

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO



Estudio de tres propuestas de distribución skew-t

Tesis para optar el grado Magister en
Estadística

Presentado por:

Alejandro Kantor Benavides

Asesor: Dr. Cristian Luis Bayes Rodríguez

Miembros del jurado:

Dr. Víctor Giancarlo Sal y Rosas Celi

Dr. Cristian Luis Bayes Rodríguez

Dr. Luis Hilmar Valdivieso Serrano

Lima, diciembre 2015

Resumen

Este trabajo compara tres distribuciones skew-t. En particular, las propuestas por Branco y Dey (2001) y Azzalini y Capitanio (2003), Fernández y Steel (1998), y Jones y Faddy (2003). Se analiza la relación entre los parámetros y el nivel de asimetría a través de la medida de Patil et al. (2014). Se propone una nueva parametrización de la distribución skew-t de Jones y Faddy (2003) que modela mejor la asimetría. Las distribuciones son ajustadas a datos reales basados en el retorno logarítmico de la tasa de cambio de PEN a USD.

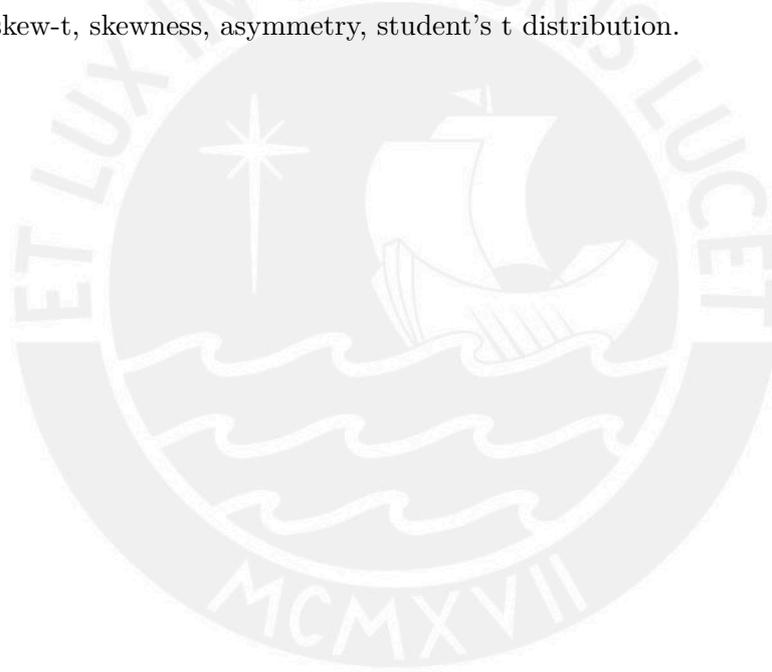
Palabras-clave: skew-t, skewness, asimetría, t-student.



Abstract

This paper compares three skew-t distributions. Namely, the ones proposed by Branco y Dey (2001) and Azzalini y Capitanio (2003), Fernández y Steel (1998), and Jones y Faddy (2003). The relationship between the parameters and the levels of asymmetry is analyzed through the asymmetry measure proposed by Patil et al. (2014). A new parameterization is proposed for the Jones y Faddy (2003) skew-t distribution that better models the asymmetry. The distributions are fitted to real data based on the logarithmic returns of the PEN USD exchange rate.

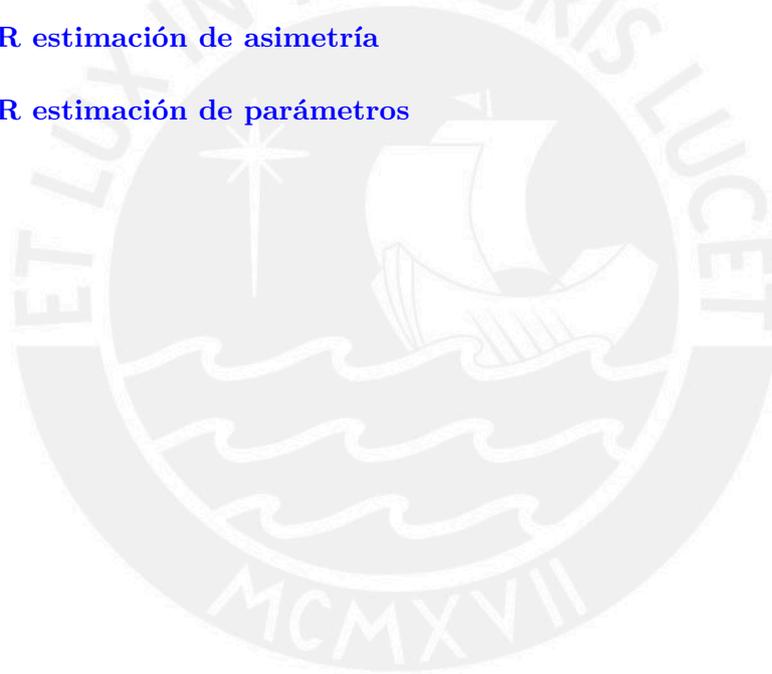
Keywords: skew-t, skewness, asymmetry, student's t distribution.



Índice general

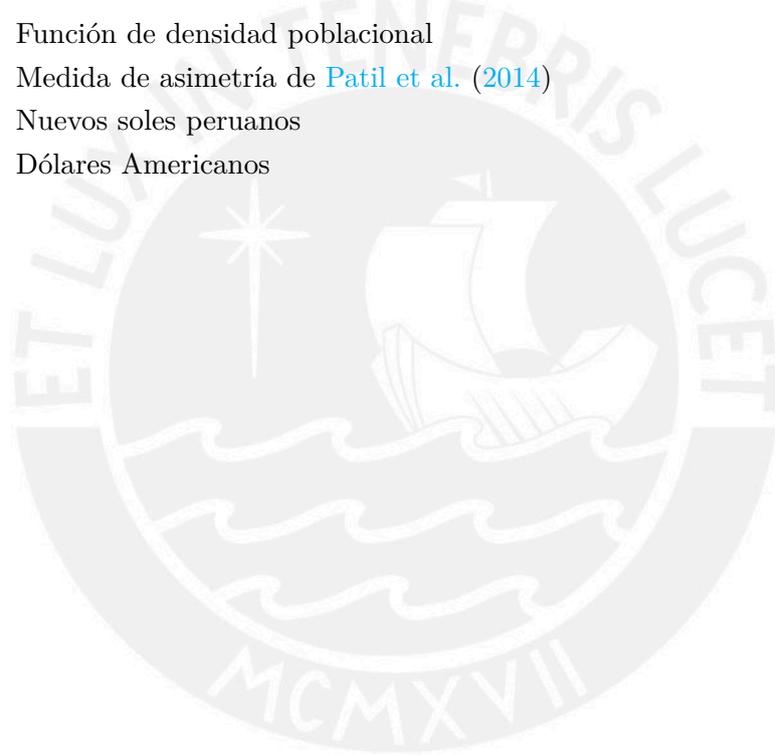
Lista de abreviaturas	VI
Índice de figuras	VII
Índice de cuadros	VIII
1. Introducción	1
1.1. Objetivos	1
1.2. Organización del Trabajo	2
2. Propuestas de distribución skew-t	3
2.1. Introducción	3
2.2. Distribución skew-t 1	4
2.2.1. Propiedades	5
2.3. Distribución skew-t 2	7
2.3.1. Propiedades	8
2.4. Distribución skew-t 3	10
2.4.1. Propiedades	11
3. Medidas de asimetría	15
3.1. Limitaciones de la medida de asimetría de Pearson	15
3.2. Medida de asimetría de Patil et al. (2014)	17
3.2.1. Relación con la medida Patil et al. (2012)	21
3.2.2. Propiedades de $\eta(\cdot)$	21
3.3. Método de estimación numérica de la medida de asimetría de Patil et al. (2014)	21
3.4. Cumplimiento de la condición $\int_{-\infty}^{\infty} f^3(x)dx < +\infty$	22
3.4.1. Cumplimiento en la skew-t 1	23
3.4.2. Cumplimiento en la skew-t 2	24
3.4.3. Cumplimiento en la skew-t 3	24
3.5. Asimetría en las distribuciones	25
3.5.1. Asimetría en la skew-t 1	25
3.5.2. Asimetría en la skew-t 2	26
3.5.3. Asimetría en la skew-t 3	26
3.5.4. Conclusion	28

4. Aplicaciones	29
4.1. Descripción y características de los datos	29
4.1.1. Datos de tipo de cambio	29
4.1.2. Depreciación logarítmica	30
4.2. Resultados	33
5. Resultados	37
6. Conclusiones	38
6.1. Sugerencias para investigaciones futuras	38
A. Estimación de parámetros	39
A.1. Principios teóricos de la EMV	39
A.2. Implementación computacional	40
A.2.1. Funciones y paquetes en R	41
B. Código R estimación de asimetría	42
C. Código R estimación de parámetros	46
Bibliografía	52



Lista de abreviaturas

AIC	Criterio de Información de Akaike (<i>Akaike Information Criterion</i>)
BCRP	Banco Central de Reserva del Perú
CAP	Coefficiente de asimetría de Pearson
EMV	Estimación por máxima verosimilitud
fd_a	Función de distribución acumulada
fd_p	Función de densidad poblacional
MAP	Medida de asimetría de Patil et al. (2014)
PEN	Nuevos soles peruanos
USD	Dólares Americanos



Índice de figuras

2.1. skew-t 1 v	6
2.2. skew-t 1 α	6
2.3. skew-t 2 v	9
2.4. skew-t 2 α	9
2.5. skew-t 3 cambio en a con b pequeño	12
2.6. skew-t 3 cambio en a con b grande	13
2.7. skew-t 3 v	13
2.8. skew-t 3 α	13
3.1. skew-t 2 CAP	16
3.2. skew-t 2 Asimetría	16
3.3. Ejemplo de medida de Patil et al. (2014) por métodos numéricos	22
3.4. Medida de asimetría skew-t 1	26
3.5. Medida de asimetría skew-t 2	26
3.6. Medida de asimetría skew-t 3 a y b	27
3.7. Medida de asimetría skew-t 3 p y q	27
3.8. Medida de asimetría skew-t 3 α y v	28
4.1. Tipo de cambio PEN/USD 1997-2015	29
4.2. Histograma de tipo de cambio nominal	30
4.3. Histograma de depreciación logarítmica a 3 meses.	31
4.4. Histograma de depreciación logarítmica a 6 meses.	31
4.5. Histograma de depreciación logarítmica a 9 meses.	32
4.6. Histograma de depreciación logarítmica a 12 meses.	32
4.7. Histograma de depreciación logarítmica a 3 meses y fdp estimadas	34
4.8. Histograma de depreciación logarítmica a 6 meses y fdp estimadas	34
4.9. Histograma de depreciación logarítmica a 9 meses y fdp estimadas	35
4.10. Histograma de depreciación logarítmica a 12 meses y fdp estimadas	35

Índice de cuadros

4.1. Estadísticas resumen	30
4.2. Parámetros estimados para depreciación logarítmica a 3 meses.	33
4.3. Parámetros estimados para depreciación logarítmica a 6 meses.	33
4.4. Parámetros estimados para depreciación logarítmica a 9 meses.	33
4.5. Parámetros estimados para depreciación logarítmica a 12 meses.	34
A.1. Funciones y paquetes en R para la implementación computacional por EMV	41



Capítulo 1

Introducción

Las distribuciones normal, log-normal, exponencial y gamma tienen limitaciones para describir algunos datos generados por la naturaleza y la economía, tales como el retorno de inversiones financieras ([Aas y Haff, 2006](#)), el ingreso en la industria cinematográfica ([Walls, 2005](#)) y la precipitación ([Marchenko y Genton, 2010](#)). Estos datos pueden ser altamente asimétricos y contar con puntos *outlier* que dificultan hacer inferencia con las distribuciones mencionadas.

En este contexto, la distribución skew- t es más flexible al permitir presencia de la asimetría y tener colas pesadas. En la literatura se encuentran diferentes propuestas de distribuciones asimétricas que tienen como caso particular o convergen a la distribución t -student; por ejemplo, las distribuciones propuestas por [Azzalini y Capitanio \(2003\)](#), [Branco y Dey \(2001\)](#), [Fernández y Steel \(1998\)](#), [Sahu et al. \(2003\)](#), [Jones y Faddy \(2003\)](#), [McNeil et al. \(2005\)](#), [Aas y Haff \(2006\)](#) y [Zhou y Galbraith \(2010\)](#). La existencia de varias propuestas de distribución skew- t amerita una comparación entre ellas, particularmente de sus propiedades de asimetría y de colas.

Medir el nivel de asimetría de una distribución es no-trivial, especialmente para casos donde la función de momentos no existe o puede no existir. En este contexto, para estudiar la asimetría de las distribuciones utilizamos la medida de asimetría definida por [Patil et al. \(2014\)](#).

Asimismo, es relevante analizar la relación de los parámetros de las distribuciones con sus colas y asimetría. En otras palabras, identificar qué parámetros modelan las colas y cuáles modelan la asimetría y si éstos son independientes entre sí.

Finalmente, para estudiar el desempeño de las distribuciones en modelar data, comparamos su ajuste a un conjunto de datos reales.

1.1. Objetivos

El objetivo general de la tesis es estudiar las propiedades de las tres propuestas de distribución skew- t más conocidas y hacer una comparación entre éstas. Para cumplir este objetivo se tienen los siguientes objetivos específicos:

- Revisar la literatura acerca de las diferentes propuestas de distribución skew- t y medidas de asimetría.

- Estudiar las propiedades de las diferentes propuestas de distribución skew-t y de la medida de asimetría seleccionada.
- Calcular o estimar los niveles de asimetría de las propuestas de distribución skew-t con respecto a sus parámetros.
- Implementar la estimación de las propuestas de distribución skew-t estudiadas utilizando programas ya existentes o propios.
- Aplicar los métodos de estimación de parámetros a datos reales.

1.2. Organización del Trabajo

Este trabajo inicia con la exposición de las tres propuestas de distribución skew-t más conocidas y sus propiedades principales en el capítulo 2 (densidad de probabilidad, función de distribución acumulada, momentos, etc); adicionalmente proponemos una nueva parametrización para la distribución skew-t de Jones y Faddy (2003).

El capítulo 3 describe la medida de asimetría de Patil et al. (2014) y estudia su aplicación a las propuestas de distribución skew-t para analizar la relación de la asimetría con los parámetros de las distribuciones exceptuando los parámetros de dispersión y posición.

El siguiente capítulo muestra la aplicación de las distribuciones a un conjunto de datos reales.

Con el fin de resumir la información de los capítulos anteriores, el capítulo 5 muestra, de manera sucinta, los resultados principales del estudio. Finalmente el capítulo 6 presenta las conclusiones y posibles extensiones que se podrían derivar de este trabajo.

Capítulo 2

Propuestas de distribución skew-t

En este capítulo presentamos las tres propuestas de distribución skew-t estudiadas en este documento. Primero, definimos la distribución t -student y luego, para cada propuesta de distribución skew-t, describimos sus propiedades y mostramos gráficos de ellas bajo diferentes valores de sus parámetros.

