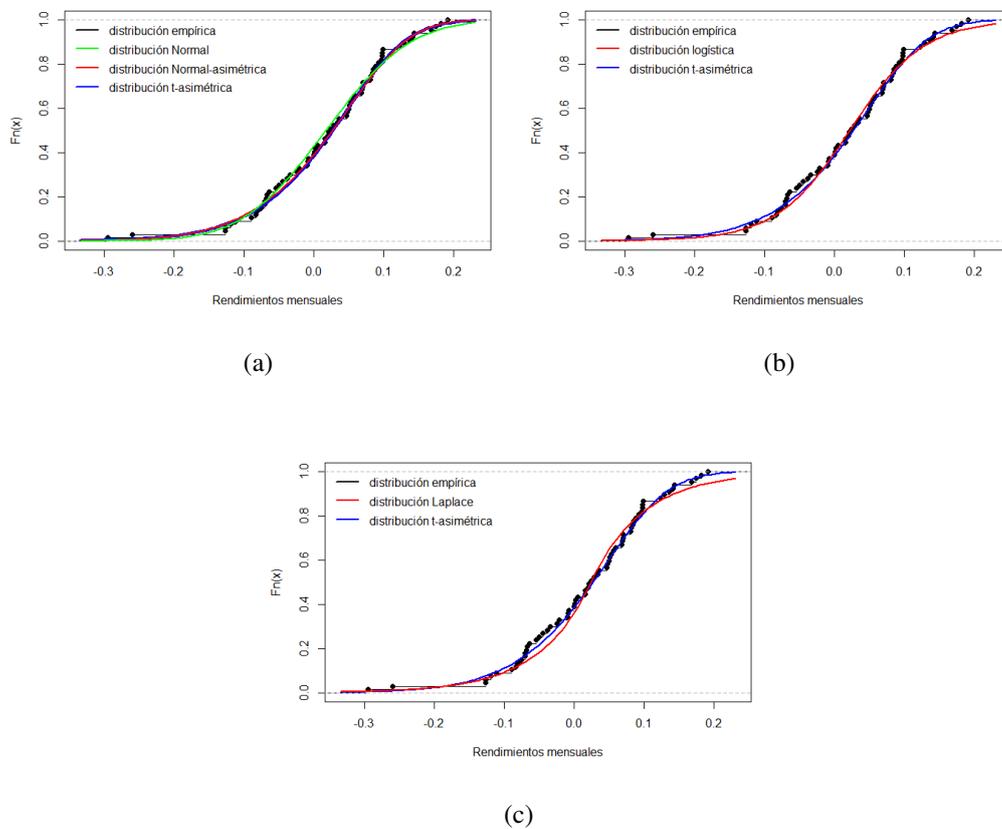
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(0.1142, 0.1301, -2.1383, 26.5126)
$N(\mu, \sigma)$	(0.0172, 0.0947)
$SN(\xi, \omega, \alpha)$	(0.1206, 0.1402, -2.4561)
$Logis(\mu, \sigma)$	(0.0220, 0.0527)
$Lap(\mu, b)$	(0.0243, 0.0739)

Cuadro 4.9: Parámetros estimados de los rendimientos mensuales del IPC 1993-1999 utilizando diferentes distribuciones.

Figura 4.10: Curvas ajustadas para los rendimientos mensuales del IPC de Julio, 1993 a Enero, 1999.



La Figura 4.10 muestra que la curva ajustada con las distribuciones Normal asimétrica y t-asimétrica es muy similar debido a que los grados de libertad estimados de la distribución t-asimétrica son considerablemente grandes (26.5). Además se puede ver que el ajuste con la distribución Normal no es bueno en parte del centro de la curva empírica y en la cola derecha esta.

En la Figura 4.10(b) se observa que en gran parte de la cola derecha de la distribución empírica, la curva t-asimétrica ajusta mejor que la curva logística. Por otra parte, en la Figura 4.10(c) se puede ver que la curva ajustada con la distribución Laplace tiene un mal ajuste tanto en el centro como en la cola derecha de la distribución empírica, contrario a lo que sucede con la curva de la distribución t-asimétrica, ya que ésta, junto con la curva de la distribución normal asimétrica, tienen un mejor ajuste tanto en las colas como en el centro de la curva empírica.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	0.5244	0.9626	0.2240	0.3321	1.4888	0.2090	0.6013
Valor-p	0.469	0.0421	0.684	0.9119	0.0304	0.67	0.084
Rechazar H_0	No	Sí	No	No	Sí	No	No

Cuadro 4.10: Resultados de las diferentes pruebas aplicadas a los rendimientos mensuales del IPC de Julio, 1993 a Enero, 1999.

El Cuadro 4.10 muestra que se rechazó la hipótesis de normalidad de los datos con un valor de probabilidad de 0.0421, de forma similar la prueba de Anderson-Darling para la distribución Laplace dio un valor de probabilidad de 0.0304 por lo cual se rechazó la hipótesis correspondiente, contrario a lo que sucede con las demás pruebas de bondad de ajuste ya que se obtuvieron valores de probabilidad mayores al nivel de significancia; hecho que no necesariamente lleva a considerar que los ajustes sean mejores, ya que por ejemplo, en la prueba W^* el valor de probabilidad obtenido fue de 0.08, mientras que el de la prueba para la distribución logística fue de 0.9119, es decir, la distribución logística parece tener un mejor ajuste.

4.2.2. Rendimientos 2015-2020

El siguiente análisis se realizó considerando los rendimientos mensuales durante el período que comprende del 01/01/2015 al 01/08/2020 con un total de 68 observaciones.

En este caso con la prueba de Box-Pierce se obtuvo un valor de probabilidad de 0.234 por lo cual no se rechazó la hipótesis de independencia. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron -1.0249 y 2.6561, respectivamente.

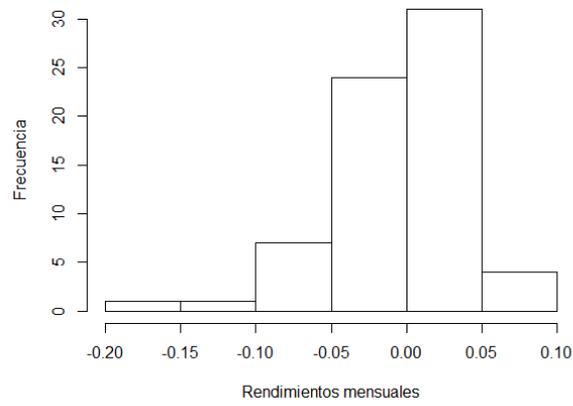
La Figura 4.11(a) muestra el histograma de frecuencias de los rendimientos mensuales en el período considerado. Se puede ver que presenta una notable asimetría negativa, lo cual provoca un alargamiento en la cola izquierda para valores inferiores a la media que tiene un valor de -0.0011. Por otro lado, la Figura 4.11(b) muestra que las mayores variaciones de los rendimientos se obtuvieron durante el 2018 y 2020, en especial en marzo de 2020 se obtuvo el menor rendimiento con un valor de -0.165.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

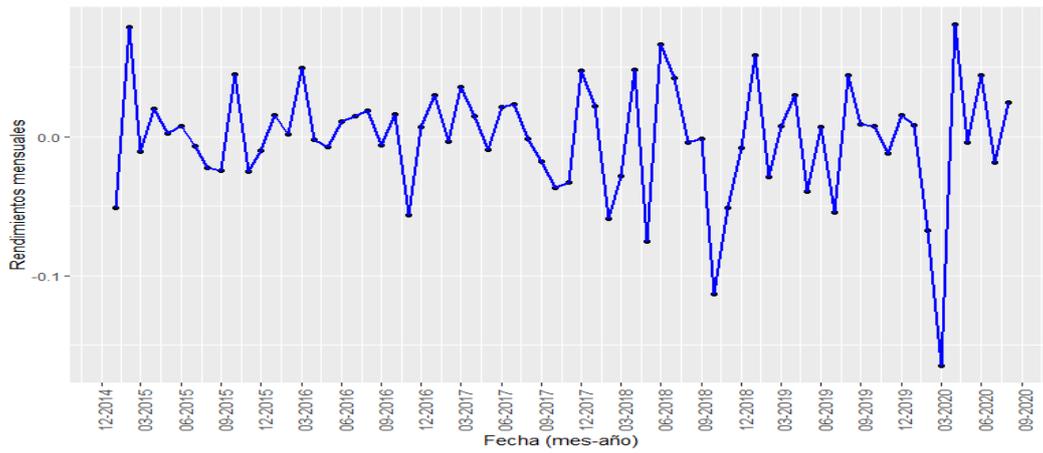
Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0029, 0.0176, 0.0774, 3.1948)
$N(\mu, \sigma)$	(-0.0012, 0.0257)
$SN(\xi, \omega, \alpha)$	(-0.0215, 0.0327, 1.2479)
$Logis(\mu, \sigma)$	(-0.0017, 0.0135)
$Lap(\mu, b)$	(0.0006, 0.0184)

Cuadro 4.11: Parámetros estimados de los rendimientos mensuales del IPC 2015-2020 utilizando diferentes distribuciones.

Figura 4.11: Rendimientos mensuales del IPC de Enero, 2015 a Agosto, 2020.

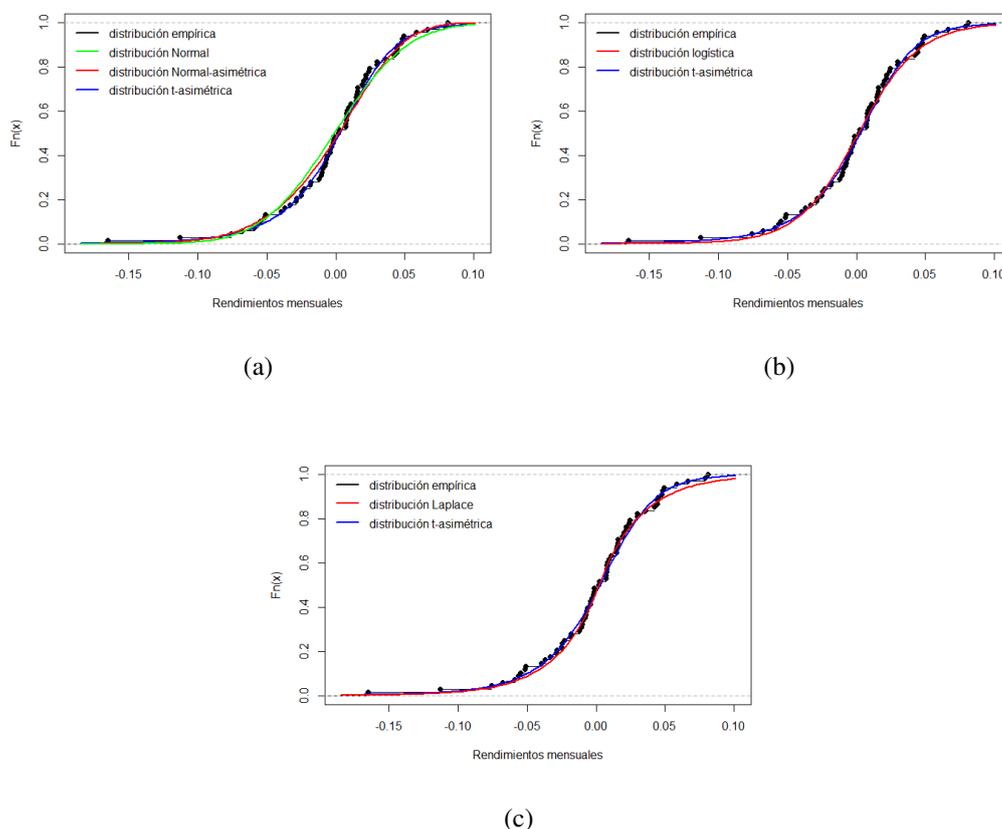


(a) Histograma de frecuencias.



(b) Gráfica en el tiempo.

Figura 4.12: Curvas ajustadas para los rendimientos mensuales del IPC de Enero, 2015 a Agosto, 2020.



La Figura 4.12 muestra la curva de la distribución empírica y la curva ajustada con las distribuciones Normal, Normal asimétrica y t-asimétrica. Se puede ver que el ajuste con la distribución Normal tanto simétrica como asimétrica es muy similar, pero la curva ajustada con la t-asimétrica tiene un mejor ajuste en el centro de la distribución empírica.

En la Figura 4.12(b) se puede ver que tanto la curva logística como la t-asimétrica tienen un ajuste similar en el centro de la curva empírica pero en gran parte de las colas, la curva t-asimétrica ajusta mejor. Por otra parte, en la Figura 4.12(c) se puede ver que la curva ajustada con la distribución Laplace tiene un mal ajuste en la cola derecha de la distribución empírica contrario a lo que sucede con la curva de la distribución la t-asimétrica.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	1.4109	0.9414	0.4917	0.3506	0.2734	0.2022	0.2557
Valor-p	0.2349	0.0031	0.1116	0.8955	0.8615	0.644	0.602
Rechazar H_0	No	Sí	No	No	No	No	No

Cuadro 4.12: Resultados de las diferentes pruebas aplicadas a los rendimientos mensuales del IPC de Enero, 2015 a Agosto, 2020.

El Cuadro 4.12 muestra que la prueba de Shapiro-Wilk se rechaza debido a que el valor de probabilidad obtenido fue cercano a 0. Note que la prueba de Anderson-Darling para la normal asimétrica no se rechaza, aunque el valor de probabilidad obtenido es relativamente pequeño, contrario a lo que sucede en las demás pruebas de bondad de ajuste. Por ejemplo, tanto la prueba de Anderson-Darling para la distribución logística como para la distribución Laplace dieron un valor de probabilidad alto, es decir, a pesar de que no se rechazaron las pruebas W y W^* se puede considerar que en este caso la distribución logística y la Laplace son modelos que ajustan correctamente los datos correspondientes.

4.3. Rendimientos para períodos largos

4.3.1. Rendimientos mensuales 2000-2020

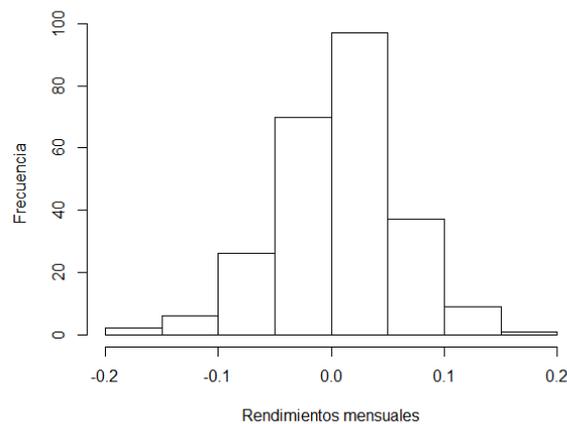
Debido a que el registro de los rendimientos mensuales contiene 12 observaciones por año es necesario considerar períodos largos para tener un número considerable de observaciones, en este caso es importante analizar lo que ha sucedido en el país en los últimos años, es por eso que se tomaron en cuenta los rendimientos mensuales que abarcan los períodos presidenciales de Vicente Fox Quesada (2000-2006), Felipe Calderón Hinojosa (2006-2012), Enrique Peña Nieto (2012-2018) y parte del período de Andrés Manuel López Obrador.

El conjunto de datos considerado contiene los rendimientos mensuales del 01/01/2000 al 01/08/2020 con un total de 248 observaciones.

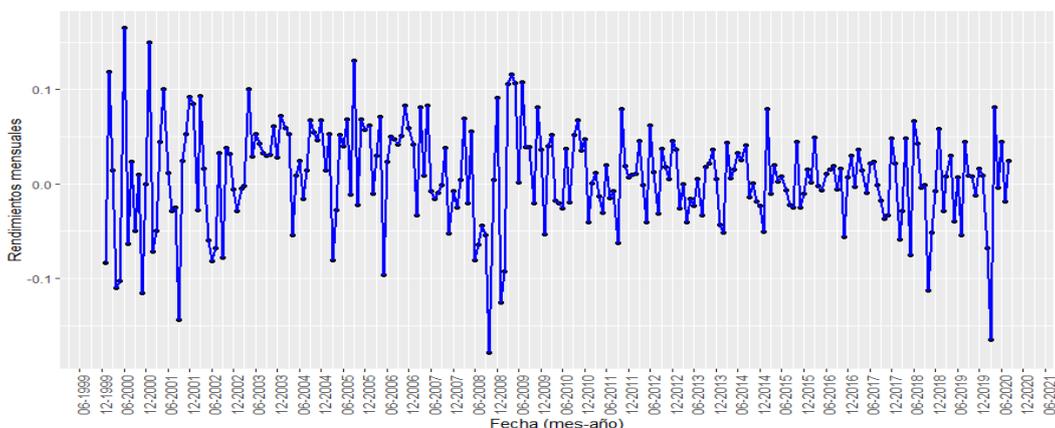
La prueba de Box-Pierce dio un valor de probabilidad de 0.5988 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron -0.3383 y 0.9311, respectivamente.

En la Figura 4.13(a) se observa que el histograma de frecuencias correspondiente a los rendimientos analizados, tiene una ligera asimetría en la parte izquierda y al presentar un exceso de curtosis pequeño tiene una forma poco achatada, ya que la mayor parte de las observaciones están concentradas alrededor de la media que fue de 0.008. Por otro lado, la Figura 4.13(b) muestra que en el período analizado los rendimientos mensuales presentaron grandes variaciones tanto positivas como negativas. Note que de junio de 2009 a diciembre de 2017 se observa que las variaciones no fueron tan grandes lo cual que hace que los rendimientos tomen valores en el intervalo $(-0.1, 0.1)$, contrario a lo que sucede de enero de 2000 a mayo de 2009 ya que en estos años se observa que los rendimientos presentaron las mayores variaciones, incluyendo el rendimiento máximo (0.1655) en junio de 2000 y el mínimo (-0.1788) en octubre de 2008.

Figura 4.13: Rendimientos mensuales del IPC de Enero, 2000 a Agosto, 2020.



(a) Histograma de frecuencias.



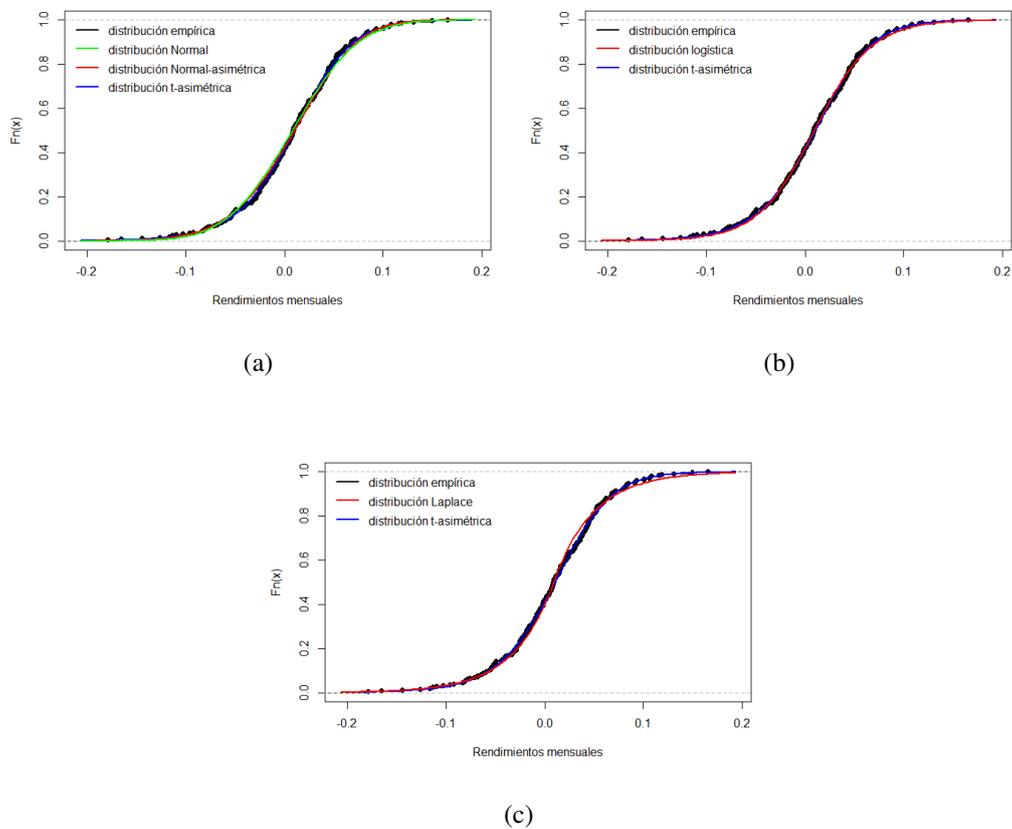
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(0.0352, 0.0512, -0.7483, 7.8594)
$N(\mu, \sigma)$	(0.0080, 0.0525)
$SN(\xi, \omega, \alpha)$	(0.0517, 0.0683, -1.3565)
$Logis(\mu, \sigma)$	(0.0094, 0.0289)
$Lap(\mu, b)$	(0.0089, 0.0400)

Cuadro 4.13: Parámetros estimados de los rendimientos mensuales del IPC 2000-2020 utilizando diferentes distribuciones.

Figura 4.14: Curvas ajustadas para los rendimientos mensuales del IPC de Enero, 2000 a Agosto, 2020.



Las curvas de la Figura 4.14(a) muestran el ajuste con las distribuciones Normal, Normal asimétrica y t-asimétrica. Note que el ajuste es similar para la distribución Normal y la Normal asimétrica, pero para el caso de la t-asimétrica la curva ajustada tiene ligeramente un mejor ajuste en el centro de la distribución empírica. El parecido entre la curva ajustada de la distribución Normal asimétrica y la t-asimétrica se debe a que los grados de libertad estimados para esta última fueron 7.859.

La curva ajustada con la distribución logística se muestra en la Figura 4.14(b), dicha curva tiene un ajuste muy similar al obtenido con la t-asimétrica, pero en parte de las colas de la distribución empírica, la curva con la t-asimétrica parece tener un mejor ajuste.

Por otro lado, en la Figura 4.14(c) se puede ver que el ajuste con la distribución Laplace se aleja de la curva empírica en la mayor parte del centro y la cola derecha.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	0.2768	0.9875	0.4381	0.2265	1.9458	0.1156	0.1233
Valor-p	0.5988	0.0297	0.166	0.9814	0.0086	0.974	0.99
Rechazar H_0	No	Sí	No	No	Sí	No	No

Cuadro 4.14: Resultados de las diferentes pruebas aplicadas a los rendimientos mensuales del IPC de Enero, 2000 a Agosto, 2020.

El cuadro anterior muestra que la hipótesis de normalidad se rechaza con un valor de probabilidad de 0.0297, aunque la hipótesis de la prueba de Anderson-Darling para la normal asimétrica no se rechazó a pesar de que el valor de probabilidad es pequeño (0.166). Por otra parte, la hipótesis de distribución Laplace se rechazó con un valor de probabilidad muy pequeño, contrario a lo que sucede para las pruebas en las que no se rechazó la hipótesis nula, por ejemplo, la prueba de Anderson-Darling para la logística y las pruebas W y W^* dieron valores de probabilidad cercanos a 1. Por lo cual se puede considerar que para el conjunto de datos analizado, las distribuciones logística y t-asimétrica son los modelos que mejor se ajustan.

CAPÍTULO 5

APLICACIÓN A LOS RENDIMIENTOS DE BONOS A 10 AÑOS

En este capítulo se analizan los rendimientos de los bonos a 10 años de México ¹ los cuales son un tipo de inversión basada en el sistema de deuda, mediante el cual se presta dinero al gobierno a cambio de una tasa de interés acordada. El gobierno los utiliza para obtener fondos que puedan ser empleados para nuevos proyectos de infraestructura y los inversores los pueden utilizar para obtener un retorno fijo en intervalos regulares. Cuando se compra un bono del estado, es como prestarle al gobierno una cantidad de dinero acordada durante un período de tiempo fijado. A cambio, el gobierno paga un porcentaje de interés fijo en periodos regulares, conocido como cupón, esto hace de los bonos un activo de renta fija.

5.1. Rendimientos semanales

Debido a que el registro de la información histórica de los bonos a 10 años de México se inició en 2009, entonces se tomarán períodos a partir de ese año.

El primer período 2009-2011 abarca parte de la crisis iniciada en 2008 y la drástica caída del (PIB) del segundo trimestre de 2009 la cual se atribuye al brote de influenza A H1N1 que sufrió el país en meses anteriores, situación que agudizó los problemas económicos que ya se venían presentando en México desde el último trimestre de 2008.

También este período abarca gran parte de la recuperación económica que México tuvo que enfrentar tras las crisis, en 2011 el PIB registró un incremento anual de 5.5 por ciento, cifra que se compara con la contracción de 6.1 por ciento observada el año anterior ².

El segundo período considerado fue de 2018-2020, éste abarca el inicio del gobierno de Andrés Manuel López Obrador y sucesos como la caída de 0.1 % del Producto Interno Bruto, primera caída en una década después de la sufrida en 2009. Además, tras la pandemia por el Covid-19

¹Fuente:Investing.com

²Fuente: Informe Anual 2010 Banxico.

en marzo y abril de 2020 se observaron fuertes disminuciones en el empleo del sector industrial y de servicios, considerando que en el mes de mayo el retroceso económico que México sufrió fue mucho mayor al vivido en la crisis de 1995.

5.1.1. Rendimientos 2009-2011

El conjunto de datos considerado contiene los rendimientos semanales de los bonos a 10 años del 05/04/2009 al 03/04/2011 con un total de 105 observaciones.

La prueba de Box-Pierce dio un valor de probabilidad de 0.8718 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron -0.6252 y 5.1733, respectivamente.

La Figura 5.1(a) muestra el histograma de frecuencias correspondiente, se puede observar que presenta asimetría negativa, lo que hace que la cola izquierda de dicho histograma sea ligeramente más alargada que la derecha. Por otra parte, la forma de dicho histograma presenta un pico alto, debido a que la mayor parte de las observaciones están cercanas a la media que tiene un valor de -0.003 y, por ende, el exceso de curtosis resultó ser bastante grande.

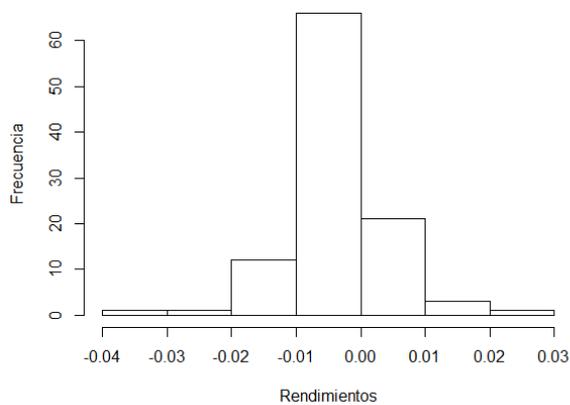
Por otra parte, la Figura 5.1(b) muestra las variaciones de los rendimientos en el período analizado. Note que las mayores variaciones se tuvieron de Agosto de 2010 a Febrero de 2011, en el resto del período se observa que la mayor parte de los rendimientos están en el intervalo (-0.02, 0.01).

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

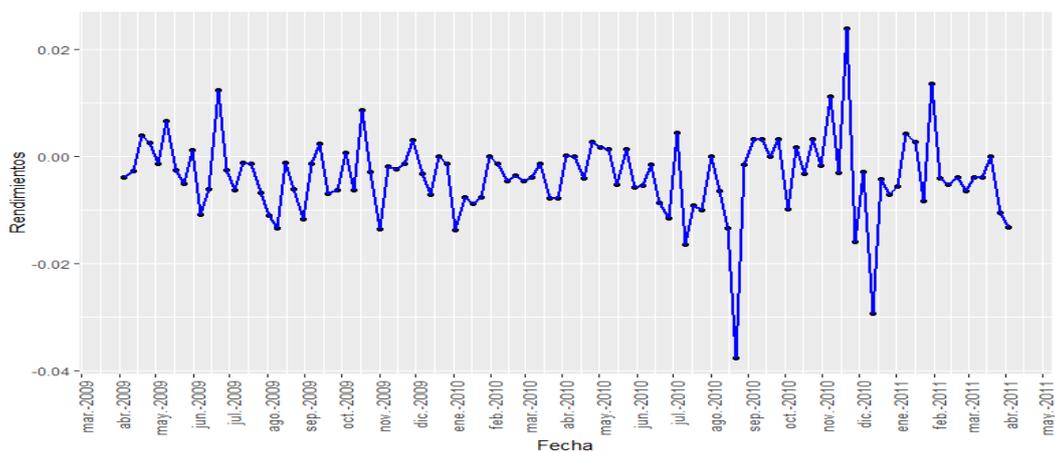
Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0024, 0.0047, -0.2298, 3.0358)
$N(\mu, \sigma)$	(-0.0035, 0.0075)
$SN(\xi, \omega, \alpha)$	(0.0021, 0.0094, -1.2220)
$Logis(\mu, \sigma)$	(-0.0034, 0.0038)
$Lap(\mu, b)$	(-0.0032, 0.0051)

Cuadro 5.1: Parámetros estimados de los rendimientos semanales de los bonos a 10 años 2009-2011 utilizando diferentes distribuciones.