2.1. Introducción

La distribución t de Student (en adelante distribución t), denotada por $X \sim T(v)$, es una familia de distribuciones con la siguiente función de densidad de probabilidad (fdp)

$$t(x; v) = \frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}}, \quad x \in \mathbb{R} \quad (2.1)$$

donde $v \in \mathbb{R}^+$, llamado grados de libertad, es un parámetro tal que a menor valor de v , las colas de la distribución son más pesadas, es decir, que las probabilidades que la v.a. tome valores extremos son menos despreciables. En comparación con la distribución normal, la distribución t permite la presencia de puntos *outliers*. Por ende, se puede utilizar para modelar una población con una muestra pequeña o cuando se cree que la población cuenta con valores *outliers*.

No obstante, al ser simétrica, la distribución t no logra modelar bien datos con algún nivel no trivial de asimetría. La literatura ha propuesto una amplia gama de distribuciones que manejan tanto asimetría y datos *outliers*. Este trabajo analiza las tres distribuciones más conocidas que permiten asimetría y tienen como caso particular a la distribución t -student; siendo éstas las distribuciones skew-t propuestas por [Azzalini y Capitanio \(2003\)](#), [Fernández y Steel \(1998\)](#) y [Jones y Faddy \(2003\)](#). Dentro de este estudio denominamos a este grupo de distribuciones *propuestas de distribución skew-t*.

Este capítulo describe las propiedades principales de las propuestas de distribución skew-t. Por simplicidad, tomamos solo los casos de las distribuciones con parámetro de escala igual a uno ($\sigma^2 = 1$) y parámetro de posicionamiento igual a cero ($\mu = 0$). Cabe notar que sin esta restricción, las distribuciones tienen la siguiente forma $f(x; \theta, \mu, \sigma) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}; \theta\right)$.

Incluimos información sobre la relación entre los parámetros y el nivel de asimetría de las distribuciones. No obstante, presentamos las justificaciones y más información sobre esta relación en la Sección [3.5](#).

2.2. Distribución skew-t 1

La primera distribución estudiada es la distribución skew-t 1 propuesta por Branco y Dey (2001) y Azzalini y Capitanio (2003). La distribución está relacionada y se construye a partir de la distribución skew normal 1 de Azzalini (1985) como detallamos en esta sección.

Ambos, la distribución skew-t 1 y la skew normal 1 son casos particulares del método de construcción propuesto en Azzalini (1985) donde definimos una fdp asimétrica a través de la multiplicación de una fdp simétrica, $g(\cdot)$, con su función de distribución acumulada (fda) correspondiente, $G(\cdot)$ evaluada en una función a escalar $h(\cdot)$ de la siguiente manera:

$$f(x; \theta_1) = 2 g(x; \theta_2) G(h(x); \theta_3)$$

A partir de este método de construcción de distribuciones asimétricas, definimos la skew normal 1 al tomar $h(x) = \alpha x$ como se detalla en la siguiente definición.

Definición 1. Una variable aleatoria X tiene distribución skew normal 1, denotado $X \sim SN_1(\alpha)$, si y solo si su función de densidad es dada por

$$f(x; \alpha) = 2 \phi(x) \Phi(\alpha x) \quad (2.2)$$

donde $\phi(\cdot)$ es la fdp de la distribución normal, $\Phi(\cdot)$ es la fda de la distribución normal y α es un parámetro de asimetría. Cuando $\alpha = 0$ la distribución es simétrica y se reduce a la distribución normal, cuando $\alpha > 0$ la distribución es asimétrica a la izquierda y para $\alpha < 0$ la distribución es asimétrica a la derecha.

La construcción de la skew-t 1 se da a partir de la skew normal 1 en combinación con la siguiente propiedad que relaciona la distribución normal con la t : si $Z_0 \sim N(0, 1)$ y $V \sim \chi_v^2/v$ (distribución Chi cuadrado con v grados de libertad entre v), son variables independientes y $X = \frac{Z_0}{\sqrt{V}}$, entonces $X \sim T(v)$. Manteniendo todo lo anterior pero reemplazando la variable aleatoria Z_0 normal con $Z_0 \sim SN_1(\alpha)$, obtenemos $X \sim ST_1(\alpha, v)$. Detallamos la definición de la distribución skew-t 1 a continuación.

Definición 2. Una variable aleatoria X tiene distribución skew-t 1, denotada por $X \sim ST_1(\alpha, v)$, si y solo si su función de densidad es dada por

$$f(x; \alpha, v) = 2 t(x; v) T\left(\alpha x \sqrt{\frac{v+1}{v+x^2}}; v+1\right) \quad (2.3)$$

donde $t(\cdot)$ es la función de densidad de la distribución t -student dada en (2.1), $T(\cdot)$ es la función de distribución acumulada de la distribución t -student y v son los grados de libertad de la t univariada. Cabe mencionar que $v \in \mathbb{R}^+$ es un parámetro tal que a menor valor de v , las colas de la distribución son más pesadas, es decir, que las probabilidades de que la v.a. tome valores extremos son menos despreciables. De otro lado, α es un parámetro de asimetría. Cuando $\alpha = 0$ la distribución es simétrica y se reduce a la distribución t -student, si $\alpha > 0$ la distribución es asimétrica a la derecha y si $\alpha < 0$ la distribución es asimétrica a la izquierda.

2.2.1. Propiedades

En esta sección describimos las principales propiedades de la distribución skew-t 1; además, presentamos gráficos de su fdp que muestran algunas de estas propiedades. Para las pruebas de estas propiedades ver [Branco y Dey \(2001\)](#) y [Azzalini y Capitanio \(2003\)](#).

P1: Si $Z_0 \sim SN_1(\alpha, v)$ y $V \sim \chi_v^2/v$, son variables independientes y $X = \frac{Z_0}{\sqrt{V}}$ entonces $X \sim ST_1(\alpha, v)$.

P2: Cuando $v \rightarrow \infty$ la distribución tiende a $SN_1(\alpha)$.

P3: Las colas tienen el siguiente comportamiento ([Aas y Haff, 2006](#)):

$$f(x) \sim \text{const} \times |x|^{-v-1} \quad \text{cuando } x \rightarrow \pm\infty$$

Por lo tanto, a menor v , las colas de la distribución son más pesadas; es decir, las probabilidades de que la v.a. tome valores extremos son menos despreciables.

P4: α es un parámetro de asimetría. Cuando $\alpha = 0$ la distribución es simétrica y se reduce a la distribución t -student. Al aumentar el valor de $|\alpha|$ la asimetría aumenta es decir la masa de la probabilidad se concentra más a un lado de la distribución. Cuando $\alpha > 0$ la distribución es asimétrica a la derecha y para $\alpha < 0$ la distribución es asimétrica a la izquierda.

P5: $f(x; \alpha) = f(-x; -\alpha)$.

P6: Los momentos de esta distribución están dados por:

$$E(X^r) = \sqrt{\frac{2}{\pi}} \frac{\text{sign}(\alpha)}{\alpha^{r+1}} K_r(\alpha^{-2}) \times \frac{\left(\frac{v}{2}\right)^{\frac{r}{2}} \Gamma\left(\frac{1}{2}(v-r)\right)}{\Gamma\left(\frac{1}{2}v\right)}, \quad \text{si } v > r$$

$$\text{donde } K_0(\alpha^{-2}) = \text{abs}(\alpha) \sqrt{\frac{\pi}{2}}, \quad K_1(\alpha^{-2}) = \frac{\alpha^2}{\sqrt{1+\alpha^{-2}}}$$

$$K_r(\alpha^{-2}) = (r-1) \alpha^2 K_{r-2}(\alpha^{-2}) + \frac{w_{r-1} \alpha^2}{(1+\alpha^2)^{\frac{r}{2}}}$$

$$\text{y } w_c = \begin{cases} 0 & , \text{ si } c \text{ es impar} \\ \frac{c!}{2^{\frac{c}{2}} \left(\frac{c}{2}\right)!} & , \text{ si } c \text{ es par} \end{cases}$$

P7: La media es dada por

$$E(X) = \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1+\alpha^2}} \frac{\left(\frac{v}{2}\right)^{\frac{1}{2}} \Gamma\left(\frac{1}{2}(v-1)\right)}{\Gamma\left(\frac{1}{2}v\right)}, \quad \text{si } v > 1$$

P8: La varianza es dada por

$$\text{Var}(X) = \frac{\left(\frac{v}{2}\right) \Gamma\left(\frac{1}{2}(v-2)\right)}{\Gamma\left(\frac{1}{2}v\right)} - \frac{v}{\pi} \frac{1}{1+\alpha^{-2}} \frac{\Gamma\left(\frac{1}{2}(v-1)\right)^2}{\Gamma\left(\frac{1}{2}v\right)^2}, \quad \text{si } v > 2$$

A continuación presentamos gráficos que muestran la forma de las fdp dada en (2.3) para diferentes valores de α y v . En la Figura 2.1 observamos la skew-t 1 para diferentes valores de v con $\alpha = 1$. A menores valores de v las dos colas de la función son más pesadas. No observamos cambios en la asimetría para un α constante.

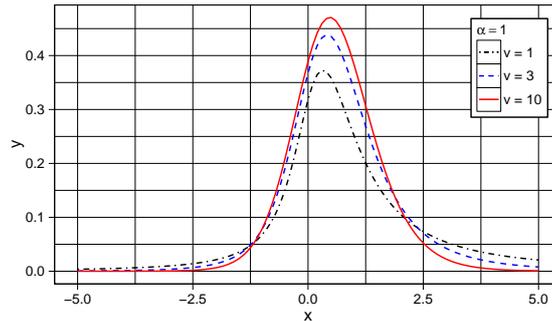


Figura 2.1: Función de densidad de la distribución skew-t 1 para $\alpha = 1$ y $v = 1, 3, 10$.

Por otro lado, observamos en la Figura 2.2 la skew-t 1 para diferentes valores de α con $v = 2$. A mayores valores de α observamos mayor asimetría. Cuando $\alpha = 0$ la distribución es simétrica y se reduce a la distribución t -student, cuando $\alpha > 0$ la distribución es asimétrica a la derecha y para $\alpha < 0$ la distribución es asimétrica a la izquierda. También observamos que $f(x, \alpha) = f(-x, -\alpha)$. Adicionalmente, vemos gráficamente que a mayor nivel de $|\alpha|$ se tiene mayor curtosis.

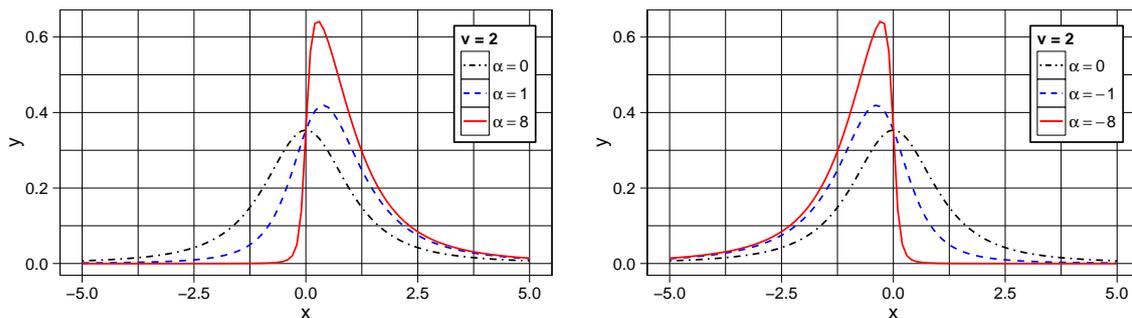


Figura 2.2: Función de densidad de la distribución skew-t 1 para $v = 2$ y diferentes valores de α . El panel izquierdo muestra la fdp para $\alpha = 0, 1, 8$ y el derecho para $\alpha = 0, -1, -8$.

Implementación computacional

El paquete *sn* de R (Azzalini, 2014) incluye rutinas que permiten calcular la función de densidad, distribución acumulada, cuantiles y generar valores aleatorios que siguen la distribución skew-t 1. Asimismo contiene la rutina *selm* para estimar los parámetros de la

distribución skew-t 1 con el método de máxima verosimilitud a través de métodos computacionales. Por otro lado, el paquete *gamlss.dist* (Stasinopoulos et al., 2014) incluye también rutinas que permiten calcular la función de densidad, distribución acumulada, cuantiles y generar valores aleatorios que siguen la distribución skew-t 1 y el paquete *gamlss* (Rigby y Stasinopoulos, 2005) contiene una rutina para estimar los parámetros de la distribución skew-t 1 con el método de máxima verosimilitud a través métodos computacionales.

2.3. Distribución skew-t 2

La segunda distribución skew-t estudiada es la distribución skew-t 2 propuesta por Fernández y Steel (1998). La distribución está relacionada a la distribución skew normal 2, como detallamos en esta sección.

Las dos distribuciones son casos particulares del método de construcción propuesto en Fernández y Steel (1998). Bajo este método, construimos una fdp asimétrica al manipular una fdp unimodal y simétrica $g(\cdot)$ de la siguiente manera:

$$f(x; \alpha, \theta) = \frac{2}{\alpha + \frac{1}{\alpha}} \left(g(\alpha x; \theta) I(x < 0) + g\left(\frac{x}{\alpha}; \theta\right) I(x \geq 0) \right) \quad (2.4)$$

donde I es la función indicadora.

En otras palabras, los autores forman la nueva fdp truncando de dos maneras a $g(\cdot)$. La primera $g(\cdot)$ truncada tiene valores diferentes a cero solo si $x < 0$; es decir, es truncada en su media; además, el área debajo de la curva es modificado al multiplicar x por α . La segunda $g(\cdot)$ truncada tiene valores diferentes a cero solo si $x \geq 0$ y el área debajo de la curva se transforma al multiplicar x por $1/\alpha$. Luego la adición de ambas funciones truncadas es multiplicada por una constante para que el área total de la nueva fdp sea uno (ver Figura 2.4).

Tomando en cuenta el método de construcción descrito anteriormente, presentamos la definición de la distribución skew normal 2 seguida por la definición de la distribución skew-t 2.

Definición 3. Una variable aleatoria X tiene distribución skew normal 2, denotada por $X \sim SN_2(\alpha)$, si y solo si su función de densidad es dada por

$$f(x; \alpha) = \frac{2}{\alpha + \frac{1}{\alpha}} \left(\phi(\alpha x) I(x < 0) + \phi\left(\frac{x}{\alpha}\right) I(x \geq 0) \right) \quad (2.5)$$

donde $\phi(\cdot)$ es la función de densidad de la distribución normal estándar y α es un parámetro de asimetría. Cuando $\alpha = 1$ la distribución es simétrica y se reduce a la distribución normal, cuando $\alpha > 1$ la distribución es asimétrica a la derecha y para $\alpha < 1$ la distribución es asimétrica a la izquierda.

Definición 4. Una variable aleatoria X tiene distribución skew-t 2, denotada por $X \sim$

$ST_2(\alpha, v)$, si y solo si su función de densidad es dada por

$$f(x; \alpha) = \frac{2}{\alpha + \frac{1}{\alpha}} \left(t(\alpha x; v) I(x < 0) + t\left(\frac{x}{\alpha}; v\right) I(x \geq 0) \right), \quad (2.6)$$

donde $t(\cdot)$ es la función de densidad de la distribución t -student dada en (2.1) y v son los grados de libertad de la t univariada. Así $v \in \mathbb{R}^+$ y a menor v las colas de la distribución son más pesadas. Aquí α es un parámetro de asimetría. Cuando $\alpha = 1$ la distribución es simétrica y se reduce a la distribución t -student, cuando $\alpha > 1$ la distribución es asimétrica a la derecha y para $\alpha < 1$ la distribución es asimétrica a la izquierda.