Figura 5.1: Rendimientos semanales de los bonos a 10 años de Abril, 2009 a Abril, 2011.

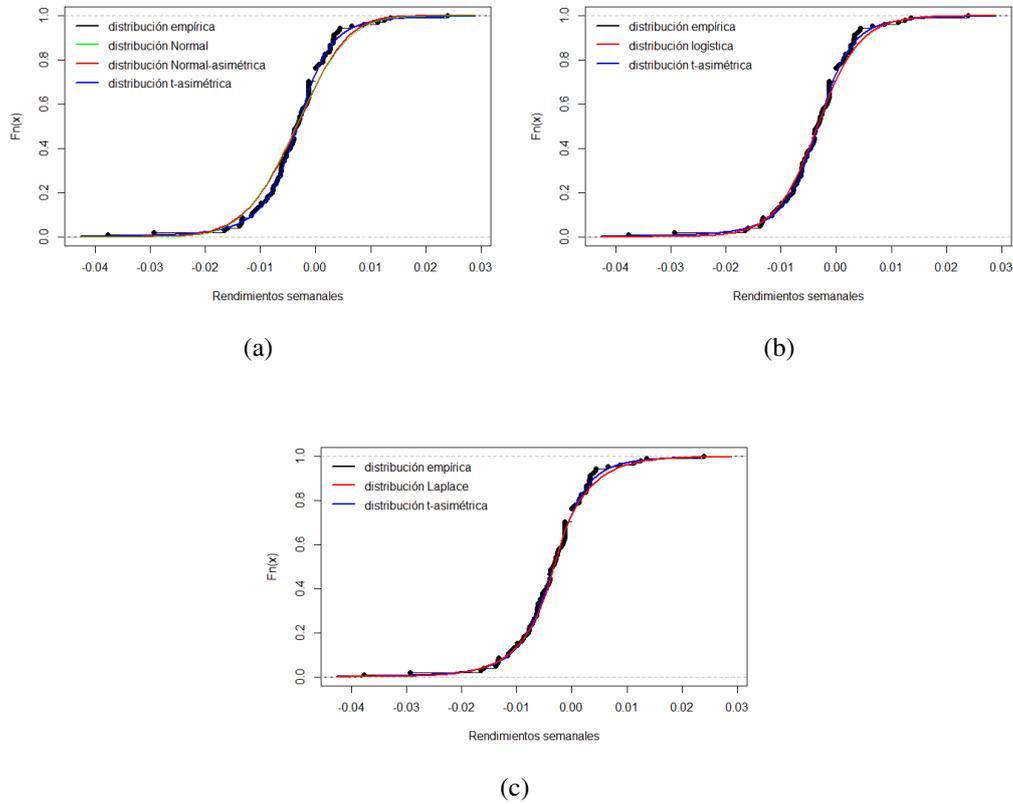


(a) Histograma de frecuencias.



(b) Gráfica en el tiempo.

Figura 5.2: Curvas ajustadas para los rendimientos semanales de los bonos a 10 años de Abril, 2009 a Abril, 2011.



En la Figura 5.2 se puede ver que las curvas ajustadas con la distribución Normal y Normal asimétrica son muy similares y presentan un mal ajuste en gran parte del centro de la distribución empírica, contrario a lo que sucede con las curvas ajustadas con la distribución logística y Laplace, ya que ambas presentan un ajuste muy similar al de la distribución t-asimétrica, pero esta última tiene ligeramente un mejor ajuste en parte del centro de la curva empírica.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	0.0260	0.9122	1.6624	0.5347	0.6880	0.1703	0.2250
Valor-p	0.8718	3.47e-06	0	0.7112	0.3015	0.734	0.716
Rechazar H_0	No	Sí	Sí	No	No	No	No

Cuadro 5.2: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de los bonos a 10 años de Abril, 2009 a Abril, 2011.

En el cuadro anterior se puede ver que el valor de probabilidad obtenido con la prueba de Shapiro-Wilk y con la prueba de Anderson-Darling para probar Normal asimétrica fue cercano a 0 en el primer caso, y en el segundo caso es 0, por lo que en ambos casos se rechazó la hipótesis de normalidad. Por otra parte, a pesar de que la distribución Laplace no se rechazó, debe notarse que el valor de probabilidad fue menor al obtenido con la prueba de Anderson-Darling para la logística y la t-asimétrica, por lo cual, se puede considerar que las distribuciones que mejor ajustan los rendimientos en este período son éstas dos últimas.

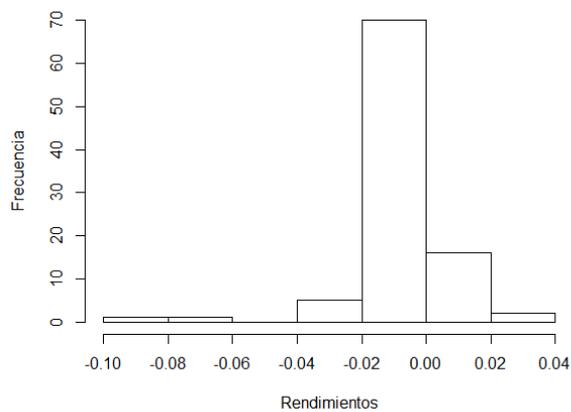
5.1.2. Rendimientos 2018-2020

El conjunto de datos considerado contiene los rendimientos semanales de los bonos a 10 años del 02/12/2018 al 20/09/2020 con un total de 95 observaciones.

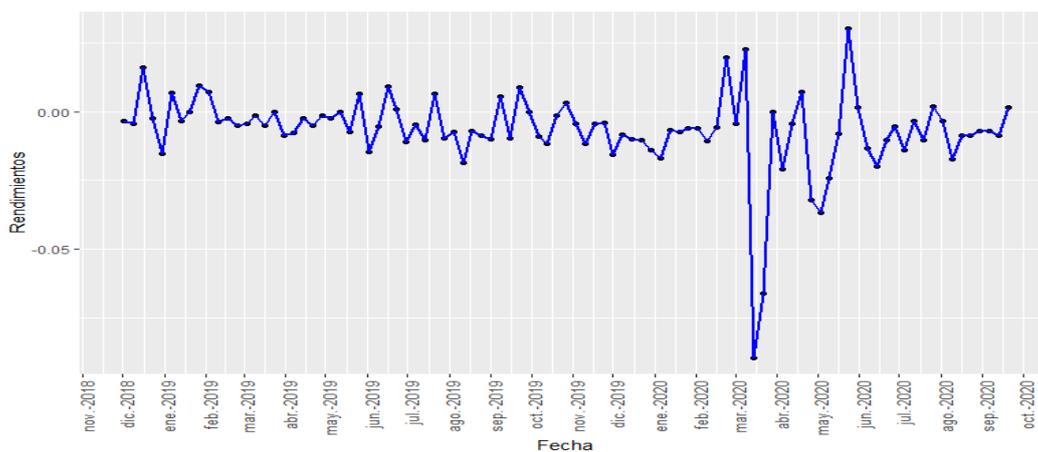
La prueba de Box-Pierce dio un valor de probabilidad de 0.05176 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron -2.5016 y 12.9798, respectivamente.

El histograma de frecuencias correspondiente se muestra en la Figura 5.3(a). Se puede ver que presenta una asimetría negativa lo cual hace que la cola izquierda de este sea más alargada que la cola derecha. Además, debido a que el exceso de curtosis es muy grande el histograma presenta un pico alto, ya que la mayor parte de los datos están concentrados alrededor de la media que tiene un valor de -0.006. Por otra parte, la Figura 5.3(b) muestra las variaciones de los rendimientos en el período analizado. Se puede ver que en general la mayor parte de los rendimientos fueron negativos y el mínimo se obtuvo en marzo de 2020.

Figura 5.3: Rendimientos semanales de los bonos a 10 años de Diciembre, 2018 a Septiembre, 2020.



(a) Histograma de frecuencias.



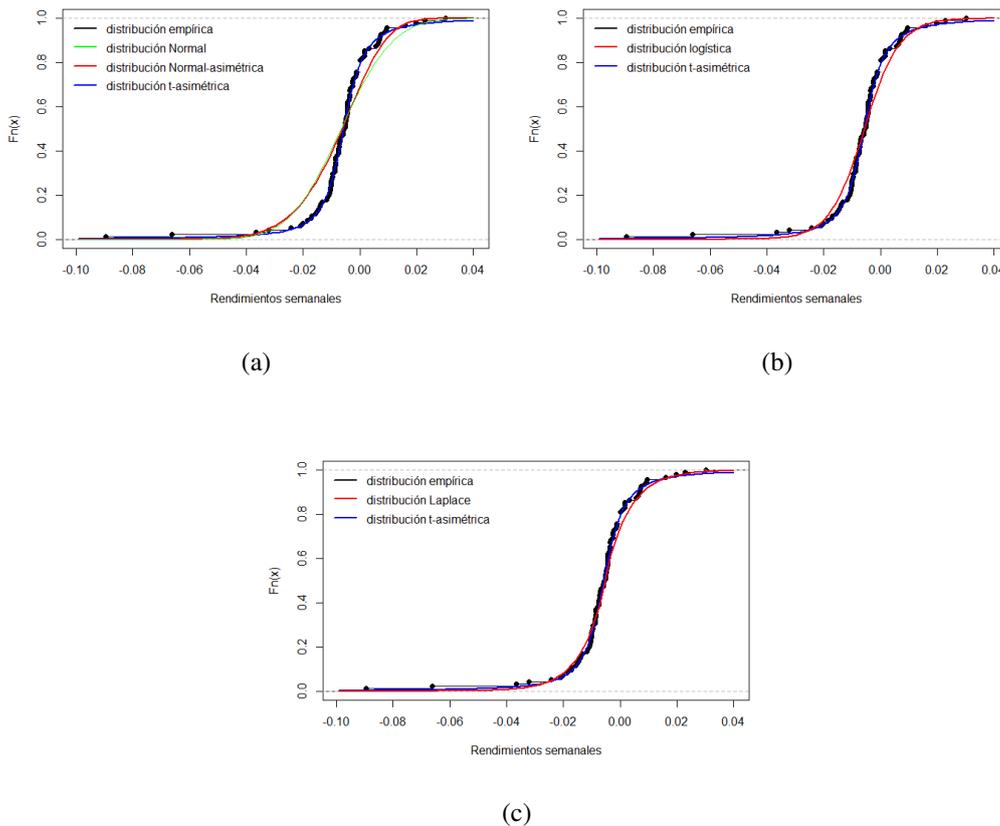
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0058, 0.0054, 0.0333, 1.7301)
$N(\mu, \sigma)$	(-0.0065, 0.0144)
$SN(\xi, \omega, \alpha)$	(0.0066, 0.0195, -2.0829)
$Logis(\mu, \sigma)$	(-0.0056, 0.0061)
$Lap(\mu, b)$	(-0.0052, 0.0081)

Cuadro 5.3: Parámetros estimados de los rendimientos semanales de los bonos a 10 años 2018-2020 utilizando diferentes distribuciones.

Figura 5.4: Curvas ajustadas para los rendimientos semanales de los bonos a 10 años de Diciembre, 2018 a Septiembre, 2020.



La Figura 5.4(a) muestra que las curvas ajustadas con la distribución Normal y Normal asimétrica difieren de la curva empírica en el centro, contrario a lo que sucede con la curva ajustada con

la t-asimétrica, ya que ésta se ajusta muy bien en toda la distribución empírica. En la Figura 5.4(b) se observa que la curva logística tiene un mal ajuste en gran parte del centro de la curva empírica y en este caso nuevamente la curva t-asimétrica tiene un mejor ajuste. Por otra parte, la Figura 5.4(c) muestra que la curva ajustada con la distribución Laplace tiene un ajuste similar al de la t-asimétrica, excepto en cierta parte del centro de la distribución.

El Cuadro 5.4 muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	3.7834	0.7612	4.3077	1.7223	1.7181	0.1755	0.3533
Valor-p	0.05176	3.995e-11	0.002	0.1314	0.0161	0.638	0.386
Rechazar H_0	No	Sí	Sí	No	Sí	No	No

Cuadro 5.4: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de los bonos a 10 años de Diciembre, 2018 a Septiembre, 2020.

Se puede ver que debido a que los valores de probabilidad obtenidos en las pruebas de Shapiro-Wilk y Anderson-Darling para Normal asimétrica fueron casi 0, entonces se rechazan las hipótesis de normalidad en ambos casos. De forma similar sucede con la prueba de Anderson-Darling para la distribución Laplace. Por otra parte, para la prueba W el valor de probabilidad resultó ser de 0.638, dicho valor es el más alto de las pruebas de bondad de ajuste en las que no se rechazó la hipótesis nula, es decir, se puede considerar que la distribución t-asimétrica provee un mejor ajuste en comparación con la distribución logística.

CAPÍTULO 6

OTRAS APLICACIONES

6.1. Rendimientos de las acciones de Wal-Mart México (WALMEX)

Wal-Mart de México es una de las cadenas comerciales más importantes de México. Opera 800 unidades comerciales distribuidas en 109 ciudades a nivel nacional, incluyendo tiendas de autoservicio, tiendas de ropa y restaurantes.

Es la segunda empresa más grande en valor de mercado bursátil, superando a TELMEX, además tiene un peso en el Índice de Precios y Cotizaciones del 11.49 %.

Maneja las tiendas de autoservicio SAM's CLUB, Bodega Aurrera, WalMart Super Center y Superama. También las tiendas de ropa Suburbia y las cadenas de restaurante VIP's y El Portón.

6.1.1. Rendimientos diarios 2020

El período considerado abarca el inicio de la pandemia por el Covid-19 en México, se tomó dicho período debido a que la cadena de autoservicios reportó el mayor incremento de ventas online de su historia y sus ingresos generales subieron 9.5 %, por el aumento de la demanda de bienes esenciales. En México, las ventas totales registraron un alza de 7.7 %, mientras que en las tiendas con más de un año en operación, las ventas aumentaron 6.3 %. Los productos más vendidos fueron juguetes, dispositivos electrónicos y alimentos.

El conjunto de datos contiene los rendimientos diarios de las acciones de Wal-Mart México del 26/03/2020 al 11/08/2020 con un total de 96 observaciones.

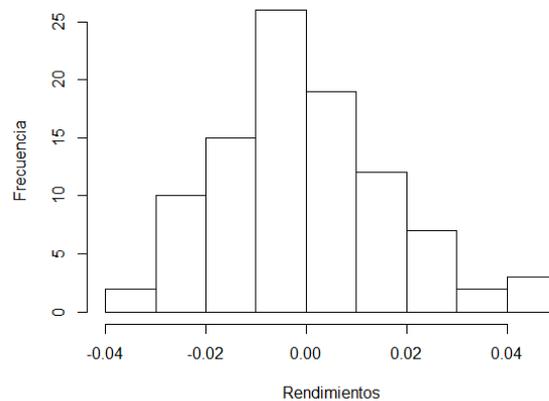
La prueba de Box-Pierce dio un valor de probabilidad de 0.2071 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otro lado, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron 0.3997 y -0.0603, respectivamente.

La Figura 6.1(a) muestra que el histograma de frecuencias correspondiente presenta un alargamiento

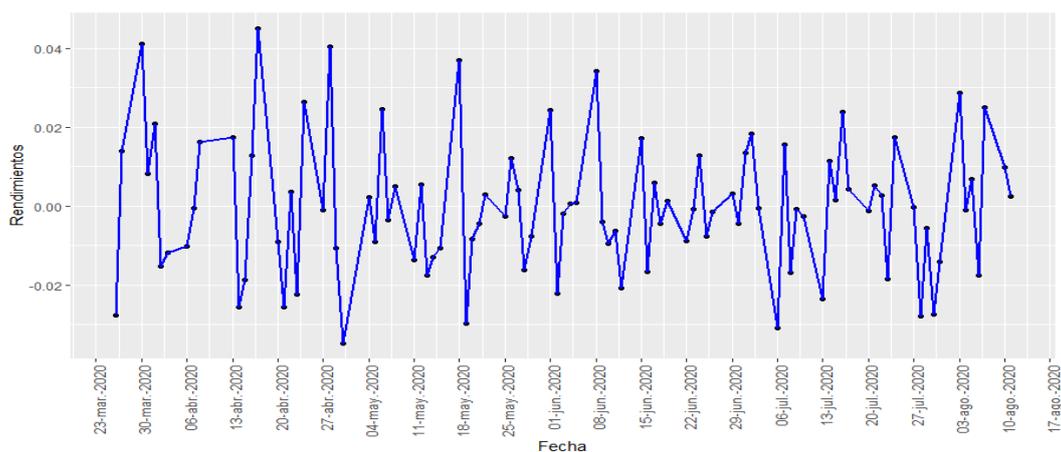
miento en la cola derecha y debido a que el exceso de curtosis resultó ser muy cercano a cero puede considerarse que la forma de éste es muy parecida a la de la distribución Normal.

En la Figura 6.1(b) se observan las variaciones en el tiempo de los rendimientos analizados, puede notarse que en general se tienen un número mayor de rendimientos negativos. Note que tanto el rendimiento máximo como el mínimo se obtuvieron en el mes de abril, debido a que en dicho mes las compras de pánico debido a que la pandemia vivida en México impactó la economía de manera significativa.

Figura 6.1: Rendimientos diarios de las acciones de Wal-Mart de Marzo a Agosto, 2020.



(a) Histograma de frecuencias.



(b) Gráfica en el tiempo.

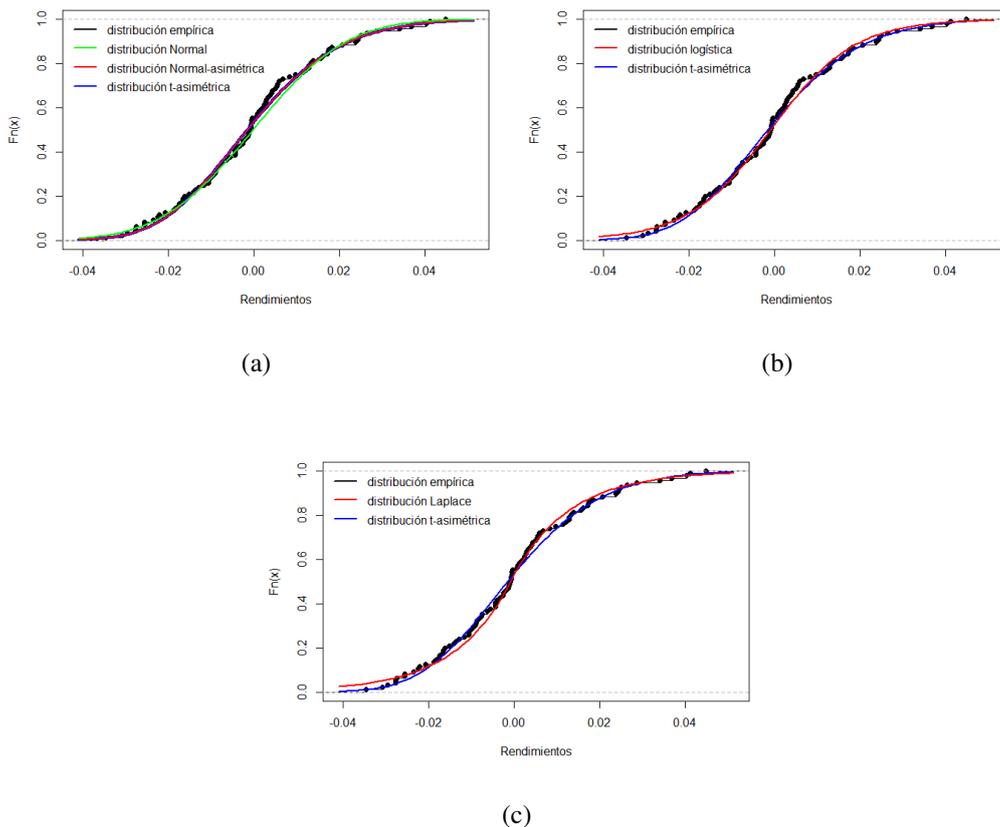
A continuación se presentan las estimaciones de los parámetros para cada distribución y su

respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0173, 0.0242, 1.9294, 9563874)
$N(\mu, \sigma)$	(-0.0002, 0.0170)
$SN(\xi, \omega, \alpha)$	(-0.0173, 0.0242, 1.9295)
$Logis(\mu, \sigma)$	(-0.0008, 0.0096)
$Lap(\mu, b)$	(-0.0008, 0.0131)

Cuadro 6.1: Parámetros estimados de los rendimientos diarios de Wal-Mart de Marzo a Agosto, 2020 utilizando diferentes distribuciones.

Figura 6.2: Curvas ajustadas para los rendimientos diarios de Wal-Mart de Marzo a Agosto, 2020.



El Cuadro 6.1 muestra que los grados de libertad estimados para la distribución t-asimétrica son considerablemente grandes y por la Propiedad (2)(pág 23) resulta que dicha distribución

converge a la Normal asimétrica y es por eso que las curvas de la Figura 6.2(a) son muy similares. Por otro lado, las curvas ajustadas con la distribución logística y la t-asimétrica dadas en la Figura 6.2(b) muestran que ambas curvas tienen un buen ajuste en gran parte de las colas de la distribución empírica, pero en el centro el ajuste parece no ser muy adecuado. De manera similar, la Figura 6.2(c) muestra el ajuste con la distribución Laplace, dicho ajuste es bueno en parte del centro de la curva empírica, pero en la cola izquierda la curva t-asimétrica tiene un mejor ajuste.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	1.5915	0.9809	0.2566	0.3402	0.7640	0.2566	0.2603
Valor-p	0.2071	0.1768	0.604	0.9049	0.2423	0.494	0.6320
Rechazar H_0	No	No	No	No	No	No	No

Cuadro 6.2: Resultados de las diferentes pruebas aplicadas a los rendimientos diarios de Wal-Mart de Marzo a Agosto, 2020.

En el cuadro anterior se observa que ninguna de las pruebas de bondad de ajuste se rechazaron. Pero debe notarse que para las pruebas de Shapiro-Wilk y Anderson-Darling para la distribución Laplace, los valores de probabilidad obtenidos fueron menores en comparación con los valores obtenidos para las demás pruebas. En este caso, puede considerarse que cualquiera de las distribuciones propuestas proporciona un buen ajuste a los datos.

6.1.2. Rendimientos semanales 2019-2020

El conjunto de datos analizado consta de 80 observaciones de los rendimientos semanales de las acciones de Wal-Mart de México del 17/03/2019 al 20/09/2020. Se consideró dicho período debido a que las ventas netas de Wal-Mart de México y Centroamérica se incrementaron 9.6 por ciento durante el segundo trimestre de 2020, al pasar de 153 mil 149 millones de pesos a 167 mil 875 millones, esto pese a que abril y mayo fueron los dos meses más complicados, pues gran parte de la sociedad permaneció en confinamiento por la pandemia de Covid-19.

No obstante, el aumento de las ventas de la minorista se vieron opacadas por un desplome de

81 por ciento anual al bajar de 8 mil 751 millones de pesos en el segundo trimestre de 2019 a mil 662 millones de pesos en igual periodo de 2020. Según revela el reporte de la empresa, es consecuencia de un mayor gasto financiero y de impuestos.

En mayo de 2020 Walmart pagó al Servicio de Administración Tributaria (SAT) 8 mil 79 millones de pesos para concluir asuntos fiscales sustanciales, los cuales, informó en su momento la minorista, se verían reflejados en sus estados financieros del segundo trimestre de este año¹.

La prueba de Box-Pierce dio un valor de probabilidad de 0.4359 por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron -0.2915 y -0.1852, respectivamente.

La Figura 6.3(a) muestra el histograma de frecuencias para los rendimientos analizados. La cola izquierda del histograma presenta un alargamiento que se ve reflejado en el valor negativo del estimador del coeficiente de asimetría, la forma achatada de este histograma es debido a que las clases con mayor frecuencia están distribuidas alrededor de las clases cercanas al cero. Por otra parte, en la Figura 6.3(b) se observa la variación en el tiempo de los rendimientos semanales en el período correspondiente. En general, no se observa una tendencia hacia ciertos valores pero se puede ver que el valor máximo (0.0606) y el mínimo (-0.0729) se obtuvieron durante la última semana de julio y la primer semana de agosto de 2020.

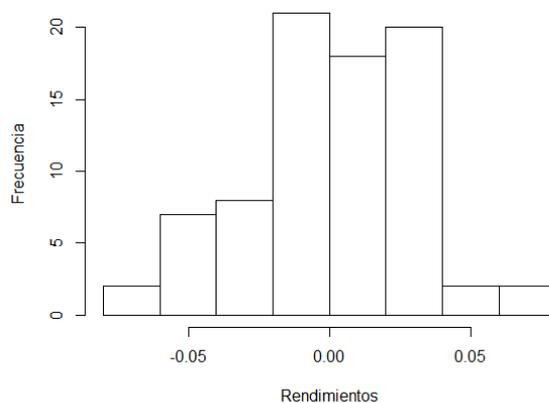
A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(0.0280, 0.0391, -1.7414, 2641919)
$N(\mu, \sigma)$	(0.0009, 0.0282)
$SN(\xi, \omega, \alpha)$	(0.0280, 0.0391, -1.7414)
$Logis(\mu, \sigma)$	(0.0019, 0.0161)
$Lap(\mu, b)$	(0.0045, 0.0228)

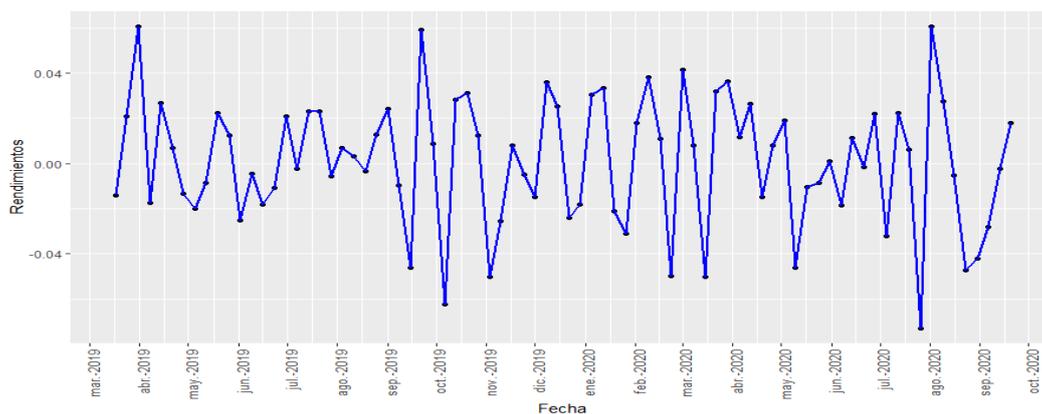
Cuadro 6.3: Parámetros estimados de los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020 utilizando diferentes distribuciones.

¹Fuente:La Jornada

Figura 6.3: Rendimientos semanales de las acciones de Wal-Mart de Marzo, 2019 a Septiembre, 2020.



(a) Histograma de frecuencias.