2.3.1. Propiedades

En esta sección describimos las propiedades principales de la distribución skew- t 2; además presentamos gráficos de su fdp. Para las pruebas de estas propiedades ver [Fernández y Steel \(1998\)](#) y [Arellano-Valle et al. \(2005b\)](#).

P1: Cuando $v \rightarrow \infty$ la distribución tiende a $SN_2(\alpha)$

P2: Las colas tienen el siguiente comportamiento ([Aas y Haff, 2006](#)):

$$f(x) \sim \text{const} \times |x|^{-v-1} \quad \text{cuando } x \rightarrow \pm\infty$$

Por lo tanto, a menor v , las colas de la distribución son más pesadas; es decir, las probabilidades de que la v.a. tome valores extremos son menos despreciables.

P3: α es el parámetro de asimetría. Cuando $\alpha = 1$ la distribución es simétrica y se reduce a la distribución t -student. Al incrementar $|\log(\alpha)|$ el nivel de asimetría aumenta. Cuando $\alpha > 1$ la distribución es asimétrica a la derecha y para $\alpha < 1$ la distribución es asimétrica a la izquierda.

P4: $f(x; \alpha, v) = f(-x; -1/\alpha, v)$.

P5: Los momentos de esta distribución vienen dados por

$$E(X^r) = \begin{cases} 2cv^{\frac{r+1}{2}} \sum_{j=0}^{\frac{r-1}{2}} \frac{(-1)^{\frac{r-1-2j}{2}}}{v-2j-1} \binom{\frac{r-1}{2}}{j} \frac{\alpha^{r+1-\frac{1}{\alpha}}}{\alpha+\frac{1}{\alpha}} & \text{si } r \text{ es impar, } v > r \\ \frac{v^{\frac{r}{2}} \times 1 \times 3 \times \dots \times (r-1)}{(v-r)(v-r+2)\dots(v-2)} \frac{\alpha^{r+1+\frac{1}{\alpha}}}{\alpha+\frac{1}{\alpha}} & \text{si } r \text{ es par, } v > r \end{cases}$$

donde $c = (v\pi)^{-1/2} \Gamma(v/2)^{-1} \Gamma((v+1)/2)$ ([Arellano-Valle et al., 2005b](#)).

P6: La media es dada por

$$E(X) = \frac{2cv(\alpha^2 - 1)}{(v-1)\alpha}, \quad \text{si } v > 1$$

P7: La varianza es dada por

$$\text{Var}(X) = \frac{v(\alpha^6 + 1)}{(v-2)\alpha^2(\alpha^2 + 1)} - \frac{4(c v (\alpha^2 - 1))^2}{(\alpha(v-1))^2}, \quad \text{si } v > 2$$

A continuación presentamos gráficos que muestran la forma de las fdp para diferentes valores de α y v . En la Figura 2.3 observamos la skew-t 1 para diferentes valores de v con $\alpha = 2$. A menores valores de v , las colas de la función son mayores. No observamos cambios en la asimetría para un α constante.

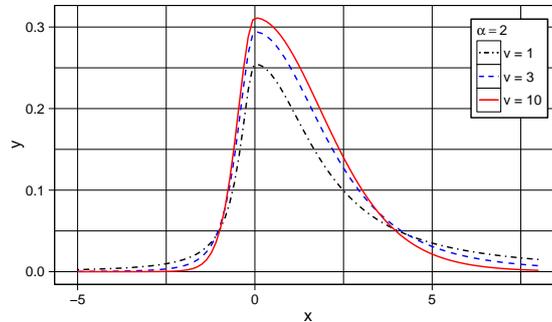


Figura 2.3: Función de densidad de la distribución skew-t 1 para $\alpha = 2$ y $v = 1, 3, 10$.

Por otro lado, observamos en la Figura 2.4 la fdp de la distribución skew-t 2 para diferentes valores de α con $v = 2$. A mayores valores de α observamos mayor asimetría. Cuando $\alpha = 1$ la distribución es simétrica y se reduce a la distribución t -student. Al alejarse α el nivel de asimetría aumenta. Cuando $\alpha > 1$ la distribución es asimétrica a la derecha y para $\alpha < 1$ la distribución es asimétrica a la izquierda. También observamos que $f(x, \alpha) = f(-x, -1/\alpha)$. Adicionalmente vemos gráficamente que a $|\log(\alpha)|$ se tiene menor nivel de asimetría. En otras palabras, al alejarse α de 1 la curtosis disminuye.

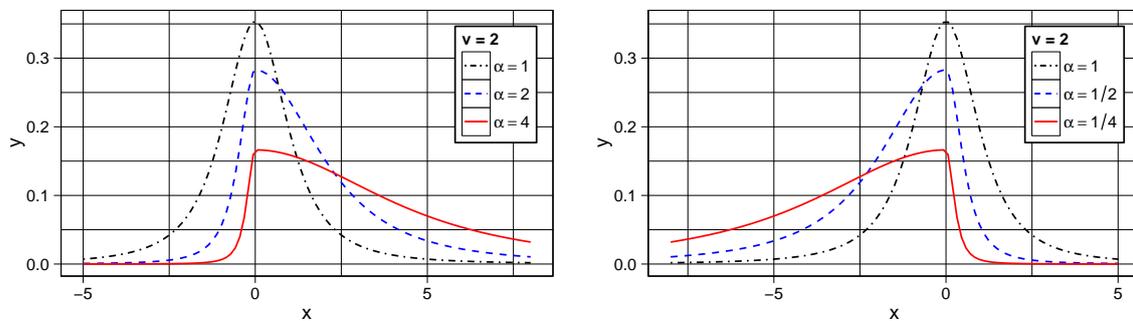


Figura 2.4: Función de densidad de la distribución skew-t 1 para $v = 2$ y diferentes valores de α . El panel izquierdo muestra la fdp para $\alpha = 1, 2, 4$ y el derecho para $\alpha = 1, \frac{1}{2}, \frac{1}{4}$.

Implementación computacional

El paquete *skewt* de R (King et al., 2012) incluye rutinas que permiten calcular la función de densidad, distribución acumulada, cuantiles y generar valores aleatorios que siguen la distribución skew-t 2. Por otro lado, el paquete *gamlss.dist* (Stasinopoulos et al., 2014) incluye también rutinas que permiten calcular la función de densidad, distribución acumulada, cuantiles y generar valores aleatorios que siguen la distribución skew-t 2 y el paquete *gamlss* (Rigby y Stasinopoulos, 2005) contiene una rutina para estimar los parámetros de la

distribución skew-t 2 con el método de máxima verosimilitud a través de métodos computacionales.

2.4. Distribución skew-t 3

La tercera distribución skew-t estudiada es la distribución skew-t 3 propuesta por Jones y Faddy (2003). A diferencia de las distribuciones skew-t 1 y 2, no hay propuesta en la literatura de una distribución skew normal asociada.

Adicionalmente, la parametrización original no separa el modelamiento del peso de las colas y de la asimetría directamente. En este contexto, estudiamos la parametrización original, la parametrización alterna mencionada en el artículo de Jones y Faddy (2003) para la Estimación por Máxima Verosimilitud (EMV) y finalmente construimos una parametrización propuesta que, como estudiaremos en la Sección 3.5, separa mejor el modelamiento de la asimetría y el peso de las colas.

A continuación, definimos la distribución y detallamos las parametrizaciones.

Definición 5. Una variable aleatoria X tiene distribución skew-t 3 si y solo si su función de densidad es dada por

$$f(x; a, b) = \frac{1}{B(a, b)(a+b)^{\frac{1}{2}} 2^{a+b-1}} \left(1 + \frac{x}{\sqrt{a+b+x^2}}\right)^{a+\frac{1}{2}} \left(1 - \frac{x}{\sqrt{a+b+x^2}}\right)^{b+\frac{1}{2}} \quad (2.7)$$

donde $B(\cdot)$ es la función beta. Denominamos a esta parametrización la *parametrización original de la distribución skew-t 3*.

Jones y Faddy (2003) proponen una parametrización alterna basada en Prentice (1975) para facilitar la construcción de la función de verosimilitud; presentamos esta parametrización a continuación.

Definición 6. La *parametrización alterna* de la función de densidad de una variable aleatoria con distribución skew-t 3 es dada por

$$\begin{aligned} q &= \frac{a-b}{\sqrt{ab(a+b)}} & \Leftrightarrow & \quad a = \frac{\frac{2}{p} + \frac{2q}{p\sqrt{2p+q^2}}}{2} \\ p &= \frac{2}{a+b} & & \quad b = \frac{\frac{2}{p} - \frac{2q}{p\sqrt{2p+q^2}}}{2} \end{aligned}$$

Observamos en la Sección 3.5 que la parametrización alterna cuenta con limitaciones importantes para capturar en un parámetro la asimetría de la distribución. En este contexto, construimos una parametrización denominada *parametrización propuesta* la cual definimos a continuación.

Definición 7. La *parametrización propuesta* de la función de densidad de una variable

aleatoria con distribución skew-t 3, denotada por $X \sim ST_3(\alpha, v)$, es dada por

$$\alpha = \frac{a-b}{a+b} \Leftrightarrow a = \frac{b(\alpha+1)}{1-\alpha} \quad \text{y} \quad b = \frac{a(1-\alpha)}{\alpha+1} \quad (2.8)$$

$$v = 2 \times b^{I_{\alpha>0}} \times a^{1-I_{\alpha \leq 0}} \Leftrightarrow \begin{cases} b = \frac{v}{2} & \text{si } \alpha > 0 \\ a = \frac{v}{2} & \text{si } \alpha \leq 0 \end{cases}$$

2.4.1. Propiedades

En esta sección, describimos las propiedades principales de la distribución skew-t 3, además de presentar gráficos de su fdp. Para las pruebas de estas propiedades ver [Jones y Faddy \(2003\)](#).

P1: Las colas tienen el siguiente comportamiento ([Aas y Haff, 2006](#)):

$$f(x) \sim \text{const} \times |x|^{-2a-1} \quad \text{cuando } x \rightarrow -\infty$$

y

$$f(x) \sim \text{const} \times |x|^{-2b-1} \quad \text{cuando } x \rightarrow +\infty$$

Bajo la parametrización alterna, asumiendo $\alpha > 0$, tenemos

$$f(x) \sim \text{const} \times |x|^{-\frac{v(\alpha+1)}{1-\alpha}-1} \quad \text{cuando } x \rightarrow -\infty$$

y

$$f(x) \sim \text{const} \times |x|^{-v-1} \quad \text{cuando } x \rightarrow +\infty$$

Complementariamente, si $\alpha \leq 0$, tenemos

$$f(x) \sim \text{const} \times |x|^{-v-1} \quad \text{cuando } x \rightarrow -\infty$$

y

$$f(x) \sim \text{const} \times |x|^{-\frac{v(1-\alpha)}{\alpha+1}-1} \quad \text{cuando } x \rightarrow +\infty$$

Por lo tanto, el peso de ambas colas aumenta para valores más pequeños de v .

El peso de la cola más pesada solo es impactado por v pero el parámetro α también impacta el peso la cola más ligera: a menor valor de α la cola es más pesada.

P2: α es el parámetro de asimetría y $|\alpha| < 1$. Cuando $\alpha = 0$, la distribución es simétrica y se reduce a la distribución t -student. Al aumentar el valor de $|\alpha|$ la asimetría aumenta es decir, la masa de la probabilidad se concentra más a un lado de la distribución. Cuando $\alpha > 0$ la distribución es asimétrica a la derecha y para $\alpha < 0$ la distribución es asimétrica a la izquierda.

P3: El parámetro v impacta la asimetría: a menor v se permite mayor asimetría.

P4: $f(x; \alpha, v) = f(-x; -\alpha, v)$.

P5: $F(x; a, b) = I_{x'}(a, b)$, donde $x' = \frac{1}{2} + \frac{x}{2\sqrt{a+b+x^2}}$ y $I_x(\cdot, \cdot)$ es la función incompleta beta.

P6: Los momentos de esta distribución son los siguientes:

$$E(X^r) = \frac{(a+b)^{\frac{r}{2}}}{B(a,b)} \sum_{i=0}^r \binom{r}{i} 2^{-i} (-1)^i B\left(a + \frac{r}{2} - i, b - \frac{r}{2}\right)$$

P7: La media es dada por

$$E(X) = \frac{(a-b)\sqrt{a+b}}{2} \frac{\Gamma\left(a - \frac{1}{2}\right) \Gamma\left(b - \frac{1}{2}\right)}{\Gamma(a+b)}$$

P8: La varianza es dada por

$$\text{Var}(X) = \frac{(a+b)}{4} \frac{(a+b)^2 + a+b-2}{(a-1)(b-1)} - \left(\frac{(a-b)\sqrt{a+b}}{2} \frac{\Gamma\left(a - \frac{1}{2}\right) \Gamma\left(b - \frac{1}{2}\right)}{\Gamma(a+b)} \right)^2$$

A continuación presentamos gráficos que muestran la forma de la fdp, primero bajo la parametrización original y luego para la parametrización propuesta. En la Figura 2.5 observamos en la fdp de la skew-t 3 el impacto de cambiar uno de los parámetros dentro de valores pequeños cuando el otro es fijado a 3. Si $a = b$, la distribución es simétrica y se reduce a la distribución t . Para un valor pequeño y fijo de b , $a < b$, a menor a la asimetría y el peso de la cola izquierda aumentan.

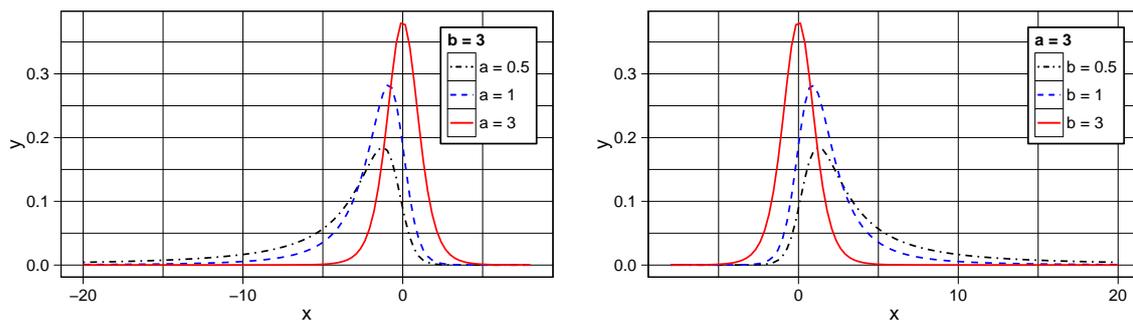


Figura 2.5: Función de densidad de la distribución skew-t 3 b o a igual a 3 y valores pequeños del otro parámetro. Panel izquierdo: $b = 3$ y $a = 0.5, 1, 3$; panel derecho: $a = 3$ y $b = 0.5, 1, 3$.

Por otro lado, vemos en la Figura 2.6 en la fdp de la skew-t 3 el impacto de cambiar uno de los parámetros dentro de valores grandes cuando el otro es fijado a 10. Observamos que si $a = b$, la distribución es simétrica y se reduce a la distribución t . Para un valor b grande y fijo y $a > b$, a mayor a aumenta la asimetría ligeramente.