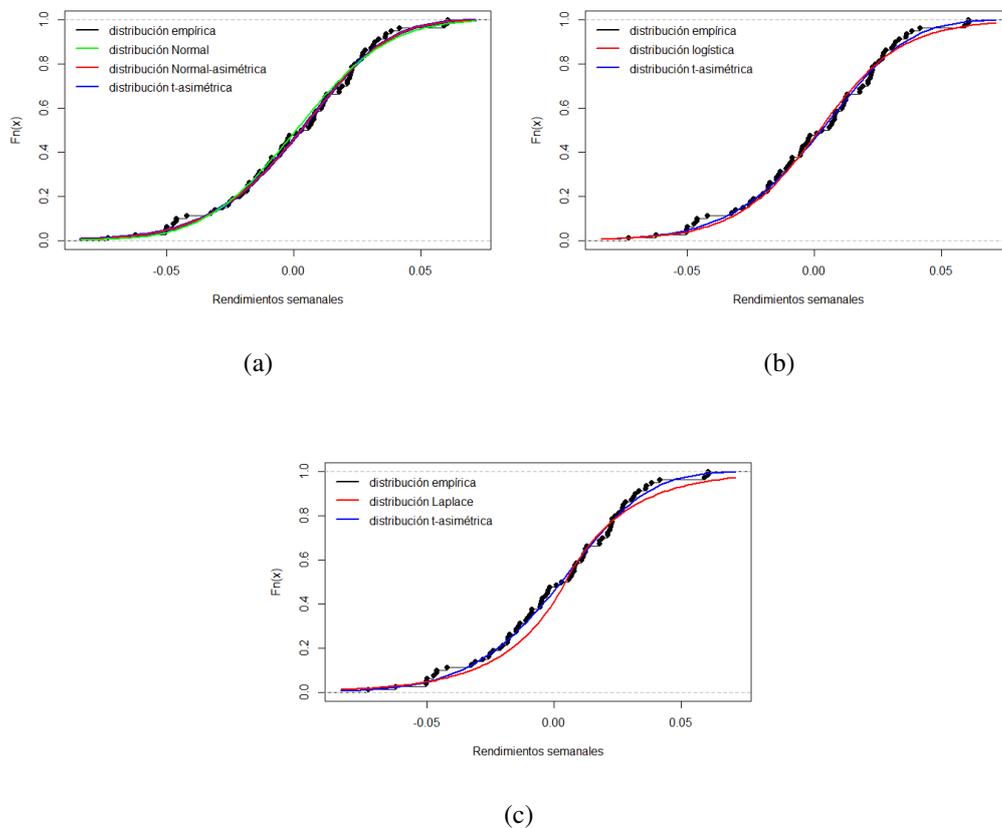
(b) Gráfica en el tiempo.

El Cuadro 6.3 muestra que los grados de libertad estimados para la distribución t-asimétrica resultaron ser bastante grandes por lo que las estimaciones de ésta distribución resultaron ser las mismas que las de la Normal asimétrica. Este hecho se ve reflejado en las curvas ajustadas en la Figura 6.4(a), aunque debe notarse que la curva ajustada con la distribución Normal se asemeja a las curvas anteriores en gran parte de las colas.

Por otra parte, la Figura 6.4(b) muestra que la curva ajustada con la distribución logística tiene un ajuste similar con el de la t-asimétrica en gran parte del centro de la curva empírica de los datos, pero en las colas de ésta, la curva con la t-asimétrica parece tener un mejor ajuste.

En la Figura 6.4(c) se puede notar que la curva ajustada con la distribución Laplace tiene un mal ajuste en casi toda la curva empírica, esto debido a que la distribución Laplace es una distribución con un pico alto y en este caso la forma de la distribución de los datos es bastante achatada.

Figura 6.4: Curvas ajustadas para los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020.



El Cuadro 6.4 muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos. Para cada prueba se utilizó un nivel de significancia de 0.05.

Se puede ver que la prueba de Anderson-Darling para la distribución Laplace dio un valor de probabilidad de 0.0019, por lo cual se rechazó la hipótesis correspondiente. Pero debe notarse que no se rechazó la hipótesis de normalidad tanto simétrica como asimétrica, aunque debe considerarse que los valores de probabilidad en estas pruebas no resultaron tan altos en comparación con el valor correspondiente a la prueba de Anderson-Darling para la distribución

logística. En este caso, se puede considerar que las distribuciones logística, la Normal asimétrica y t-asimétrica proveen un buen ajuste para los datos.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	0.6071	0.9828	0.2576	0.4354	2.4908	0.2576	0.3072
Valor-p	0.4359	0.3596	0.598	0.8125	0.0019	0.454	0.5160
Rechazar H_0	No	No	No	No	Sí	No	No

Cuadro 6.4: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020.

6.1.3. Rendimientos mensuales 2013-2020

El conjunto de datos consta de 85 observaciones correspondientes a los rendimientos mensuales de las acciones del Wal-Mart de septiembre de 2013 a septiembre de 2020.

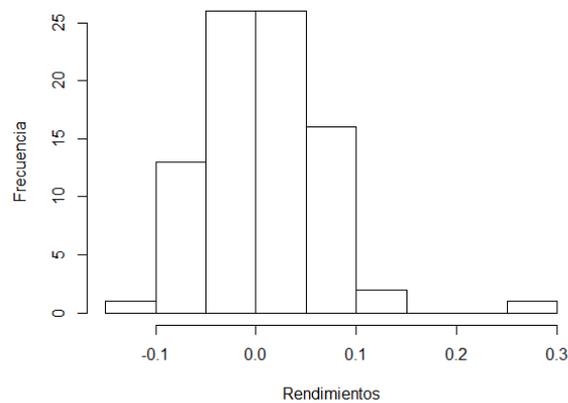
Se consideró este período debido a que en los últimos años la empresa ha incrementado tanto el número de tiendas como el número de empleados alrededor del país. Además de que en este período se tuvieron sucesos importantes que afectaron las ganancias de dicha empresa. Por ejemplo, en septiembre de 2013, Wal-Mart reportó una disminución de 4.7 por ciento. Mientras en diciembre de ese mismo año registró una baja de 1.6 por ciento, ni siquiera durante 2008, año en que estalló la crisis financiera internacional, o 2009, con la emergencia sanitaria originada por la gripe AH1N1 en México, hubo un resultado anual tan negativo.

La prueba de Box-Pierce dio un valor de probabilidad de 0.2838, por lo cual no se rechazó la hipótesis de independencia entre las observaciones. Por otra parte, las estimaciones de los coeficientes de asimetría y exceso de curtosis que se obtuvieron fueron 0.9901 y 3.2410, respectivamente.

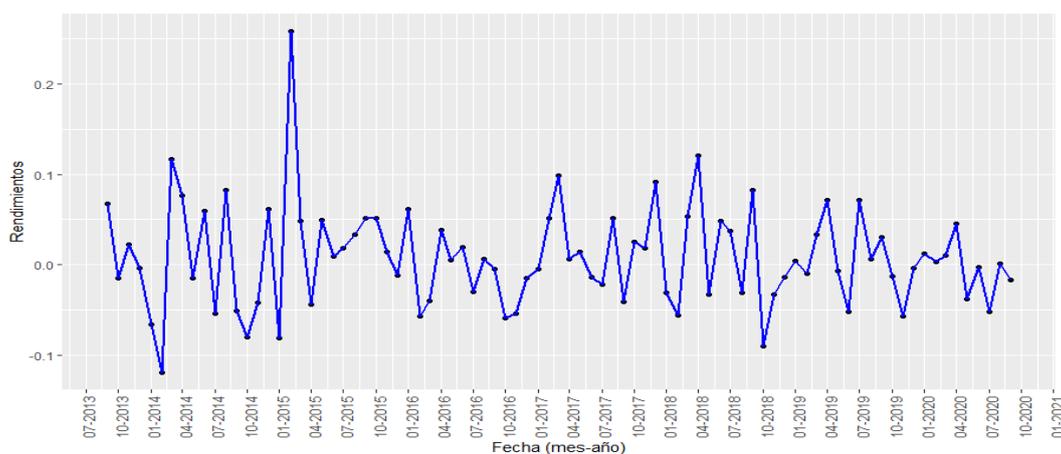
El histograma de frecuencias de la Figura 6.5(a) muestra que la distribución de los datos analizados presenta una asimetría positiva que se ve reflejada en el alargamiento de la cola derecha. Además, debido al alto exceso de curtosis se tiene una gran concentración de observaciones entorno a la media que resultó ser 0.0077 y por ende la forma del histograma presenta un pico. La Figura 6.5(b) muestra que el rendimiento mensual máximo en los últimos años se obtuvo

en febrero de 2015 con un valor de 0.2586, mientras que el rendimiento mínimo se obtuvo en febrero de 2014 con un valor de -0.1197. Debe notarse que a pesar de las condiciones económicas que México ha enfrentado en los últimos años Wal-Mart no se ha visto afectado debido a que al ser una cadena de tiendas que abastecen a toda la república, su consumo es indispensable para ciertos sectores de la población.

Figura 6.5: Rendimientos semanales de las acciones de Wal-Mart de Marzo, 2019 a Septiembre, 2020.



(a) Histograma de frecuencias.



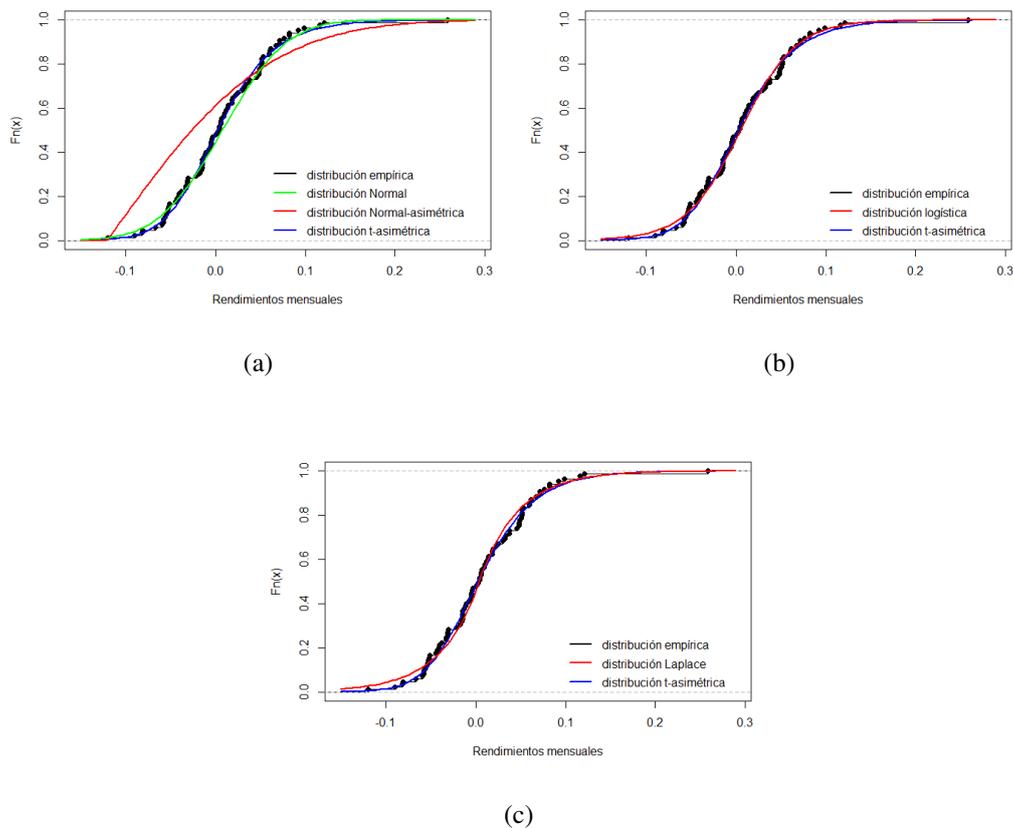
(b) Gráfica en el tiempo.

A continuación se presentan las estimaciones de los parámetros para cada distribución y su respectiva curva ajustada.

Distribución	Estimaciones
$ST(\xi, \omega, \alpha, \nu)$	(-0.0371, 0.06257, 1.4062, 8.9229)
$N(\mu, \sigma)$	(0.0077, 0.0561)
$SN(\xi, \omega, \alpha)$	(-0.1202, 0.1395, 183.4461)
$Logis(\mu, \sigma)$	(0.0051, 0.0304)
$Lap(\mu, b)$	(0.004, 0.0422)

Cuadro 6.5: Parámetros estimados de los rendimientos mensuales de Wal-Mart de Septiembre, 2013 a Septiembre, 2020 utilizando diferentes distribuciones.

Figura 6.6: Curvas ajustadas para los rendimientos mensuales de Wal-Mart de Septiembre, 2013 a Septiembre, 2020.



El Cuadro 6.5 muestra que la estimación del parámetro de asimetría α de la distribución Normal

asimétrica resultó ser bastante grande por lo que la curva ajustada con dicha distribución la cual se muestra en la Figura 6.6(a) tiene un mal ajuste comparado con el de la distribución t-asimétrica y la Normal. Pero debe notarse que la curva de la distribución t-asimétrica tiene un mejor ajuste comparado con el de la distribución Normal esto debido a que la distribución t-asimétrica capta la asimetría y el alto exceso de curtosis de los datos. Por otra parte la Figura 6.6(b) muestra que la curva ajustada con la distribución logística es muy similar a la ajustada con la t-asimétrica, excepto en la cola izquierda de la distribución, ya que en este caso la curva con la t-asimétrica tiene un mejor ajuste.

La Figura 6.6(c) muestra el ajuste con la distribución Laplace, la cual tiene un buen ajuste en parte del centro de la distribución empírica pero en las colas de ésta la distribución t-asimétrica presenta un mejor ajuste.

El siguiente cuadro muestra los resultados obtenidos al aplicar la prueba de independencia y cada una de las pruebas de bondad de ajuste a los datos, para cada prueba se utilizó un nivel de significancia de 0.05.

Prueba	BP	SW	AD_N	AD_L	AD_{LP}	W	W^*
Valor del estadístico	1.1486	0.9478	10.0482	0.3261	1.431	0.2363	0.4232
Valor-p	0.2838	0.0017	0	0.9171	0.0358	0.504	0.162
Rechazar H_0	No	Sí	Sí	No	Sí	No	No

Cuadro 6.6: Resultados de las diferentes pruebas aplicadas a los rendimientos semanales de Wal-Mart de Marzo, 2019 a Septiembre, 2020.

En el cuadro anterior se puede ver que se rechazó la hipótesis de normalidad tanto simétrica como asimétrica y la hipótesis de la prueba de Anderson-Darling para la distribución Laplace. Note que los valores de probabilidad más grandes se obtuvieron con las pruebas de Anderson-Darling para la logística y la t-asimétrica y por tanto se puede considerar que dichas distribuciones proveen un buen ajuste a los datos.

CAPÍTULO 7

DISCUSIÓN DE RESULTADOS Y CONCLUSIONES

Se propusieron dos pruebas de bondad de ajuste para la distribución t-asimétrica con parámetros desconocidos.