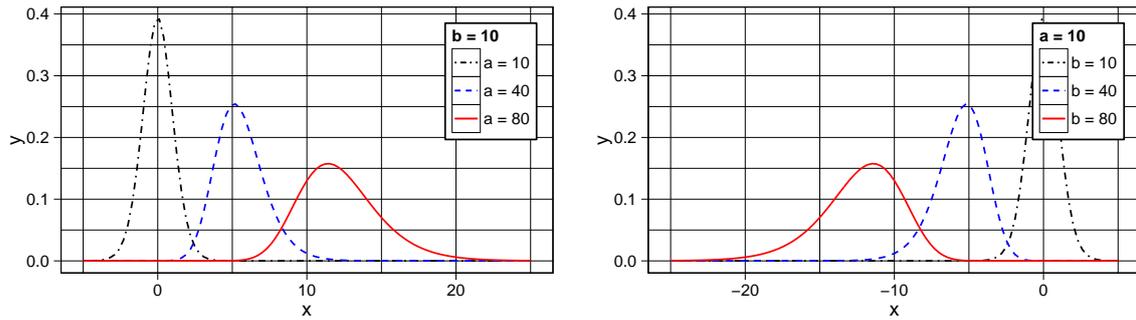


Figura 2.6: Función de densidad de la distribución skew-t 3 b o a igual a 10 y valores grandes del otro parámetro. Panel izquierdo: $b = 10$ y $a = 10, 40, 80$; panel derecho: $a = 10$ y $b = 10, 40, 80$.

Bajo la parametrización propuesta, observamos en la Figura 2.7 el impacto en la fdp de modificar v cuando $\alpha = 0.5$. A menores valores la cola más pesada y la asimetría se incrementa.

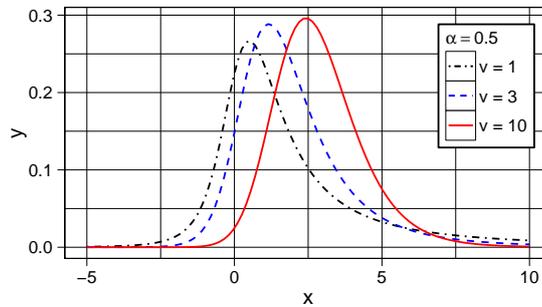


Figura 2.7: Función de densidad de la distribución skew-t 3 $\alpha = 0.5$ y $v = 1, 3, 10$.

Finalmente en la Figura mostramos 2.8 la fdp para diferentes valores de α y $v = 1$. El parámetro α afecta la asimetría dejando fijo el peso de la cola más pesada y modificando el peso de la cola más ligera. También observamos que $f(x, \alpha, v) = f(-x, -\alpha, v)$. Adicionalmente vemos gráficamente que a mayor valor de $|\alpha|$ se tiene mayor nivel de curtosis.

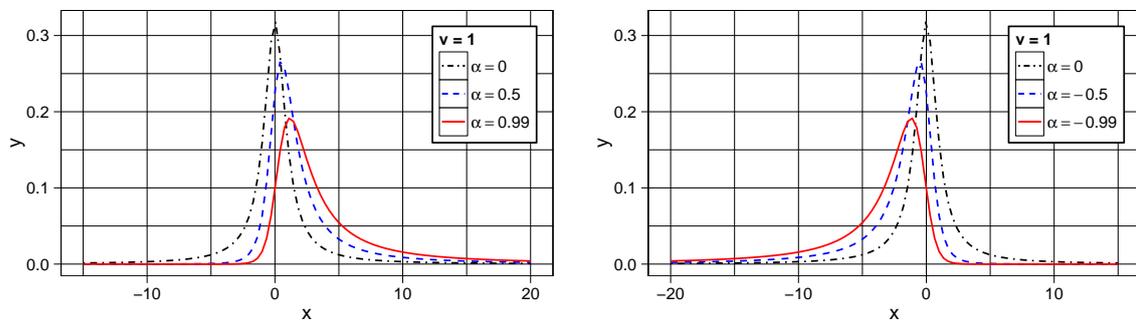


Figura 2.8: Función de densidad de la distribución skew-t 3 para $v = 1$ y diferentes valores de α . El panel izquierdo muestra la fdp para $\alpha = 0, 0.5, 0.99$ y el derecho para $\alpha = 0, -0.5, -0.99$.

Implementación computacional

El paquete *gamlss.dist* (Stasinopoulos et al., 2014) incluye rutinas que permiten calcular la función de densidad, distribución acumulada, cuantiles y generar valores aleatorios que siguen la distribución skew-t 3 y el paquete *gamlss* (Rigby y Stasinopoulos, 2005) contiene una rutina para estimar los parámetros de la distribución skew-t 3 con el método de máxima verosimilitud a través de métodos computacionales.



Capítulo 3

Medidas de asimetría

En este capítulo, analizamos las propiedades de asimetría de las propuestas de distribución skew-t. Para ello, presentamos la medida de asimetría a ser utilizada, en particular la de [Patil et al. \(2014\)](#) y luego, a través de ésta, analizamos el nivel de asimetría que permiten las distribuciones y la relación entre los parámetros y la asimetría.

La Sección 3.1 explica las limitaciones de la medida tradicional de asimetría: el *coeficiente de asimetría de Pearson*, respecto a las propuestas de distribución skew-t. Luego en 3.2 definimos a detalle la medida de asimetría de [Patil et al. \(2014\)](#). Seguido, explicamos en la Sección 3.3 cómo estimamos la asimetría a través de métodos numéricos. Finalmente, la Sección 3.5 presenta la aplicación de la medida de [Patil et al. \(2014\)](#) a cada propuesta de distribución skew-t.

3.1. Limitaciones de la medida de asimetría de Pearson

La medida más conocida de asimetría es el *coeficiente de asimetría de Pearson* (CAP). Como se aprecia en la ecuación (3.1) el CAP solo existe cuando los momentos de orden 1, 2 y 3 de la distribución existen. Lo anterior, implica una limitación importante si deseamos utilizar el CAP para estudiar el nivel de asimetría de las distribuciones skew-t, debido a que los momentos 1, 2 y 3 no siempre existen para las propuestas de distribución skew-t, particularmente cuando las colas son más pesadas.

$$\text{CAP} = E\left(\left(\frac{X - E(X)}{\text{Var}(X)}\right)^3\right) = \frac{E(X^3) - 3E(X)\text{Var}(X) - E^3(X)}{\text{Var}^3(X)} \quad (3.1)$$

Para la distribución skew-t 2, construimos el CAP con las siguientes ecuaciones:

$$\begin{aligned} E(X) &= \frac{2cv(\alpha^2 - 1)}{(v - 1)\alpha}, \quad \text{si } v > 1 \\ \text{Var}(X) &= \frac{v(\alpha^6 + 1)}{(v - 2)\alpha^2(\alpha^2 + 1)} - \frac{4(c v (\alpha^2 - 1))^2}{(\alpha(v - 1))^2}, \quad \text{si } v > 2 \\ E(X^3) &= \frac{2cv^2(\alpha^2 - 1)(\alpha^4 + 1)(2v - 4)}{(v - 3)(v - 1)\alpha^3}, \quad \text{si } v > 3 \end{aligned}$$

donde $c = (v\pi)^{-1/2}\Gamma(v/2)^{-1}\Gamma((v + 1)/2)$.

Mostramos a continuación, para diferentes valores de sus parámetros, el CAP de la distribución skew-t 2 en la Figura 3.1. Observamos que los valores de CAP no tienen una relación monótona respecto a α . Por ejemplo, si consideramos $v = 4$, tenemos que el CAP toma los valores de 0, 1.46 y 0.03 para $\alpha = 0, 1.5$ y 4, respectivamente. Estos resultados contradicen a la Figura 3.2, panel izquierdo, donde la asimetría aumenta a medida que α toma mayores valores. De hecho, se observa que la mayor asimetría se presenta para $\alpha = 4$; sin embargo, el CAP da un valor próximo de 0.

Por otro lado, si consideramos $\alpha = 4$, tenemos que el CAP toma los valores de 0.13, 0.18 y 0.35 para $v = 3.5, 4$ y 5, respectivamente. Estos resultados contradicen lo presentado en la Figura 3.2, panel derecho, donde observamos que v tiene un impacto muy pequeño en la asimetría; sin embargo, el CAP da valores distintos.

En general, se observa visualmente que para la distribución skew-t 2 la asimetría aumenta cuando $|\log(\alpha)|$ aumenta y v tiene un impacto despreciable. Por lo tanto, el CAP falla en capturar la asimetría visual de la distribución skew-t y no es una medida apta para estudiar la asimetría de las distribuciones skew-t.

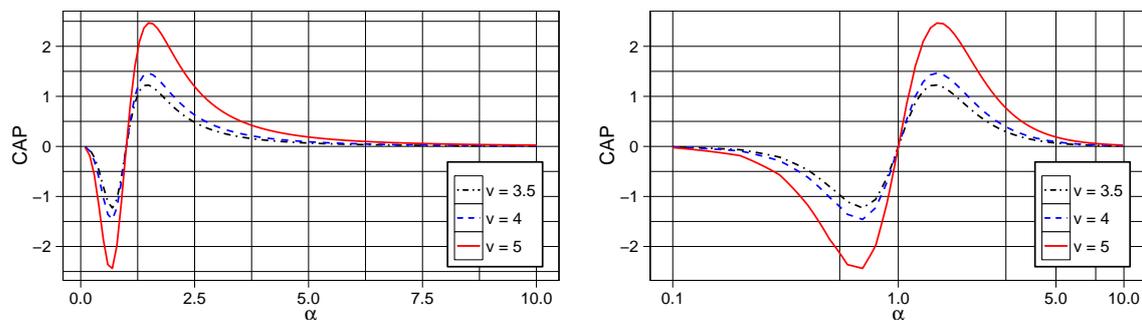


Figura 3.1: CAP de skew-t 2 para diferentes valores de α y v . La figura a la derecha es igual a la de la izquierda pero con escala logarítmica.

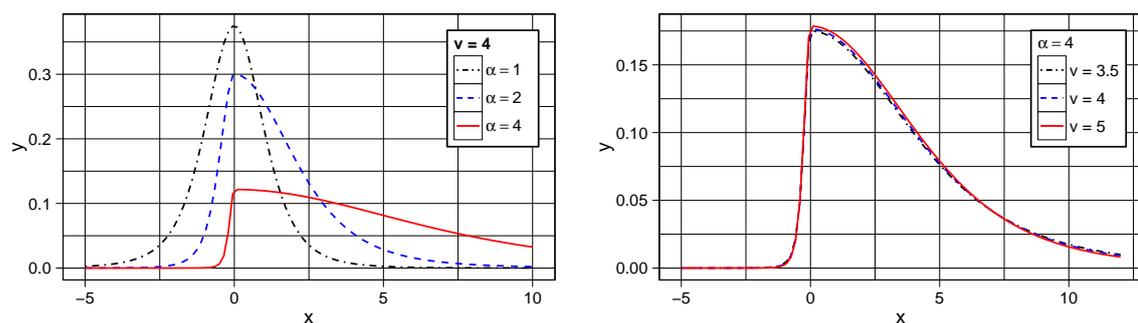


Figura 3.2: Distribución skew-t 2 para diferentes valores de α y v

Es importante mencionar que [Li y Morris \(1991\)](#) y [Patil et al. \(2012\)](#) consideran que el coeficiente de asimetría de Pearson y otras medidas de *skewness* no son apropiadas ya que son influenciadas no trivialmente por las colas de la distribución y pueden no representar la asimetría visual de la fdp.

3.2. Medida de asimetría de Patil et al. (2014)

Para estudiar la asimetría de las distribuciones skew-t, utilizamos la medida de asimetría propuesta por Patil et al. (2014) que es una extensión de la propuesta por Patil et al. (2012). Ambas medidas se basan en calcular la correlación entre la función de densidad de probabilidad (fdp) y la función de densidad acumulada (fda) de la distribución.

La medida de Patil et al. (2014) es resultado de una condición suficiente y necesaria donde hay asimetría si y solo si la medida tiene un valor de 0. Además la medida tiene un rango entre -1 y 1 y, por ende, es acotada y diferencia entre asimetría positiva y negativa. A continuación, presentamos su construcción.

Sea $p \in [1/2, 1)$ y ξ_p tal que $F(\xi_p) = p$, donde $F(x)$ es una fda. Por cada p definimos un par de fdp truncadas $f_p(x)$ y $f_p^*(x)$ tal que

$$f_p(x) = \begin{cases} \frac{f(x)}{p} & \text{si } x \leq \xi_p \\ 0 & \text{si } x > \xi_p \end{cases}$$

y

$$f_p^*(x) = \begin{cases} \frac{f(x)}{p} & \text{si } x \geq \xi_{1-p} \\ 0 & \text{si } x < \xi_{1-p}. \end{cases}$$

Esto equivale a

$$F_p(x) = \begin{cases} \frac{F(x)}{p} & \text{si } x \leq \xi_p \\ 1 & \text{si } x \geq \xi_p \end{cases}$$

y

$$F_p^*(x) = \begin{cases} 0 & \text{si } x \leq \xi_{1-p} \\ 1 - \frac{1-F(x)}{p} & \text{si } x \geq \xi_{1-p}. \end{cases}$$

Definamos ahora

$$\delta_p = \text{Cov}_{f_p}(f_p(X), F_p(X)) \quad \text{y} \quad \delta_p^* = \text{Cov}_{f_p^*}(f_p^*(X), F_p^*(X)) \quad (3.2)$$

donde para $p = 1$, tenemos $\delta_1^* = \delta_1$.

Una condición suficiente y necesaria para que f sea simétrica es que $\delta_p + \delta_p^* = 0 \quad \forall p \in [1/2, 1)$. A continuación formalizamos la aseveración anterior y mostramos su prueba.

Proposición 1. Sea X una variable aleatoria con fdp $f(x)$ y fda $F(x)$ tal que $\int_{-\infty}^{\infty} f^2(x)dx < +\infty$. Si se cumple que, dentro del interior de su soporte, $f(x)$ es (i) continuo y (ii) positivo en todo punto en el enfoque de Lebesgue, entonces $f(x)$ es simétrico si y solo si para todo $p \in [1/2, 1)$,

$$\delta_p + \delta_p^* = 0. \quad (3.3)$$

Demostración. Inicialmente es conveniente escribir (3.3) en términos de la fdp y fda de la variable aleatoria. Notemos que para una variable aleatoria $X \sim f_p(X)$, $\delta_p = \text{Cov}_{f_p}(f_p(X), F_p(X)) = E(f_p(X)F_p(X)) - E(f_p(X))E(F_p(X))$ y se tiene las siguientes relaciones

$$E(f_p(X)F_p(X)) = \int_{-\infty}^{\infty} f_p(x)F_p(x)f_p(x)dx = \frac{1}{p^3} \int_{-\infty}^{\xi_p} F(x)f^2(x)dx,$$

$$E(f_p(X)) = \int_{-\infty}^{\infty} f_p(x)f_p(x)dx = \frac{1}{p^2} \int_{-\infty}^{\xi_p} f^2(x)dx,$$

$$F_p(X) \sim U(0, 1) \quad \text{y} \quad E(F_p(x)) = \frac{1}{2}.$$

Por lo tanto,

$$\delta_p = \frac{1}{p^3} \int_{-\infty}^{\xi_p} F(x)f^2(x)dx - \frac{1}{2p^2} \int_{-\infty}^{\xi_p} f^2(x)dx.$$

Similarmente, encontramos que

$$\delta_p^* = -\frac{1}{p^3} \int_{\xi_{1-p}}^{\infty} f^2(x)(1 - F(x))dx + \frac{1}{2p^2} \int_{\xi_{1-p}}^{\infty} f^2(x)dx.$$

Entonces $\delta_p + \delta_p^* = 0$ equivale a

$$\int_{-\infty}^{\xi_p} F(x)f^2(x)dx - \frac{p}{2} \int_{-\infty}^{\xi_p} f^2(x)dx - \int_{\xi_{1-p}}^{\infty} f^2(x)(1 - F(x))dx + \frac{p}{2} \int_{\xi_{1-p}}^{\infty} f^2(x)dx = 0 \quad (3.4)$$

Primero probamos que si $f(x)$ es simétrica entonces se cumple (3.4) y por ende que $\delta_p + \delta_p^* = 0$. Sin perdida de generalidad, asumimos que $f(x)$ esta centrada alrededor de 0: $\forall x \in \mathbb{R} f(x) = f(-x)$ y $F(-x) = 1 - F(x)$ lo cual implica $\xi_p = -\xi_{1-p}$, $\forall x \in \mathbb{R}$. Por ende tenemos

$$\int_{-\infty}^{\xi_p} f^2(x)dx = \int_{-\infty}^{-\xi_{1-p}} f^2(-x)dx = \int_{\xi_{1-p}}^{\infty} f^2(x)dx.$$

Lo anterior implica que $\frac{p}{2} \int_{-\infty}^{\xi_p} f^2(x)dx = \frac{p}{2} \int_{\xi_{1-p}}^{\infty} f^2(x)dx$ y

$\int_{-\infty}^{\xi_p} F(x)f^2(x) = \int_{\xi_{1-p}}^{\infty} f^2(x)(1 - F(x))dx$. Por lo tanto se cumple (3.4) y $\delta_p + \delta_p^* = 0$. De la misma manera, mostramos que si $\delta_p + \delta_p^* = 0$ entonces $f(x)$ es simétrica. Derivamos la parte izquierda de (3.4) con respecto a p .

$$\frac{p}{2} (f(\xi_p) - f(\xi_{1-p})) + \frac{1}{2} \left(\int_{-\infty}^{\xi_p} f^2(x)dx - \int_{\xi_{1-p}}^{\infty} f^2(x)dx \right) = 0$$

o equivalentemente

$$p \frac{d}{dp} g(p) + g(p) = 0, \quad (3.5)$$

donde

$$g(p) = \int_{-\infty}^{\xi_p} f^2(x)dx - \int_{\xi_{1-p}}^{\infty} f^2(x)dx$$

Notemos que la solución general de la ecuación diferencial (3.5) es

$$g(p) = \frac{c}{p}$$

donde c es un constante y se debe cumplir en particular que $g\left(\frac{1}{2}\right) = 0$. Lo anterior implica que $c = 0$ y por lo tanto que $g(p) = 0$. Entonces se tiene,

$$\int_{-\infty}^{\xi_p} f^2(x)dx = \int_{\xi_{1-p}}^{\infty} f^2(x)dx. \quad (3.6)$$

Por definición de simetría, $f(a+x) = f(a-x) \forall x$, donde $f(x)$ es centrada alrededor de a . Probar que se cumple lo anterior equivale a probar que $f(\xi_p) = f(\xi_{1-p}) \forall p \in [1/2, 1)$.

Para cada $p \in (1/2, 1)$, definimos una secuencia de números $S = (p_j)_{j=1}^{\infty} \subset (1/2, 1)$ tal que se cumplan las siguiente propiedades:

- $f(\xi_{p_j}) > 0, \forall p_j \in S$
- $f(\xi_{1-p_j}) > 0, \forall p_j \in S$
- $\lim_{j \rightarrow \infty} p_j = p$

Vemos que al derivar la parte izquierda de la ecuación (3.6) obtenemos lo siguiente,

$$\left. \frac{d}{dp} \left[\int_{-\infty}^{\xi_p} f^2(x)dx \right] \right|_{p=p_j} = f^2(\xi_{p_j}) \times \left. \frac{d\xi_p}{dp} \right|_{p=p_j}.$$

Notemos que

$$F(\xi_p) = p \quad \text{por ende} \quad \frac{dF(\xi_p)}{dp} = f(\xi_p) \times \frac{d\xi_p}{dp} = 1 \quad \text{y} \quad \left. \frac{d\xi_p}{dp} \right|_{p=p_j} = \frac{1}{f(\xi_{p_j})}.$$

Por ende tenemos

$$\left. \frac{d}{dp} \left[\int_{-\infty}^{\xi_p} f^2(x)dx \right] \right|_{p=p_j} = f^2(\xi_{p_j}) \times \frac{1}{f(\xi_{p_j})} = f(\xi_{p_j}).$$

Similarmente encontramos que al derivar la parte derecha de la ecuación (3.6) obtenemos

$$\left. \frac{d}{dp} \left[\int_{\xi_{1-p}}^{\infty} f^2(x)dx \right] \right|_{p=p_j} = f(\xi_{1-p_j}).$$

Lo anterior muestra que la ecuación (3.6) implica que

$$f(\xi_{p_j}) = f(\xi_{1-p_j}).$$

Por lo tanto

$$\lim_{j \rightarrow \infty} f(\xi_{p_j}) = \lim_{j \rightarrow \infty} f(\xi_{1-p_j}).$$

y

$$f(\xi_p) = f(\xi_{1-p}).$$

como requerimos. □

Cabe notar que, como indican los autores, se cumple $\delta_p + \delta_p^* < 0$ para casos de asimetría positiva y $\delta_p + \delta_p^* > 0$ para casos de asimetría negativa. Como la escala afecta el valor de $\delta_p + \delta_p^*$ los autores proponen estandarizar estos valores con:

$$\rho_p = \text{corr}_{f_p}(f_p(X), F_p(X)) \quad \text{y} \quad \rho_p^* = \text{corr}_{f_p^*}(f_p^*(X), F_p^*(X)) \quad p \in [1/2, 1) \quad (3.7)$$

$$\text{y} \quad \rho_1 = \rho_1^* = \text{corr}_f(f(X), F(X))$$

Corolario 1. Podemos formular las relaciones anteriores en términos de la función de densidad original de la siguiente manera

$$\rho_p = \frac{2\sqrt{3}}{p} \frac{\int_{-\infty}^{\xi_p} f^2(x)F(x)dx - \frac{p}{2} \int_{-\infty}^{\xi_p} f^2(x)dx}{\left(p \int_{-\infty}^{\xi_p} f^3(x)dx - \left(\int_{-\infty}^{\xi_p} f^2(x)dx\right)^2\right)^{\frac{1}{2}}} \quad (3.8)$$

y

$$\rho_p^* = \frac{2\sqrt{3}}{p} \frac{\int_{\xi_{1-p}}^{\infty} f^2(x)(1-F(x))dx + \frac{p}{2} \int_{\xi_{1-p}}^{\infty} f^2(x)dx}{\left(p \int_{\xi_{1-p}}^{\infty} f^3(x)dx - \left(\int_{\xi_{1-p}}^{\infty} f^2(x)dx\right)^2\right)^{\frac{1}{2}}} \quad (3.9)$$

Demostración. Notemos que

$$\rho_p = \text{corr}_{f_p}(f_p(X), F_p(X)) = \frac{\text{Cov}(f_p(X), F_p(X))}{\left(\text{Var}(f_p(X))\text{Var}(F_p(X))\right)^{\frac{1}{2}}},$$

$$\text{Var}(f_p(X)) = \text{E}(f_p(X)^2) - \text{E}^2(f_p(X)) = \frac{1}{p^3} \int_{-\infty}^{\xi_p} f^3(x)dx - \left(\frac{1}{p^2} \int_{-\infty}^{\xi_p} f^2(x)dx\right)^2,$$

$$F(X) \sim U(0, 1) \quad \text{y} \quad \text{Var}(F_p(x)) = \frac{1}{12}.$$

Con las formulas anteriores y lo obtenido en (3.2) la prueba de (3.8) es directa. El mismo argumento se puede utilizar para obtener (3.9) □

Definición 8. Definimos como medida de asimetría de Patil et al. (2014), denotada $\eta(X)$ por la siguiente relación:

$$\eta(X) = -\frac{1}{2} \text{signo}(\rho_1) \max_{\frac{1}{2} \leq p \leq 1} |\rho_p + \rho_p^*|, \quad (3.10)$$

donde tenemos como condición necesaria para la existencia de $\eta(X)$ que $\int_{-\infty}^{\infty} f^3(x)dx < +\infty$

Observación 1. Por el corolario 1 observamos que $f(x)$ es simétrica si y solo si $\rho_p + \rho_p^* = 0$ para todo $p \in [1/2, 1)$. Por otro lado, el valor máximo teórico de $|\rho_p + \rho_p^*|$ es 2. Finalmente $-\text{signo}(\rho_1)$ es positivo para asimetría positiva y negativo para asimetría negativa.

Tomando en cuenta lo anterior, $\eta(X)$ es 0 si y solo si $f(x)$ es simétrica, tiene como valor máximo teórico 1 y como mínimo teórico -1 , es positivo si $f(x)$ cuenta con asimetría positiva y es negativo si $f(x)$ cuenta con asimetría negativa.

3.2.1. Relación con la medida Patil et al. (2012)

Definición 9. La medida de asimetría de Patil et al. (2012) es definida como

$$\eta_w(X) = -\text{corr}_f(f(x), F(x)) \tag{3.11}$$

Observamos que $\eta(X) = \eta_w(X)$ solo si se maximiza $\max_{\frac{1}{2} \leq p \leq 1} |\rho_p + \rho_p^*|$ cuando $p = 1$. Por lo tanto la medida de asimetría de Patil et al. (2014) es una generalización de la medida de asimetría de Patil et al. (2012).

$\eta_w(X)$ tiene la debilidad de ser necesaria pero no suficiente para determinar si una distribución es simétrica. Los autores dan el ejemplo de una distribución Weibull con parámetro de escala σ y de forma α . Para $\alpha = \left(2 - \frac{\ln(2)}{\ln(1.5)}\right)^{-1}$ tenemos $\eta_w(X) = 0$ pero la distribución no es estrictamente simétrica. Por otro lado, $\eta(X)$ es diferente a 0.

3.2.2. Propiedades de $\eta(\cdot)$

A continuación describimos algunas de las propiedades principales de $\eta(\cdot)$. Para mayor información y las pruebas ver Patil et al. (2014).

P1: Para una variable aleatoria simétrica X , $\eta(X) = 0$

P2: Si $Y = aX + b$ donde $a > 0$ y $b \in \mathbb{R}$, entonces $\eta(X) = \eta(Y)$

P3: Si $Y = -X$, entonces $\eta(X) = -\eta(Y)$

P4: $|\eta_w(X)| \leq |\eta(X)|$

P5: $|\eta(X)| \leq 1$

3.3. Método de estimación numérica de la medida de asimetría de Patil et al. (2014)

Aplicar la medida de asimetría de Patil et al. (2014) a las propuestas de distribución skew-t no es inmediato ya que la medida asume que podemos calcular directamente las dos siguientes correlación teóricas $\text{corr}_{f_p}(f_p(X), F_p(X))$ y $\text{corr}_{f_p^*}(f_p^*(X), F_p^*(X))$. Dentro del

contexto de las ecuaciones (3.8) y (3.9) y de las fdp de las skew-t, es necesario calcular integrales que no pueden ser expresadas como combinaciones de funciones elementales. Por lo tanto, en lugar de calcular analíticamente los valores teóricos, los estimamos a través de métodos de Monte Carlo. En particular, a través de la simulación de vectores aleatorios con las distribuciones estudiadas y los estimadores de las estadísticas que construyen la medida de asimetría.

Los pasos a seguir para estimar la medida son los siguientes:

1. Generar n observaciones X_1, \dots, X_n de una variable aleatoria X con propuesta de distribución skew-t
2. Calcular ρ_p y ρ_p^* con la función de correlación muestral para varios valores de $p \in [1/2, 1)$.
3. Calcular η para los valores del paso anterior.

Para ilustrar este proceso, presentamos en la Figura 3.3 $f_p(x_i)$ y $F_p(x_i)$ para 10,000 observaciones con distribución skew-t 1 con $v = 1$ y $\alpha = -10, 0, 10$, donde $p = 1$, es el valor que, para estos parámetros, maximiza η . Como parte del cálculo de η para cada curva se calcula la correlación para $f(X_i)_p$ y $F(X_i)_p$ para todos los valores simulados. Considerando lo anterior, se ve como para $\alpha = -10, 0, 10$ obtenemos $\eta = -0.87, 0, 0.87$.

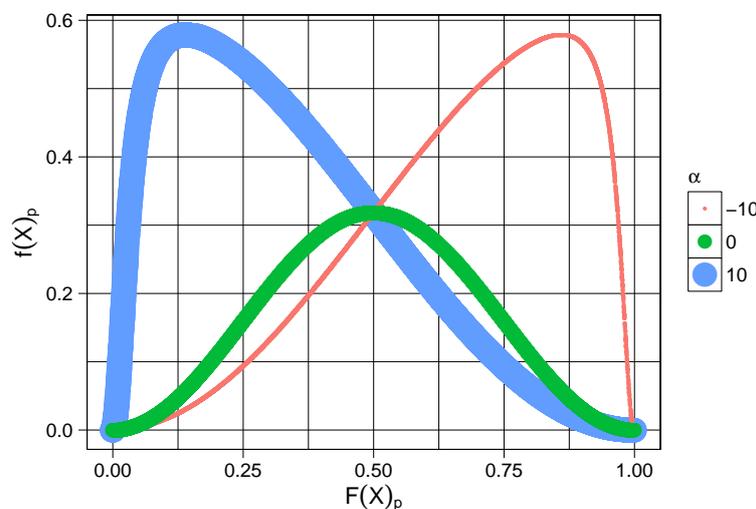


Figura 3.3: $f_p(X)$ y $F_p(X)$ simulada de la distribución skew-t 1 para ρ que maximiza η .

3.4. Cumplimiento de la condición $\int_{-\infty}^{\infty} f^3(x)dx < +\infty$

Como detallamos en la Sección 3.2 para utilizar la medida de asimetría de Patil et al. (2014) se debe cumplir que $\int_{-\infty}^{\infty} f^3(x)dx < +\infty$, donde $f(\cdot)$ es la fdp. En esta sección probamos que las tres propuestas de skew-t cumplen con esta condición.