En general las dos pruebas propuestas (W y W^*) preservan el tamaño nominal de prueba para tamaños de muestras menores a 500 y en algunos casos en que las muestras son de tamaño 500 no se tiene control de éste.

Las dos pruebas tienen alta potencia para distribuciones alternativas relacionadas con la distribución Uniforme(0,1), la Beta(0.5, 0.5), la exponencial y la Laplace asimétrica con parámetro de asimetría entre 0.5 y 3.

Para las distribuciones alternativas en que los grados de libertad estimados son cercanos a 1 se obtuvieron potencias muy pequeñas a medida que el tamaño de muestra aumentó a 500.

El análisis realizado a los rendimientos del IPC en diferentes períodos mostró que las observaciones no son independientes cuando se consideran conjuntos de datos que contienen rendimientos diarios y rendimientos semanales en períodos largos con más de 200 observaciones. Además, en la gran mayoría de conjuntos de datos analizados las estimaciones del coeficiente de asimetría fueron diferentes de 0 y las estimaciones del coeficiente de exceso de curtosis fueron mayor a 0.3.

En el caso de los rendimientos semanales del IPC (para períodos cortos) en tres de los cuatro períodos analizados se rechazó la hipótesis de normalidad usando la prueba de Shapiro-Wilk. Análogamente, en el caso de los rendimientos mensuales siempre se rechazó dicha hipótesis. Por lo anterior, se tiene evidencia suficiente para concluir que en general los rendimientos semanales y mensuales no siguen dicha distribución. Debe notarse que sólo en un caso se rechazó la hipótesis de normalidad asimétrica, pero los ajustes realizados gráficamente con dicha distribución muestran en varios casos un mal ajuste a los datos, lo cual indica que la prueba de bondad de ajuste utilizada tiene baja potencia en dichos casos. Por otra parte, en ningún caso

se rechazó la hipótesis para la distribución logística y en tres de los siete casos se rechazó la hipótesis para la distribución Laplace, lo cual lleva a considerar la distribución logística como un posible modelo apropiado para el estudio de la variable de interés. Pero los resultados no aportan suficiente evidencia para considerar la distribución Laplace como un modelo válido.

Los resultados de los análisis realizados para los rendimientos del IPC dan evidencia de que la distribución t-asimétrica es un modelo adecuado para aproximar la distribución de los rendimientos semanales y mensuales.

En el análisis realizado con los rendimientos de los bonos a 10 años se obtuvo que en general los rendimientos diarios y mensuales no se pueden considerar independientes.

Para los dos períodos de los rendimientos semanales de los bonos a 10 años se obtuvo que en ambos casos el coeficiente de asimetría estimado fue negativo y el coeficiente estimado de exceso de curtosis fue mayor a 5, mostrando que la distribución de los rendimientos en dichos casos presenta un pico alto. En los dos casos analizados se rechazó la hipótesis de normalidad utilizando la prueba de bondad de ajuste de Shapiro-Wilk y la hipótesis de normalidad asimétrica usando la prueba de Anderson-Darling. Por tanto, se puede considerar que en general la variable estudiada no sigue una distribución normal. Los resultados al aplicar la prueba de Anderson-Darling para la distribución logística en los dos períodos estudiados dieron como resultado que en ambos casos no se rechaza la hipótesis correspondiente, por lo que la distribución logística se puede considerar como un modelo adecuado para el análisis de los rendimientos semanales de los bonos a 10 años. De forma similar, se aplicó la prueba de Anderson-Darling para probar la distribución Laplace y en uno de los dos conjuntos de datos se rechazó la hipótesis correspondiente.

Los resultados de aplicar a los rendimientos de los bonos a 10 años las pruebas de Anderson-Darling propuestas (W y W^*) llevan a considerar la distribución t-asimétrica como un modelo apropiado para el estudio de éstos.

En el análisis realizado para los rendimientos de las acciones de Wal-Mart, se obtuvo que el valor de probabilidad de la prueba de independencia de Box-Pierce para los rendimientos diarios fue mayor al nivel de significancia utilizado, lo cual permitió realizar un análisis estadístico completo a dichos datos. En los tres períodos analizados se obtuvo un coeficiente de asimetría estimado diferente de 0 y en el caso de los rendimientos mensuales el coeficiente fue cercano a 1. Por otra parte, el coeficiente de exceso de curtosis obtenido para los rendimientos diarios

y semanales fue negativo. Para estos casos, el histograma de frecuencias presentó una forma achatada. Para los rendimientos mensuales el coeficiente de exceso de curtosis resultó ser mayor a 3 por lo cual su histograma de frecuencias presentó un pico alto.

Los resultados de las pruebas de bondad de ajuste aplicadas a los rendimientos de las acciones de Wal-Mart permiten descartar a la distribución Laplace como un modelo apropiado para su estudio, ya que en dos de los tres casos analizados se rechazó la hipótesis correspondiente de la prueba de Anderson-Darling para dicha distribución.

Por otro lado, no se rechazó la hipótesis de normalidad aplicando la prueba de Shapiro-Wilk a los rendimientos diarios y semanales de las acciones de Wal-Mart, pero debe notarse que los valores de probabilidad obtenidos no son tan grandes como los obtenidos con la prueba de Anderson-Darling para la normal asimétrica. Además, debido a que para los rendimientos mensuales se rechazó la hipótesis de normalidad tanto simétrica como asimétrica, no se puede afirmar que la distribución normal sea un modelo apropiado para la variable de interés. Análogamente, en dos de los tres casos anteriores se rechazó la hipótesis correspondiente de la prueba de Anderson-Darling para la distribución Laplace, pero en ningún caso se rechazó para la distribución logística y t-asimétrica.

En general, los resultados obtenidos llevan a considerar las distribuciones logística y t-asimétrica como modelos probabilísticos que ajustan de forma adecuada los rendimientos económicos del IPC, de los bonos a 10 años y de las acciones de Wal-Mart. Los ajustes realizados gráficamente con ambas distribuciones y los resultados de las pruebas de bondad de ajuste aplicadas apoyan dicha afirmación.

BIBLIOGRAFÍA

- [1] T. W. Anderson and D. A. Darling. Asymptotic theory of certain. *Ann. Math. Statist.*, (2):193–212, 06.
- [2] Reinaldo B. Arellano-Valle and Adelchi Azzalini. The centred parameterization and related quantities of the skew-t distribution. *Journal of Multivariate Analysis*, 113:73 – 90, 2013. Special Issue on Multivariate Distribution Theory in Memory of Samuel Kotz.
- [3] A. Azzalini. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2):171–178, 1985.
- [4] A Azzalini. R package sn: the skew-normal and skew-t distributions. *Version 1.6-1. Padova, Italy: Universita di Padova*, 2020.
- [5] Adelchi Azzalini. *The skew-normal and related families*, volume 1. Cambridge University Press, 2014.
- [6] G. E. P. Box and David A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526, 1970.
- [7] MARCIA D’ELIA BRANCO, MARC G. GENTON, and BRUNERO LISEO. Objective bayesian analysis of skew-t distributions. *Scandinavian Journal of Statistics*, 40(1):63–85, 2013.
- [8] Eleanor Burke, Richard Perry, and Simon Brown. An extreme value analysis of uk drought and projections of change in the future. *Journal of Hydrology*, 388:131–143, 06 2010.
- [9] R. De Rose, M. Lanuzza, F. Crupi, G. Siracusano, R. Tomasello, G. Finocchio, M. Carpentieri, and M. Alioto. A variation-aware timing modeling approach for write operation in hybrid cmos/stt-mtj circuits. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(3):1086–1095, 2018.

- [10] Stephan Dreiseitl and Lucila Ohno-Machado. Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, 35(5):352 – 359, 2002.
- [11] James Eastman and Lee J. Bain. A property of maximum likelihood estimators in the presence of location-scale nuisance parameters. *Communications in Statistics*, 2(1):23–28, 1973.
- [12] B. Efron. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1–26, 01 1979.
- [13] F. Ghaderinezhad, C. Ley, and N. Loperfido. Bayesian inference for skew-symmetric distributions. *Symmetry*, 12(4), 2020. cited By 0.
- [14] A. Gonzalez-Blanks, J.M. Bridgewater, and T.M. Yates. Statistical approaches for highly skewed data: Evaluating relations between maltreatment and young adults’ non-suicidal self-injury. *Journal of Clinical Child and Adolescent Psychology*, 49(2):147–161, 2020. cited By 0.
- [15] Elizabeth González-Estrada and Waldenia Cosmes. Shapiro–wilk test for skew normal distributions based on data transformations. *Journal of Statistical Computation and Simulation*, 89(17):3258–3272, 2019.
- [16] Elizabeth González-Estrada and José A. Villaseñor. A ratio goodness-of-fit test for the laplace distribution. *Statistics Probability Letters*, 119:30 – 35, 2016.
- [17] Elizabeth González-Estrada and Jose A. Villaseñor-Alva. *Tests of Fit for some Probability Distributions*, 2020. R package version 1.3.6.
- [18] K. Guerra-Peña, Z.E. García-Batista, S. Depaoli, and L.E. Garrido. Class enumeration false positive in skew-t family of continuous growth mixture models. *PLoS ONE*, 15(4), 2020. cited By 0.
- [19] AK Gupta. Multivariate skew t-distribution. *Statistics: A Journal of Theoretical and Applied Statistics*, 37(4):359–363, 2003.
- [20] Ramesh Gupta and Nicole Brown. Reliability studies of the skew-normal distribution and its application to a strength-stress model. *Communications in Statistics-theory and Methods - COMMUN STATIST-THEOR METHOD*, 30:2427–2445, 01 2001.

- [21] Richard Harris and C. Kucukozmen. The empirical distribution of stock returns: Evidence from an emerging european market. *Applied Economics Letters*, 8:367–71, 02 2001.
- [22] Ya-Ling Huang. Forecasting the demand for tourism in china by applying the skew-t distribution model. *The Chinese Economy*, 46(4):49–62, 2013.
- [23] Olga Julià and José Rego. A microbiology application of the skew-laplace distribution. *Sort: Statistics and Operations Research Transactions, ISSN 1696-2281, Vol. 32, N° 2, 2008, pags. 141-150*, 32, 07 2008.
- [24] John Marsaglia Adrian Baddeley Julian Faraway, George Marsaglia. *Classical Goodness-of-Fit Tests for Univariate Distributions*, 2019. R package version 1.2-2.
- [25] Lukasz Komsta and Frederick Novomestky. *Moments, cumulants, skewness, kurtosis and related tests*, 2015. R package version 0.14.
- [26] Vasileios M. Koutras, Konstantinos Drakos, and Markos V. Koutras. A polynomial logistic distribution and its applications in finance. *Communications in Statistics - Theory and Methods*, 43(10-12):2045–2065, 2014.
- [27] B. Lagos Álvarez and M.D. Jiménez Gamero. A note on bias reduction of maximum likelihood estimates for the scalar skew t distribution. *Journal of Statistical Planning and Inference*, 142(2):608 – 612, 2012.
- [28] Yonghui Liu, Guohua Mao, Víctor Leiva, Shuangzhe Liu, and Alejandra Tapia. Diagnostic analytics for an autoregressive model under the skew-normal distribution. *Mathematics*, 8:693, 05 2020.
- [29] G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.
- [30] M. Bahrami M. Maghami. Goodness of fit test for the skew-t distribution. 14(4):274–283, 2015.
- [31] M.M. Maghami, M. Bahrami, and F.A. Sajadi. On bias reduction estimators of skew-normal and skew-t distributions. *Journal of Applied Statistics*, 2020. cited By 0.
- [32] Alexander McFarlane Mood, Franklin A Graybill, and Duane C Boes. *Introduction to the Theory of Statistics 1974*. McGraw-Hill Kogakusha, 1974.

- [33] J. Nakajima. Skew selection for factor stochastic volatility models. *Journal of Applied Statistics*, 47(4):582–601, 2020. cited By 0.
- [34] Kheirolah Okhli, M. Mozafari, and M. Naderi. Skew laplace finite mixture modelling. *Journal of the Iranian Statistical Society*, 16:97–110, 01 2017.
- [35] F Paz. La crisis y sus efectos en la economía mexicana. *Revista Economía Informal* (362), pages 93–107, 2010.
- [36] ES Pearson and H Hartley. 0.(1972). *biometrika tables for statisticians*, vol. 2, 1981.
- [37] Amado Peiro. The distribution of stock returns: International evidence. *Applied Financial Economics*, 4:431–39, 02 1994.
- [38] Brian Ripley. *Support Functions and Datasets for Venables and Ripley’s MASS*, 2020. R package version 7.3-51.6.
- [39] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965.
- [40] Student. The probable error of a mean. *Biometrika*, 6(1):1–25, 1908.
- [41] ”B. Swihart and J. Lindsey”. R package rmutil: Utilities for nonlinear regression and repeated measurements models. *Version 1.1.4.*, 2020.
- [42] R Core Team. R core team. r: A language and environment for statistical computing. *Foundation for Statistical Computing*, 2020.
- [43] Pierre-François Verhulst. Resherches mathematiques sur la loi d’accroissement de la population. *Nouveaux memoires de l’academie royale des sciences*, 18:1–41, 1845.
- [44] Ronald E Walpole, Raymond H Myers, Sharon L Myers, and Keying Ye. *Probability and statistics for engineers and scientists*, volume 5. Macmillan New York, 1993.
- [45] Claudia Yeap, s.T. Choy, and Simon Kwok. *The Skew-t Option Pricing Model*, pages 309–326. 01 2018.