Para la skew-t 1 y 2 utilizamos la propiedad que la integral de la distribución t-student sobre toda la recta real es igual a uno, esto es,

$$\int_{-\infty}^{\infty} \frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}} dx = 1, \quad \text{donde } v \in \mathbb{R}^+.$$

Equivalentemente,

$$\int_{-\infty}^{\infty} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}} dx = \frac{\sqrt{\pi v} \Gamma(\frac{1}{2}v)}{\Gamma(\frac{1}{2}(v+1))}, \quad \text{donde } v \in \mathbb{R}^+.$$

Para la skew-t 3 utilizamos la propiedad que la integral de la distribución skew-3 sobre toda la recta real es igual a uno, esto es

$$\int_{-\infty}^{\infty} \frac{1}{B(a,b)(a+b)^{\frac{1}{2}} 2^{a+b-1}} \left(1 + \frac{x}{\sqrt{a+b+x^2}}\right)^{a+\frac{1}{2}} \left(1 - \frac{x}{\sqrt{a+b+x^2}}\right)^{b+\frac{1}{2}} dx = 1,$$

donde $a > 0$ y $b > 0$. Equivalentemente,

$$\int_{-\infty}^{\infty} \left(1 + \frac{x}{\sqrt{a+b+x^2}}\right)^{a+\frac{1}{2}} \left(1 - \frac{x}{\sqrt{a+b+x^2}}\right)^{b+\frac{1}{2}} dx = B(a,b)(a+b)^{\frac{1}{2}} 2^{a+b-1},$$

donde $a > 0$ y $b > 0$.

3.4.1. Cumplimiento en la skew-t 1

$$\begin{aligned} \int_{-\infty}^{\infty} f^3(x) dx &= \int_{-\infty}^{\infty} \left(2 t(x;v) T\left(\alpha x \sqrt{\frac{v+1}{v+x^2}}; v+1\right)\right)^3 dx \\ &< 8 \int_{-\infty}^{\infty} (t(x;v))^3 dx \\ &= 8 \int_{-\infty}^{\infty} \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}}\right)^3 dx \\ &= 8 \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)}\right)^3 \int_{-\infty}^{\infty} \left(1 + \frac{x^2}{v}\right)^{-\frac{(3v+2)+1}{2}} dx \end{aligned}$$

Haciendo la substitución de variable $x = y\sqrt{\frac{v}{3v+2}}$ en el último término tenemos $dx =$

$\sqrt{\frac{v}{3v+2}} dy$ y podemos escribir la integral como

$$\begin{aligned} 8 \int_{-\infty}^{\infty} (t(x; v))^3 dx &= 8 \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \right)^3 \sqrt{\frac{v}{3v+2}} \int_{-\infty}^{\infty} \left(1 + \frac{y^2}{3v+2} \right)^{-\frac{(3v+2)+1}{2}} dy \\ &= 8 \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \right)^3 \sqrt{\frac{v}{3v+2}} \frac{\sqrt{\pi(3v+2)} \Gamma(\frac{1}{2}3v+2)}{\Gamma(\frac{1}{2}(3v+3))} < +\infty \end{aligned}$$

3.4.2. Cumplimiento en la skew-t 2

Suponiendo, sin pérdida de generalidad, que $\alpha \geq 1$

$$\begin{aligned} \int_{-\infty}^{\infty} f^3(x) dx &= \int_{-\infty}^{\infty} \left(\frac{2}{\alpha + \frac{1}{\alpha}} \right)^3 \left(t(\alpha x; v) I(x < 0) + t\left(\frac{x}{\alpha}; v\right) I(x \geq 0) \right)^3 dx \\ &\leq \left(\frac{2}{\alpha + \frac{1}{\alpha}} \right)^3 \int_{-\infty}^{\infty} \left(t\left(\frac{x}{\alpha}; v\right) \right)^3 dx \\ &= \left(\frac{2}{\alpha + \frac{1}{\alpha}} \right)^3 \int_{-\infty}^{\infty} \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \left(1 + \frac{x^2}{\alpha^2 v} \right)^{-\frac{v+1}{2}} \right)^3 dx \\ &= \left(\frac{2}{\alpha + \frac{1}{\alpha}} \right)^3 \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \right)^3 \int_{-\infty}^{\infty} \left(1 + \frac{x^2}{\alpha^2 v} \right)^{-\frac{(3v+2)+1}{2}} dx \end{aligned}$$

Haciendo la substitución de variable $x = y\alpha\sqrt{\frac{v}{3v+2}}$ tenemos $dx = \alpha\sqrt{\frac{v}{3v+2}} dy$ y el último término podemos expresar como

$$\begin{aligned} &= \left(\frac{2}{\alpha + \frac{1}{\alpha}} \right)^3 \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \right)^3 \alpha \sqrt{\frac{v}{3v+2}} \int_{-\infty}^{\infty} \left(1 + \frac{y^2}{3v+2} \right)^{-\frac{(3v+2)+1}{2}} dy \\ &= \left(\frac{2}{\alpha + \frac{1}{\alpha}} \right)^3 \alpha \left(\frac{\Gamma(\frac{1}{2}(v+1))}{\sqrt{\pi v} \Gamma(\frac{1}{2}v)} \right)^3 \sqrt{\frac{v}{3v+2}} \frac{\sqrt{\pi(3v+2)} \Gamma(\frac{1}{2}3v+2)}{\Gamma(\frac{1}{2}(3v+3))} < +\infty \end{aligned}$$

3.4.3. Cumplimiento en la skew-t 3

$$\begin{aligned} \int_{-\infty}^{\infty} f^3(x) dx &= c_1 \int_{-\infty}^{\infty} \left(\left(1 + \frac{x}{\sqrt{a+b+x^2}} \right)^{a+\frac{1}{2}} \left(1 - \frac{x}{\sqrt{a+b+x^2}} \right)^{b+\frac{1}{2}} \right)^3 dx \\ &= c_1 \int_{-\infty}^{\infty} \left(1 + \frac{x}{\sqrt{a+b+x^2}} \right)^{(3a+1)\frac{1}{2}} \left(1 - \frac{x}{\sqrt{a+b+x^2}} \right)^{(3b+1)\frac{1}{2}} dx \end{aligned}$$

donde $c_1 = \frac{1}{B(a, b)(a+b)^{\frac{1}{2}} 2^{a+b-1}}$, una constante. Tomando $x = y\sqrt{\frac{a+b}{3a+3b+2}}$ tenemos

$$dx = \sqrt{\frac{a+b}{3a+3b+2}} dy \text{ y}$$

$$\begin{aligned} \frac{x}{\sqrt{a+b+x^2}} &= \frac{y\sqrt{\frac{a+b}{3a+3b+2}}}{\sqrt{a+b+y^2\frac{a+b}{3a+3b+2}}} \\ &= \frac{y}{\sqrt{\frac{3a+3b+2}{a+b}}\sqrt{a+b+y^2\frac{a+b}{3a+3b+2}}} \\ &= \frac{y}{\sqrt{3a+3b+2+y^2}} \end{aligned}$$

Así

$$\begin{aligned} &\int_{-\infty}^{\infty} f^3(x) dx \\ &= c_1 \int_{-\infty}^{\infty} \left(1 + \frac{y}{\sqrt{3a+3b+2y^2}}\right)^{(3a+1)\frac{1}{2}} \left(1 - \frac{y}{\sqrt{3a+3b+2+y^2}}\right)^{(3b+1)\frac{1}{2}} \sqrt{\frac{a+b}{3a+3b+2}} dy \\ &= \left(\frac{1}{B(a,b)(a+b)^{\frac{1}{2}} 2^{a+b-1}}\right)^3 \times \frac{a+b}{3a+3b+2} \times B(3a+1, 3b+1)(3a+3b+2)^{\frac{1}{2}} 2^{3a+3b+1} \\ &< +\infty \end{aligned}$$

3.5. Asimetría en las distribuciones

En esta sección presentamos los resultados de aplicar la medida de asimetría de [Patil et al. \(2014\)](#) a través de procesos numéricos a las tres propuestas de distribución skew-t. En particular, utilizando la metodología descrita en la Sección 3.3 con 51 valores de p y 2,000,000 valores de muestra.

Analizamos el nivel de asimetría que permite cada distribución, al igual que la relación entre la asimetría y los parámetros que no son parámetros de posición o dispersión (esto es, excluyendo μ y σ del análisis).

3.5.1. Asimetría en la skew-t 1

Presentamos en la Figura 3.4 el nivel de asimetría para la distribución skew-t 1 para diferentes valores de α y v . Observamos que la distribución aparenta permitir el rango completo de asimetría de entre -1 y 1 . Por otro lado, el parámetro que impacta principalmente la asimetría es α y presenta una aparente relación monótonica con la asimetría. v impacta la asimetría de manera mínima e inversa (a menor v la asimetría es ligeramente mayor).

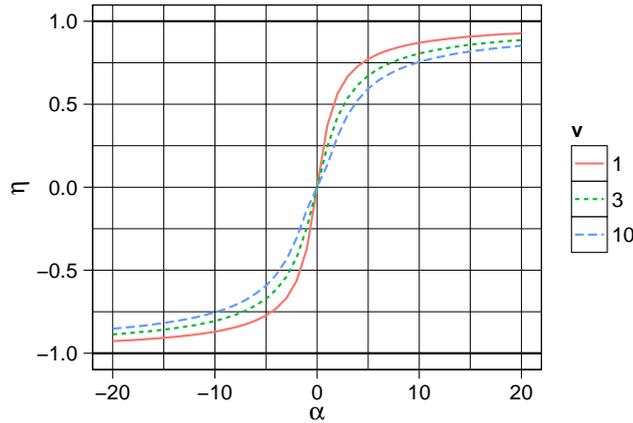


Figura 3.4: Asimetría en skew-t 1 para diferentes valores de α y v .

3.5.2. Asimetría en la skew-t 2

Presentamos en la Figura 3.5 el nivel de asimetría para la distribución skew-t 2 para diferentes valores de α y v . Observamos que la distribución aparenta permitir el rango completo de asimetría de entre -1 y 1 . El parámetro que impacta principalmente la asimetría es α y presenta una aparente relación monótonica con la asimetría. v impacta la asimetría de manera casi insignificante e inversa (a menor v la asimetría es ligeramente mayor).

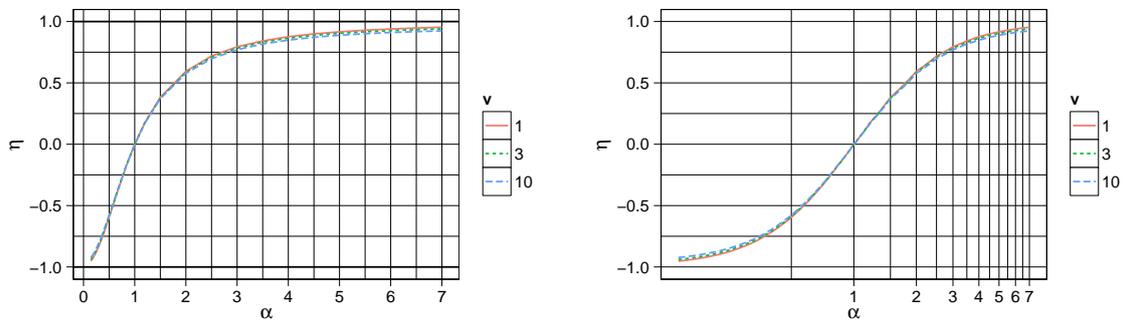


Figura 3.5: Asimetría en skew-t 2 para diferentes valores de α y v . El panel izquierdo con escala lineal para α y panel derecho con escala logarítmica para α .

3.5.3. Asimetría en la skew-t 3

Presentamos en la Figura 3.6 el nivel de asimetría para la distribución skew-t 3 sobre la base de la parametrización original para diferentes valores de a y b . Observamos que al fijar b , por lo general, se tiene mayor asimetría para valores pequeños de a . Para mayores valores de a , la asimetría aumenta pero solo hasta un nivel limitado. Esto implica que si las dos colas de la distribución son ligeras, no se puede tener alta asimetría y por ende no se puede separar el peso de las colas con la asimetría.

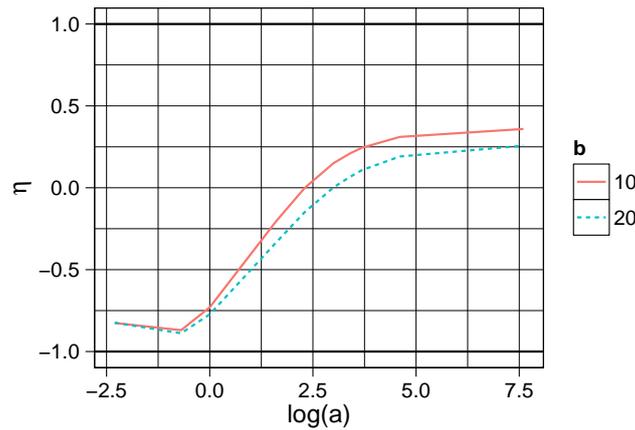


Figura 3.6: Asimetría en skew-t 3 para diferentes valores de a y b .

Por otro lado, la Figura 3.7 muestra el nivel de asimetría bajo la parametrización alterna de Jones y Faddy (2003). Se observa que hay una relación no monótona entre la asimetría y el parámetro de asimetría q . Esto sugiere que esta parametrización no permite capturar bien la asimetría a través los parámetros p y q .

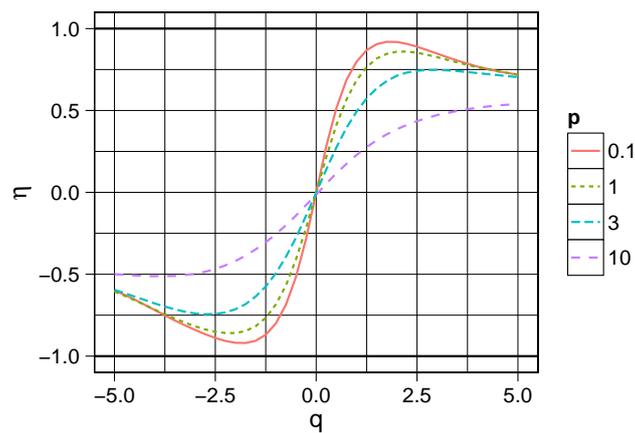


Figura 3.7: Asimetría en skew-t 3 para diferentes valores de q y p .

Finalmente presentamos el nivel de asimetría para diferentes valores de α y v bajo la parametrización propuesta en la Figura 3.8. Observamos que la distribución aparenta permitir el rango completo de asimetría de entre -1 y 1 . El parámetro que impacta principalmente la asimetría es α y presenta una aparente relación monótonica con la asimetría. v tiene un impacto importante en la asimetría. A menor valor de v la asimetría es mayor.

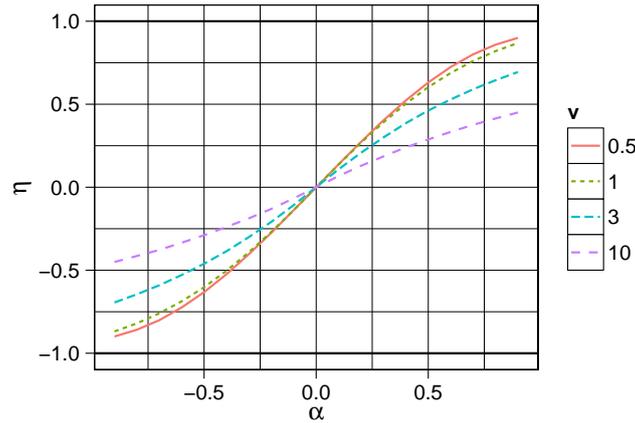


Figura 3.8: Asimetría en skew-t 3 para diferentes valores de α y v .

3.5.4. Conclusion

Tomando en cuenta los análisis previos se concluye lo siguiente:

- Las tres distribuciones permiten un rango de asimetría cercano al máximo de η entre -1 y 1.
- El parámetro que principalmente impacta la asimetría para los tres casos es α
- Para menores valores de v la asimetría es mayor en las tres distribuciones. No obstante, el impacto es mayor para la skew-t 3, pequeño en la skew-t 1 e insignificante en la skew-t 2.
- La distribución skew-t 2 cuenta con la desventaja que si sus dos colas son ligeras no permite altos niveles de asimetría.

Capítulo 4

Aplicaciones

Este Capítulo presenta los resultados de estimar las tres propuestas de distribución skew-t a un conjunto de datos reales; siendo ésta, la depreciación logarítmica de la tasa de cambio de soles a dólares para diferentes periodos de 3, 6, 9 y 12 meses. Este conjunto de datos es modelado con las tres propuestas de distribución skew-t al igual que las skew normal 1, t-student y normal. En la Sección 4.1 describimos los datos y en la Sección 4.2 mostramos los resultados de la aplicación y comparamos el ajuste de las diferentes distribuciones utilizando el Criterio de Información de Akaike (AIC).

4.1. Descripción y características de los datos

4.1.1. Datos de tipo de cambio

Los datos se construyen con la tasa de cambio nominal diaria de Nuevos Soles Peruanos (PEN) a Dólares Americanos (USD) para el periodo comprendido entre el 2 de febrero de 1997 al 31 de marzo de 2015 (BCRP, 2015). La Figura 4.1 muestra esta serie de tiempo.

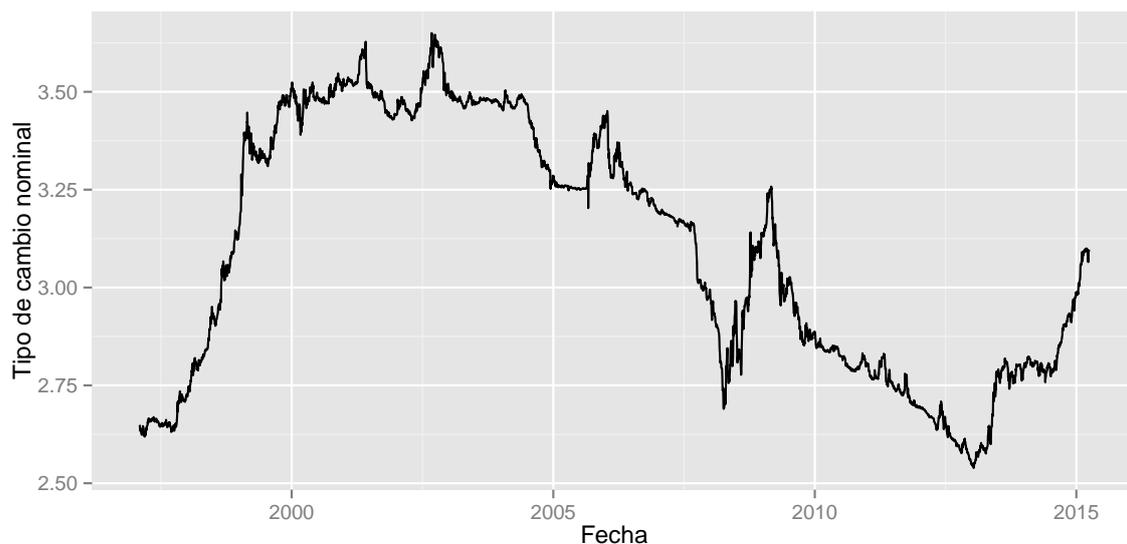


Figura 4.1: Serie de tiempo del tipo de cambio PEN/USD entre 02/02/1997 y 31/03/2015.

Observamos la distribución de estos datos en la Figura 4.2. La tasa de cambio tiene un rango de entre 2.539 y 3.65 y presenta dos modas. Complementariamente, en el Cuadro 4.1 presentamos varias estadísticas resumen de la tasa de cambio.

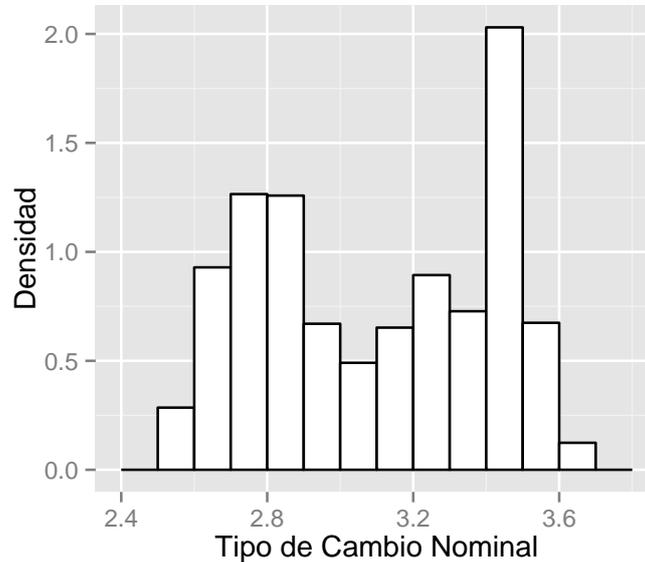


Figura 4.2: Histograma de tipo de cambio nominal.

Estadístico	Valor
Número de observaciones	4522
Fecha inicio	1997-02-03
Fecha fin	2015-03-31
Tipo de cambio inicial	2.649
Tipo de cambio final	3.097
Tipo de cambio min	2.539
Tipo de cambio max	3.65
Tipo de cambio promedio	3.1
Tipo de cambio medio	3.127
Desviación estándar	0.32

Cuadro 4.1: Estadísticas resumen

4.1.2. Depreciación logarítmica

Transformamos el conjunto de datos descrito en la Sección 4.1.1 a retornos logarítmicos de la tasa de cambio. El retorno logarítmico es un indicador financiero que se define de la siguiente manera:

$$R = \log \left(\frac{V_{t+n}}{V_t} \right). \tag{4.1}$$

Cabe notar que para el caso de la tasa de cambio el retorno logarítmico es equivalente a la depreciación logarítmica de la moneda.

Esta transformación permite observar la depreciación del PEN respecto al USD para diferentes periodos de tiempo. Aplicamos el logaritmo natural al ratio, como mostramos en (4.1), para que exista en la recta. Por ejemplo, si en una fecha el tipo de cambio PEN a USD es de 3.00 y un año después es de 3.50 sabemos que la moneda se ha depreciado en $3.50/3.00 - 1 = 16.7\%$ y la depreciación logarítmica es de $\log(3.50/3.00) = 0.15$.

Observamos la distribución de la depreciación logarítmica a diferentes trimestres financieros, en particular a 3, 6, 9, 12 meses. Esto permite simular el impacto que tendría el comprar soles en cierta fecha y esperar diferentes periodos antes de venderlo. Tomando estos tiempos comparamos la distribución (nivel de asimetría y puntos *outliers*) de la depreciación logarítmica de la moneda dependiendo del número de trimestres. Las Figuras 4.3, 4.4, 4.5 y 4.6 muestran los histogramas correspondientes.

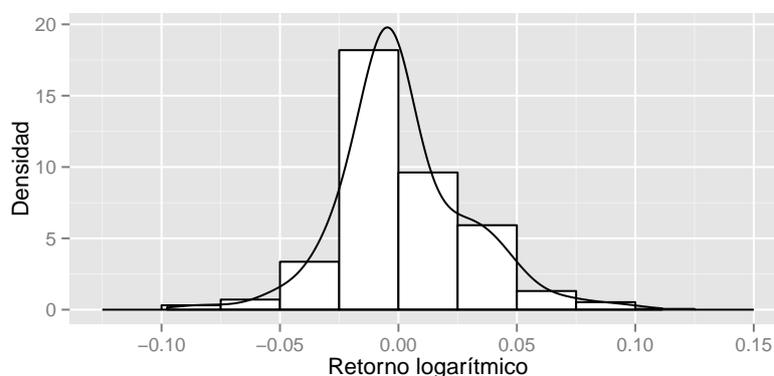


Figura 4.3: Histograma de depreciación logarítmica a 3 meses con estimación de densidad.

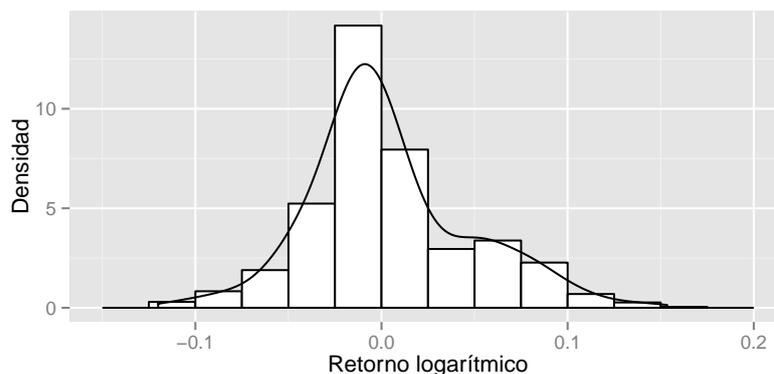


Figura 4.4: Histograma de depreciación logarítmica a 6 meses con estimación de densidad.

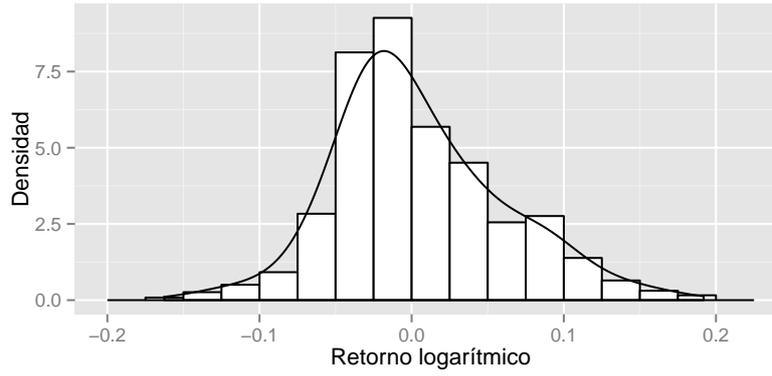


Figura 4.5: Histograma de depreciación logarítmica a 9 meses con estimación de densidad.

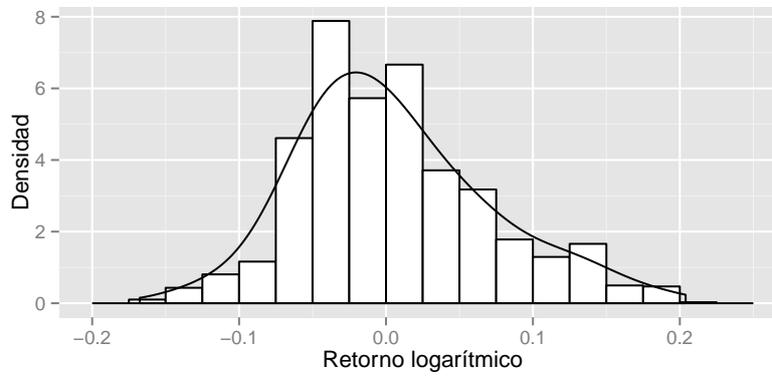
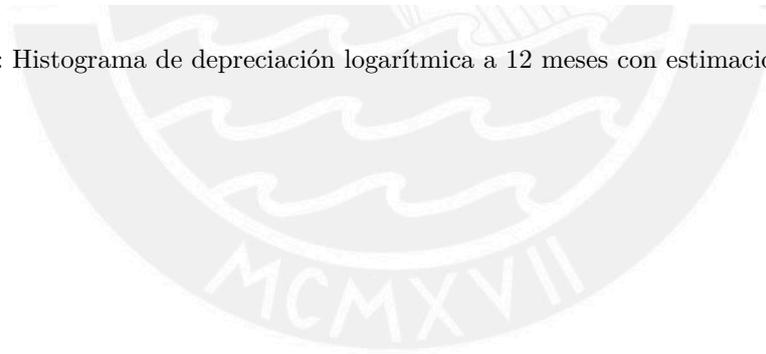


Figura 4.6: Histograma de depreciación logarítmica a 12 meses con estimación de densidad.



4.2. Resultados

Esta sección presenta los resultados de aplicar los métodos de estimación de parámetros referenciados en el Anexo A a los datos de depreciación logarítmico descritos en la sección anterior. Este conjunto de datos es modelado con las tres propuestas de distribución skew-t y así como la skew normal 1, t-student y normal, tomadas como punto de comparación. Para este análisis asumimos independencia entre las observaciones.

Los Cuadros 4.2, 4.3, 4.4 y 4.5 muestran el ajuste medido por AIC y los parámetros estimados de las distribuciones: μ, σ, α, v . Aquí η ha sido estimado con los métodos descritos en la Sección 3.3 para las distribuciones que permiten asimetría. Asimismo, las Figuras 4.7, 4.8, 4.9 y 4.10 muestran las distribuciones estimadas y los histogramas de los datos.

Distribución	AIC	μ	σ	α	v	η
Skew-t 2	-19739.29	-0.01	0.02	1.28	3.30	0.23
Skew-t 1	-19724.56	-0.01	0.02	0.96	3.12	0.23
Skew-t 3	-19697.00	-0.01	0.02	0.19	2.72	0.20
t-student	-19546.30	-0.00	0.02	-	3.57	-
Skew normal 1	-19369.66	-0.02	0.04	1.57	-	0.16
Normal	-19242.26	0.00	0.03	-	-	-

Cuadro 4.2: Parámetros estimados para depreciación logarítmica a 3 meses. Cuando la distribución no cuenta con un parámetro ponemos “-”. Los resultados están ordenados por AIC.

Distribución	AIC	μ	σ	α	v	η
Skew-t 2	-15288.09	-0.01	0.03	1.35	4.46	0.28
Skew-t 1	-15276.88	-0.03	0.04	1.40	4.29	0.28
Skew-t 3	-15246.35	-0.02	0.03	0.26	3.46	0.25
Skew normal 1	-15160.81	-0.04	0.06	2.09	-	0.25
t-student	-15015.20	-0.00	0.03	-	5.16	-
Normal	-14920.79	0.00	0.04	-	-	-

Cuadro 4.3: Parámetros estimados para depreciación logarítmica a 6 meses. Cuando la distribución no cuenta con un parámetro ponemos “-”. Los resultados están ordenados por AIC.

Distribución	AIC	μ	σ	α	v	η
Skew-t 2	-12955.88	-0.02	0.04	1.41	5.97	0.32
Skew-t 1	-12907.34	-0.04	0.06	1.74	6.23	0.30
Skew-t 3	-12862.92	-0.04	0.04	0.30	5.19	0.24
Skew normal 1	-12836.51	-0.05	0.08	2.10	-	0.25
t-student	-12668.82	0.00	0.05	-	9.87	-
Normal	-12627.15	0.00	0.06	-	-	-

Cuadro 4.4: Parámetros estimados para depreciación logarítmica a 9 meses. Cuando la distribución no cuenta con un parámetro ponemos “-”. Los resultados están ordenados por AIC.