ANEXOS

```
#Prueba de Anderson-Darling W
library(sn)
#La siguiente funcion calcula el estadistico
#de Anderson-Darling
AD_stat2<- function(x){
  n<- length(x) # sample size
  w<- st.mple(y = x, opt.method = "nlminb")$dp
  s<- sort(x)
  theop<- pst(s, w[1], w[2], w[3],w[4])
  ad_calc<- - n - sum((2*(1:n)-1)*log(theop) +
  (2*n+1-2*(1:n))*log(1-theop))/n
  return(ad_calc)
}
# La siguiente funcion regresa 1 si Ho es rechazada
# y 0 de otra forma.
# Argumentos
# x: datos (vector)
# B: numero de muestras Bootstrap
# size: tamaño de la prueba

AD_test2 <- function(x, B, size)
{
  n<- length(x)
  ad_calc <- AD_stat2(x)
  w <- st.mple(y=x, opt.method="nlminb")$dp
  #bootstrap
```

```

AD_dist <- replicate(B, AD_stat2(rst(n, xi = w[1],
omega=w[2], alpha = w[3], nu=w[4])))
orden<- sort(AD_dist)
cuant<- B*(1-size)
k_alpha <- orden[cuant]
return(ifelse(ad_calc > k_alpha, 1, 0))
}
#La siguiente funcion calcula el Valor-p de la prueba
pvalor<- function(x, B){
  n<- length(x)
  ad_calc<- AD_stat2(x)
  w<- st.mple(y=x, opt.method="nlminb")$dp
  #bootstrap
  AD_dist<- replicate(B, AD_stat2(rst(n, xi = w[1],
omega = w[2], alpha= w[3], nu=w[4])))
  p.value<- 1-sum(ad_calc > AD_dist)/B
  return(p.value)
}

#Prueba de Anderson-Darling W*
library(sn)
ADstat <- function(x){
  n<- length(x) # sample size
  w<- st.mple(y=x, opt.method="nlminb", symmetr = TRUE)$dp
  s <- sort(x)
  theop <- pst(s, w[1], w[2], 0, w[3])
  ad_calc <- -n- sum(((2*(1:n)-1)*log(theop)
+ (2*n+1-2*(1:n))*log(1-theop)))/n
  return(ad_calc)
}

ADtest<- function(x, B, size)
{

```

```

n<- length(x)
ad_calc<- ADstat(x)
w<- st.mple(y=x, opt.method="nlminb", symmetr = TRUE)$dp
AD_dist <- replicate(B, ADstat(rst(n,w[1],w[2],0,w[3])))
orden  <- sort(AD_dist)
cuant  <- B*(1-size)
k_alpha <- orden[cuant]
return(ifelse(ad_calc > k_alpha, 1, 0))
}
# La siguiente funcion regresa 1 si Ho es rechazada
# y 0 de otra forma.
# Argumentos
# x: datos (vector)
# B: numero de muestras Bootstrap
# size: tamaño de la prueba

prueba<-function(x,B,size) {
  n<-length(x)
  estimadores<-st.mple(y = x, opt.method = "nlminb")$dp
  x1<-(x-estimadores[1])/estimadores[2]
  A<-rbinom(n,1, 0.5) #variable de 0 y 1
  D<-matrix(NA, ncol = 1, nrow = n)
  for(i in 1:n){
    if (A[i]==1 ) {D[i]<-x1[i]} else D[i]<--(x1[i])}
  return(ADtest(D,B, size))
}
#Valor-p de la prueba
pvalorWstar<- function(x, B){
  n<-length(x)
  estimadores<-st.mple(y = x, opt.method = "nlminb")$dp
  x1<-(x-estimadores[1])/estimadores[2]
  A<-rbinom(n,1, 0.5) #variable de 0 y 1
  # variables transformadas
  D<-matrix(NA, ncol = 1, nrow = n)

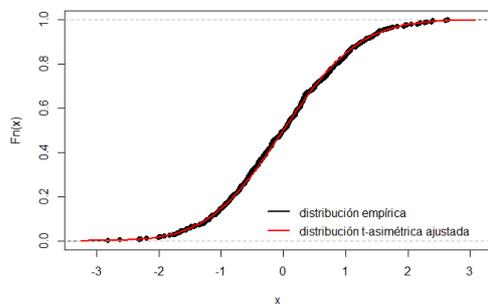
```

```

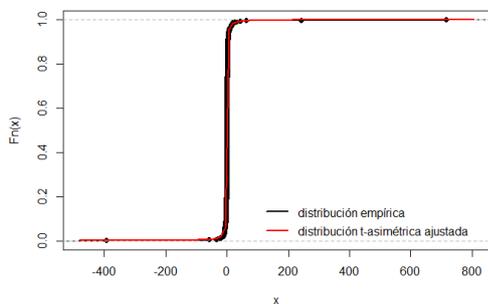
for(i in 1:n){
  if (A[i]==1 ) {D[i]<-x1[i]} else D[i]<--(x1[i])}
ad_calc <- ADstat(D)
w<- st.mple(y=D, opt.method="nlminb",symmetr = TRUE)$dp
AD_dist<- replicate(B, ADstat(rst(n,w[1],w[2],0,w[3])))
p.value <- 1-sum(ad_calc > AD_dist)/B
return(p.value)
}

```

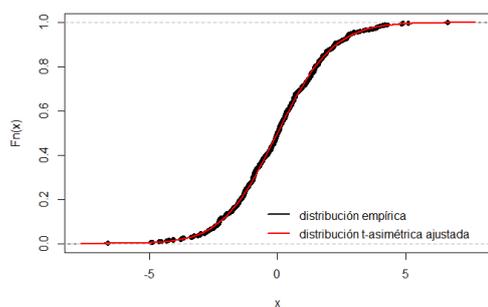
Figura 7.1: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=500$, parte I.



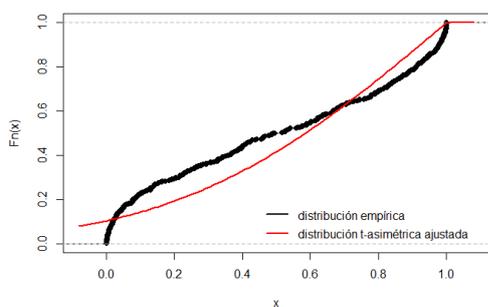
(a) Normal(0, 1)



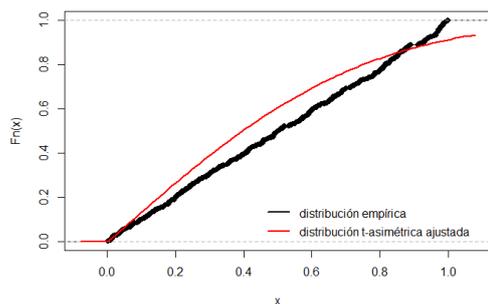
(b) Cauchy(0, 1)



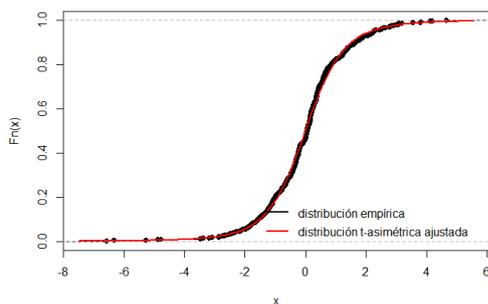
(c) Logística(0, 1)



(d) Beta(0.5, 0.5)

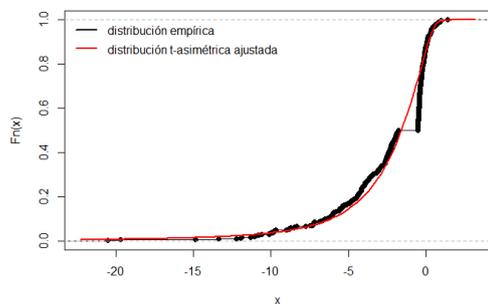


(e) Uniforme(0, 1)

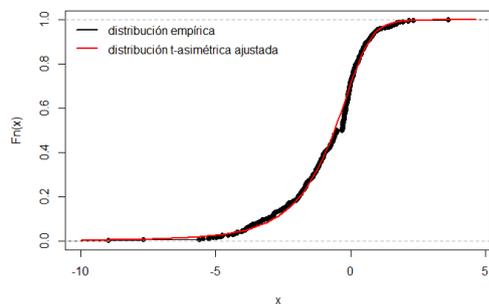


(f) Laplace(0, 1)

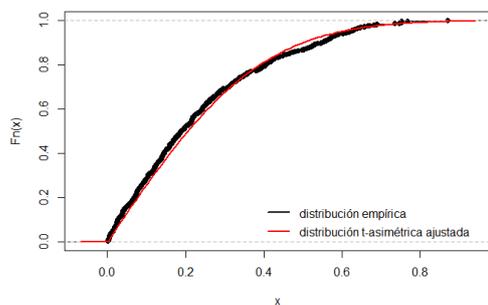
Figura 7.2: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=500$, parte II.



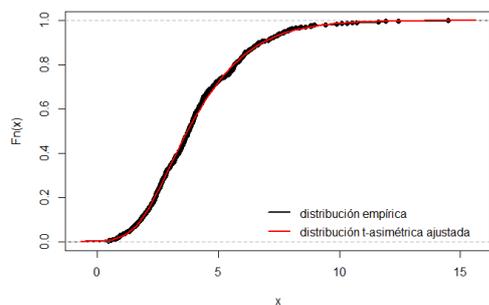
(a) sLaplace(0 , 1, 3)



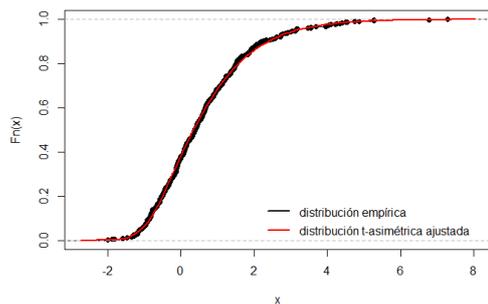
(b) sLaplace(0 , 1, 1.5)



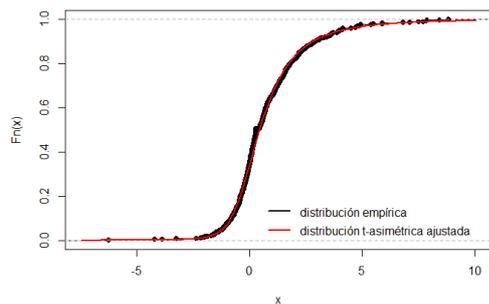
(c) Beta(1, 3)



(d) Gamma(4, 1)

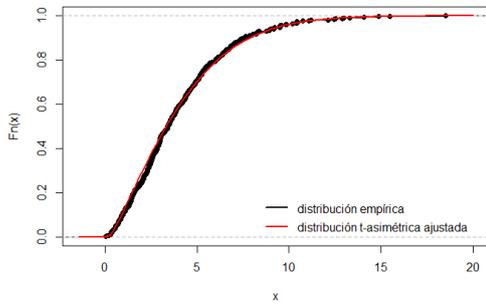


(e) Gumbel(0, 1)

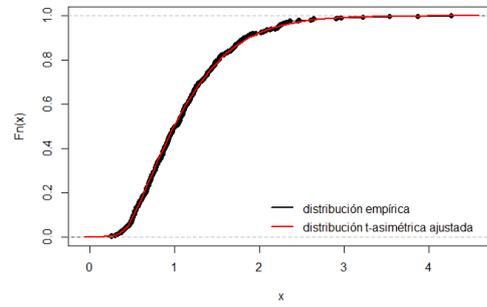


(f) sLaplace(0, 1, 0.7)

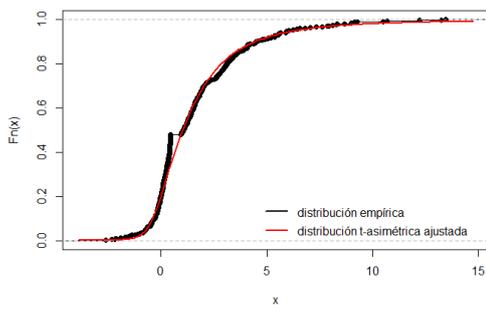
Figura 7.3: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=500$, parte III.



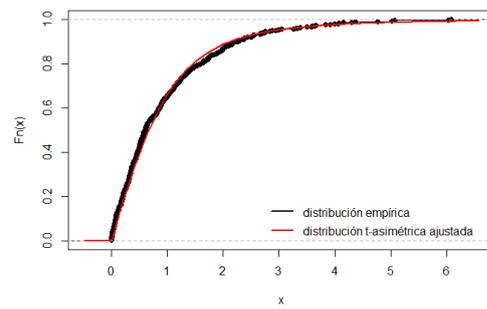
(a) χ_4^2



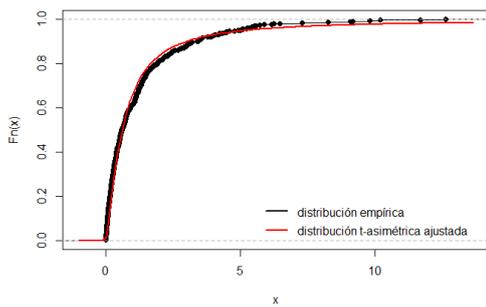
(b) Log-Normal(0, 0.5)



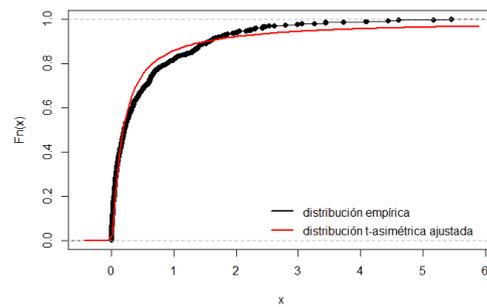
(c) sLaplace(0, 1, 0.5)



(d) Exp(1)

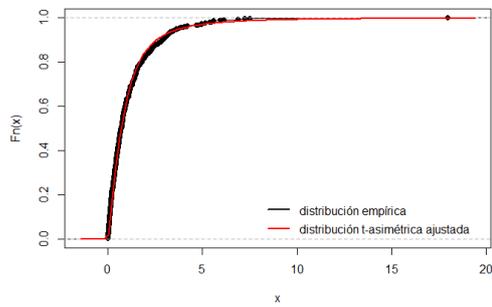


(e) Weibull(0.75, 1)

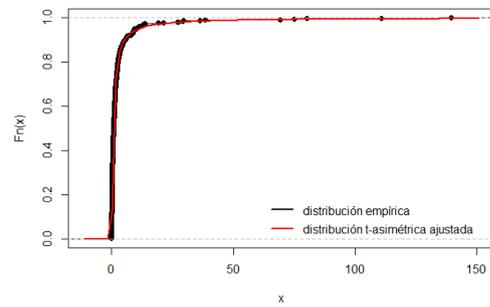


(f) Gamma(0.5, 1)

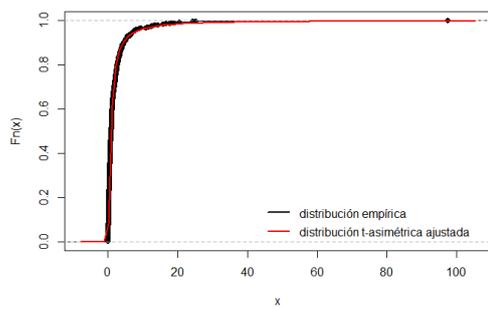
Figura 7.4: FDE y distribución t-asimétrica ajustada a muestras aleatorias de las distribuciones alternativas con $n=500$, parte IV.



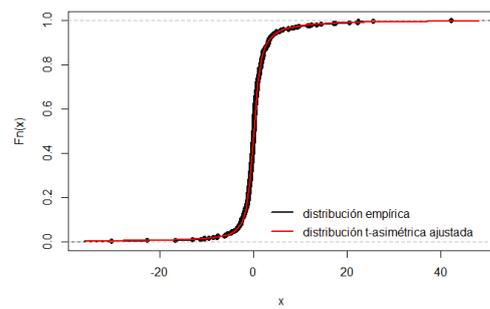
(a) ParetoG(0.15, 1)



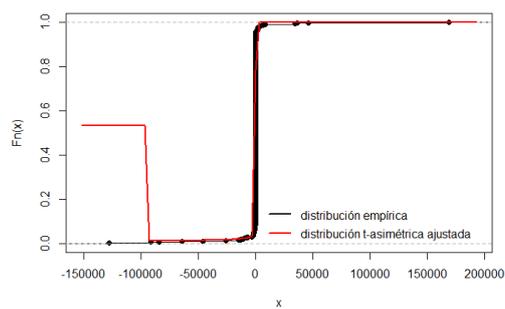
(b) Log-Normal(0, 1.5)



(c) ParetoG(0.5, 1)

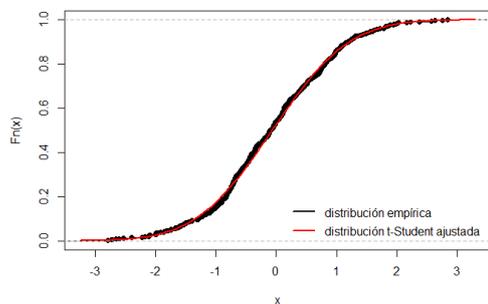


(d) DParetoG(0.4, 1)

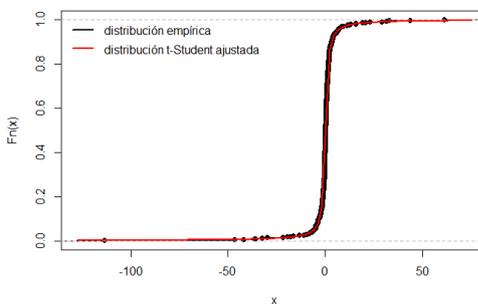


(e) DParetoG(2.5, 1)

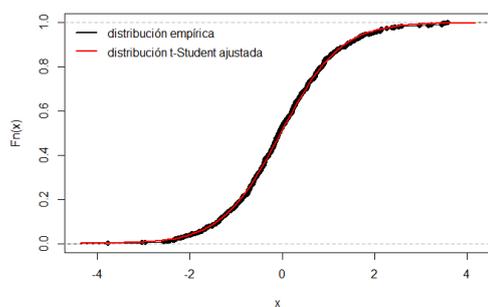
Figura 7.5: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte I.



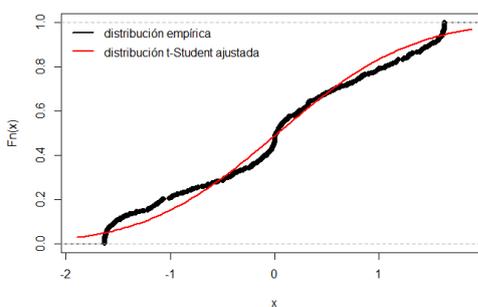
(a) Normal(0, 1)



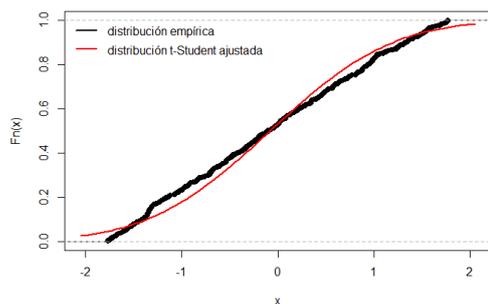
(b) Cauchy(0, 1)



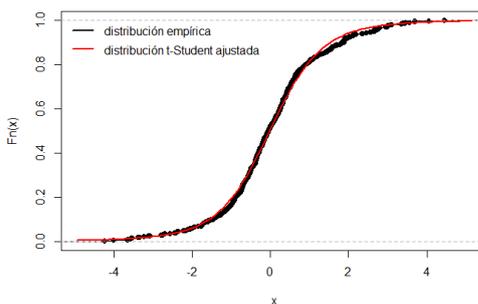
(c) Logística(0, 1)



(d) Beta(0.5, 0.5)

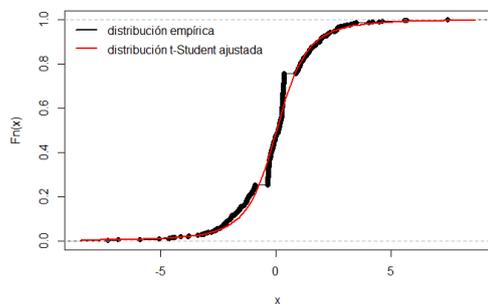


(e) Uniforme(0, 1)

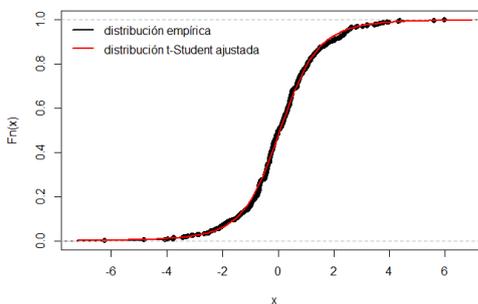


(f) Laplace(0, 1)

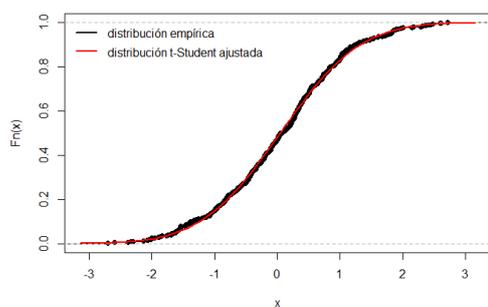
Figura 7.6: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte II.



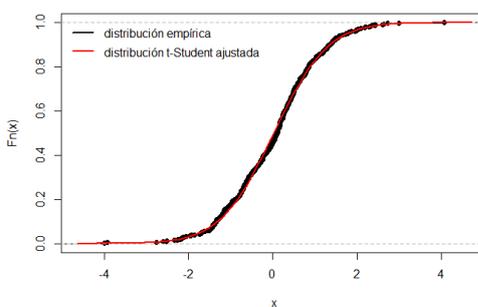
(a) $sLaplace(0, 1, 3)$



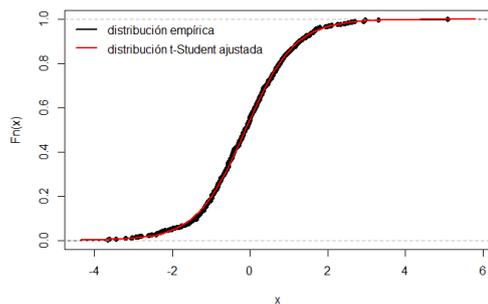
(b) $sLaplace(0, 1, 1.5)$



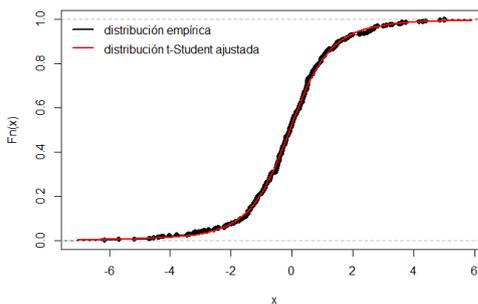
(c) $Beta(1, 3)$



(d) $Gamma(4, 1)$

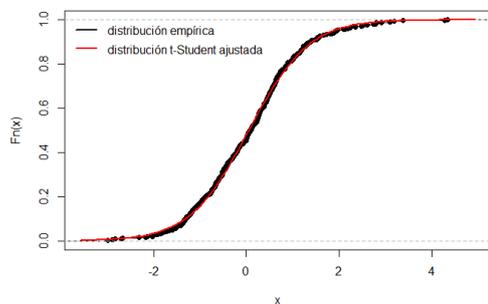


(e) $Gumbel(0, 1)$

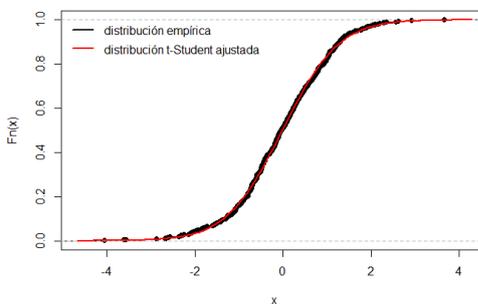


(f) $sLaplace(0, 1, 0.7)$

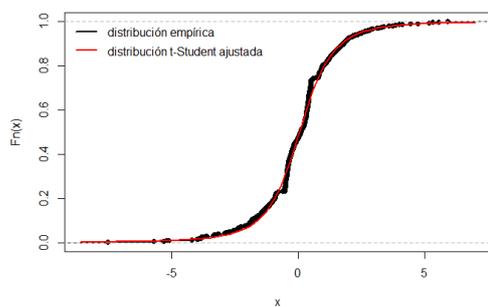
Figura 7.7: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte III.



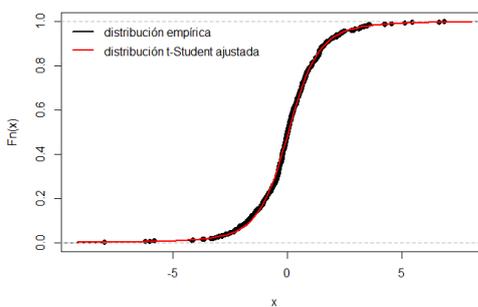
(a) χ_4^2



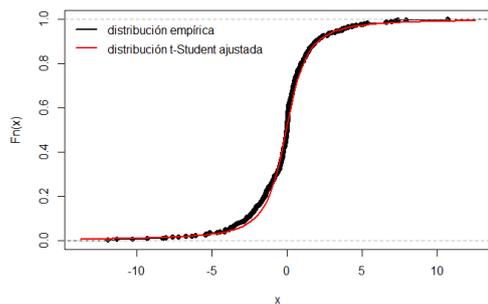
(b) Log-Normal(0, 0.5)



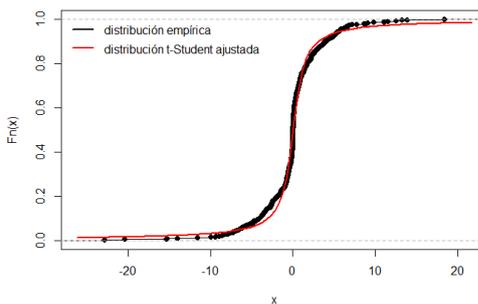
(c) sLaplace(0, 1, 0.5)



(d) Exp(1)

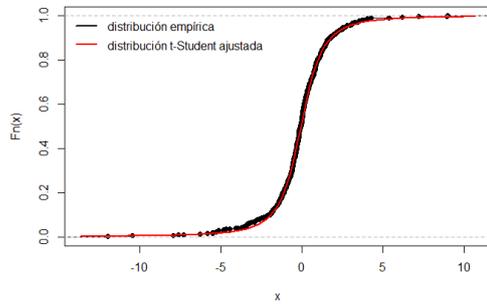


(e) Weibull(0.75, 1)

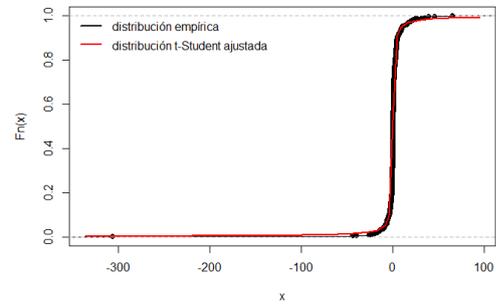


(f) Gamma(0.5, 1)

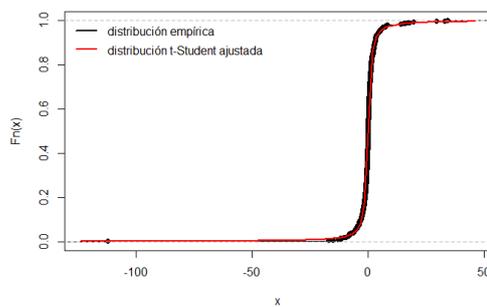
Figura 7.8: FDE y distribución t-Student ajustada a muestras aleatorias transformadas (usando 3.1) de las distribuciones alternativas con $n=500$, parte IV.



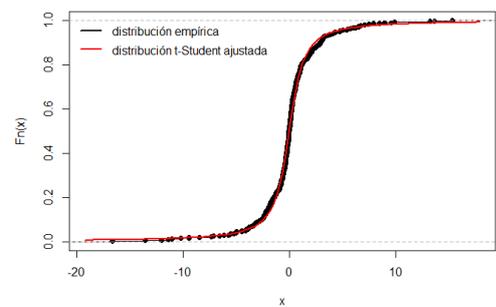
(a) ParetoG(0.15, 1)



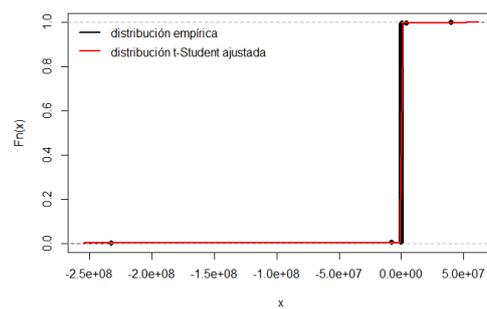
(b) Log-Normal(0, 1.5)



(c) ParetoG(0.5, 1)



(d) DParetoG(0.4, 1)



(e) DParetoG(2.5, 1)