Distribución	AIC	μ	σ	α	v	η
Skew-t 2	-11534.70	-0.03	0.05	1.45	12.08	0.34
Skew-t 1	-11515.88	-0.06	0.09	2.34	19.86	0.32
Skew normal 1	-11511.82	-0.07	0.10	2.53	-	0.32
Skew-t 3	-11475.85	-0.12	0.04	0.48	10.25	0.27
t-student	-11226.22	0.00	0.06	-	12.84	-
Normal	-11217.43	0.00	0.06	-	-	-

Cuadro 4.5: Parámetros estimados para depreciación logarítmica a 12 meses. Cuando la distribución no cuenta con un parámetro ponemos “-”. Los resultados están ordenados por AIC.

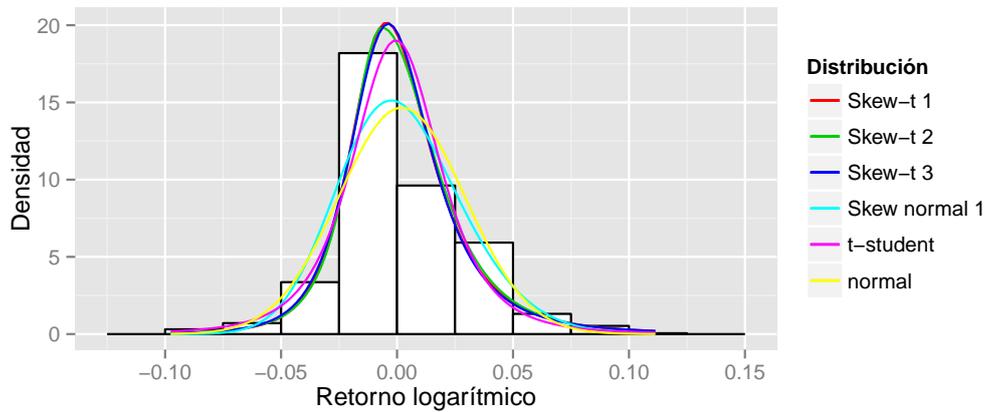


Figura 4.7: Histograma de depreciación logarítmica a 3 meses y fdp estimadas.

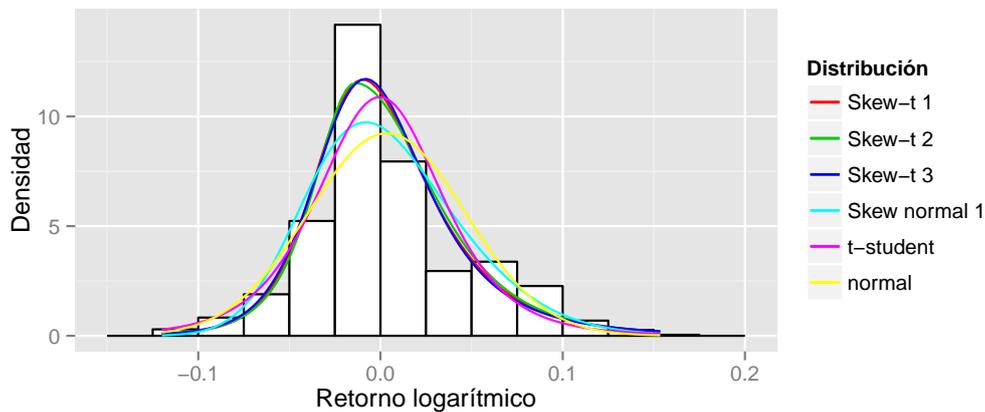


Figura 4.8: Histograma de depreciación logarítmica a 6 meses y fdp estimadas.

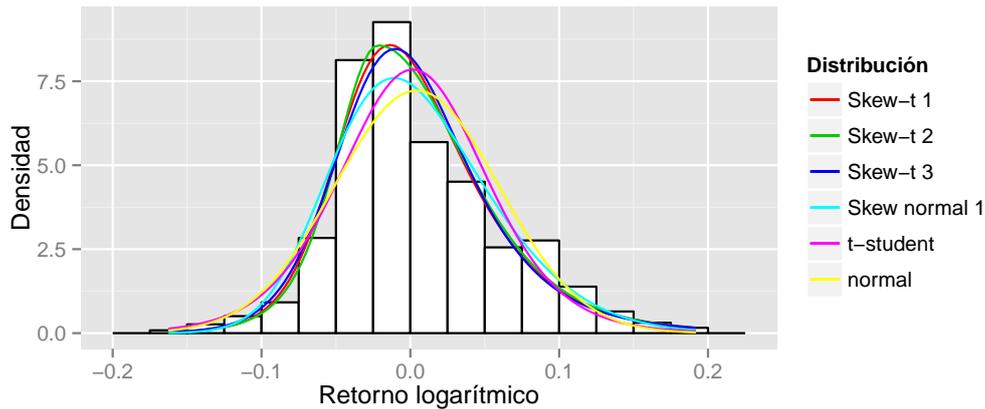


Figura 4.9: Histograma de depreciación logarítmica a 9 meses y fdp estimadas.

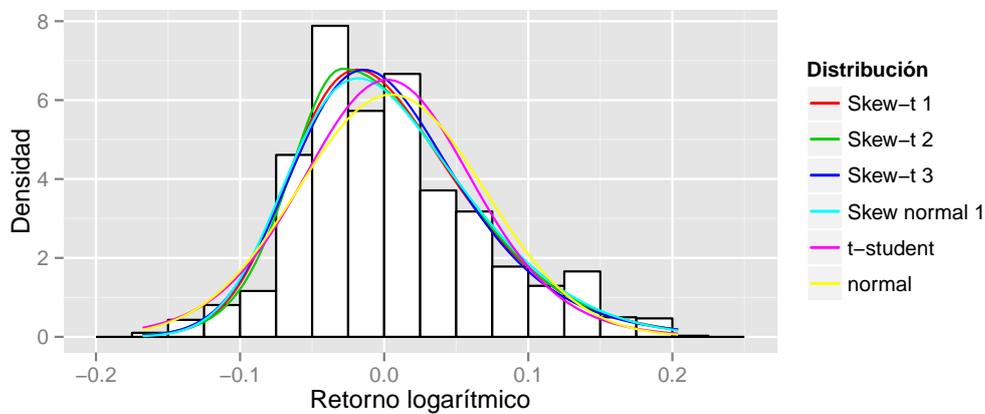


Figura 4.10: Histograma de depreciación logarítmica a 12 meses y fdp estimadas.

Para los cuatro periodos la distribución con mejor ajuste según el AIC es la skew-t 2, seguida por la skew-t 1. La distribución skew-t 3 tiene el tercer mejor ajuste con excepción del depreciación logarítmica correspondiente al periodo de 12 meses, donde toma el cuarto lugar. La distribución normal tiene el peor ajuste en los cuatro periodos.

En general, para cada periodo observamos un valor similar de μ y σ , indistintamente de la distribución. Lo anterior, sugiere que las distribuciones tienen un posicionamiento y dispersión similar.

Por otro lado, α cuenta con valores distintos para cada distribución, sin embargo sugiere que hay asimetría a la derecha en todos los casos. La interpretación de los valores de α corresponde a cada distribución de manera independiente, para más información ver Capítulo 2 y 3.

Además, encontramos que v toma valores diferentes dependiendo de la distribución. No obstante, estos valores no son tan disímiles considerando que $v \in (0, \infty)$.

Finalmente, en general η presenta valores similares entre las distribuciones, lo cual indica que las distribuciones modelan un nivel de asimetría similar para los datos. Los valores de η sugieren que los datos tienen un nivel de asimetría a la derecha intermedia considerando que

$\eta \in [-1, 1]$.

Observamos cambios en los valores de los parámetros a través de los periodos. Tomando como referencia a la skew-t 2, por ser la distribución con el mejor ajuste, observamos que los valores de v, α, η aumentan a medida que el rango del periodo incrementa.

Inicialmente $v = 3.30$ y aumenta hasta llegar a 12.08, sugiriendo que a medida que se espere más tiempo entre la compra y venta de PEN hay menor valores de depreciación logarítmica *outliers*. Esto concuerda con lo que uno esperaría; a corto plazo hay mayor presencia de cambios abruptos en el depreciación logarítmica, pero para un periodo más largo hay menos volatilidad en la tasa de cambio.

Por su parte, el aumento de η (o equivalentemente el de α) a través de los periodos indica que la asimetría aumenta con el tiempo. $\eta = 0.23$ a 3 meses y llega a 0.34. Considerando que μ siempre está cercano a 0, esto significa que si uno espera más tiempo entre la compra y venta de PEN hay mayor concentración de valores negativos de depreciación logarítmica. Lo anterior sugiere que, dentro del periodo estudiado, a mayores plazos entre compra y venta hay mayor tendencia a que el PEN se aprecie. Esto tiene sentido considerando que dentro del periodo el PEN por lo general se apreció.

Considerando los análisis previos llegamos las siguiente conclusiones:

- La distribución que presenta el mejor ajuste a los datos es la skew-t 2 seguida por la skew-t 1.
- Para los datos de depreciación logarítmica, a menor rango de periodo,
 - la asimetría es menor y
 - hay mayor presencia de puntos *outliers*.

Capítulo 5

Resultados

Los resultados principales de este trabajo son los siguientes:

1. Las tres propuestas de distribución skew-t
 - Cumplen con la condición necesaria para la existencia de la medida de [Patil et al. \(2014\)](#).
 - Aparentan existir en el rango completo de asimetría de la medida de [Patil et al. \(2014\)](#) (estimado por métodos numéricos).
 - El parámetro α principalmente impacta la asimetría.
 - Para menores valores de v , la asimetría es mayor en las tres distribuciones. No obstante, el impacto es mayor en la skew-t 3, pequeño en la 1 y muy pequeño en la 2.
2. Respecto a las parametrizaciones de la distribución skew-t 3
 - Las dos colas tienen diferentes pesos independiente de la parametrización.
 - Los parámetros no modelan independientemente la asimetría y el peso de las colas.
 - La parametrización alterna de la distribución skew-t 3 de [Jones y Faddy \(2003\)](#) no cuenta con una relación monótonica respecto a sus parámetros y la medida de asimetría de [Patil et al. \(2014\)](#).
 - La parametrización propuesta de la distribución skew-t 3 cuenta con una relación monótonica con α y la medida de asimetría de [Patil et al. \(2014\)](#).
 - La distribución no logra altos niveles de asimetría cuando sus dos colas son ligeras. Consecuentemente, no hay una skew normal que tenga el rango completo de nivel de asimetría asociada a la distribución.
3. En relación a la aplicación de las propuestas de distribución skew-t a datos reales
 - La distribución que presenta el mejor ajuste a los datos es la skew-t 2 seguida por la skew-t 1.
 - Para los datos de retorno logarítmico, a menor rango de periodo, la asimetría es menor y hay mayor presencia de puntos outliers.

Capítulo 6

Conclusiones

Las conclusiones principales de este trabajo son las siguientes:

1. La distribución skew-t 3 tiene una limitación importante al no permitir altos niveles de asimetría cuando sus dos colas son ligeras. No obstante, al contar con colas de diferentes pesos puede presentar un mejor ajuste a datos reales frente a las otras dos propuestas de distribución skew-t.
2. La parametrización propuesta en este trabajo de la skew-t 3 es superior a la alterna porque separa mejor la asimetría y el peso de las colas y cuenta con una relación monótonica entre su parámetro de asimetría y la medida de asimetría de [Patil et al. \(2014\)](#).
3. La propuesta de skew-t en nuestra aplicación que mejor separa la asimetría y pesos de las colas es la 2, seguida por la 1 y finalmente la 3.
4. Las distribuciones skew-t 2 y 1 presentan resultados similares en la aplicación a datos reales.
5. Las propuestas de skew-t pueden ser utilizadas en describir el nivel de asimetría y *outliers* en un conjunto de datos. Esto puede ser útil para comparar el nivel de asimetría y *outliers* de diferentes bases de datos.

6.1. Sugerencias para investigaciones futuras

Sugerimos que futuras investigaciones

- Estudiar otras parametrizaciones que resulten en una relación más lineal entre el parámetro de asimetría de las distribución y la medida de asimetría de [Patil et al. \(2014\)](#)
- Ampliar el estudio a otras distribuciones que tienen como caso particular la *t*-student.
- Estudiar la curtosis de las distribuciones con una medida que resulte apropiada para las propuestas de skew-t.
- Aplicar las distribuciones skew-t a otros datos reales no simétricos donde se desee estudiar el nivel de asimetría y *outliers* de los datos.

Apéndice A

Estimación de parámetros por máxima verosimilitud

Las propuestas de distribución skew-t estudiadas cuentan con los parámetros μ , σ , α y v . Bajo la perspectiva clásica, estos parámetros se estiman para un conjunto de datos utilizando estimación por máxima verosimilitud (EMV).

En general, la EMV requiere algoritmos que utilizan la segunda derivada de la función de máxima verosimilitud. No obstante, encontramos que las implementaciones computacionales para las propuestas de skew-t utilizan métodos numéricos para aproximar la primera y segunda derivada.

Este capítulo describe brevemente la teoría de la EMV y referencia los métodos utilizados en las implementaciones computacionales.

A.1. Principios teóricos de la EMV

En esta sección describimos los principios teóricos de la distribución asintótica de EMV, tomando como referencia [Fahrmeir et al. \(2013\)](#).

Sea Y_1, \dots, Y_n una muestra aleatoria de una distribución continua con realizaciones y_1, \dots, y_n . Sea la la fdp conjunta

$$f(Y_1 = y_1, \dots, Y_n = y_n | \theta)$$

la cual depende de un vector desconocido $\theta = (\theta_1, \dots, \theta_p) \in \Theta$ el cual se requiere estimar. Dadas las realizaciones y_1, \dots, y_n la fdp conjunta es considerada una función de θ y es denominada verosimilitud:

$$\mathcal{L}(\theta) = f(Y_1 = y_1, \dots, Y_n = y_n | \theta).$$

El principio de máxima verosimilitud postula que el estimador de máxima verosimilitud (EMV) $\hat{\theta}$ para θ es el valor de $\hat{\theta}$ el cual maximiza $\mathcal{L}(\theta)$.

En el caso continuo, este principio dice: determina el EMV $\hat{\theta}$ tal que la fdp asume su máximo valor para $\theta = \hat{\theta}$, haciendo que la muestra tenga mayor probabilidad de ocurrir. En la mayoría de casos, $\mathcal{L}(\theta)$ es diferenciable con respecto a θ y la maxima se puede determinar al fijar las primeras derivadas a cero y resolviendo el sistema de ecuaciones para $\theta = \hat{\theta}$. Por razones técnicas, la maximación

normalmente no se hace a la función de verosimilitud sino al log-verosimilitud $\ell(\boldsymbol{\theta}) = \log(\mathcal{L}(\boldsymbol{\theta})) \dots$

En el caso más simple Y_1, \dots, Y_n son una muestra independientemente e idénticamente distribuida de $f(y|\boldsymbol{\theta}) \dots$ Por ende la verosimilitud es

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^n f(y_i|\boldsymbol{\theta})$$

y la log-verosimilitud es

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \log f(y_i|\boldsymbol{\theta}).$$

Propiedades del Estimador de Máxima Verosimilitud (EMV)

P1: Invarianza. Si $\hat{\boldsymbol{\theta}}$ es el EMV de $\boldsymbol{\theta}$, entonces para cualquier función $h(\cdot)$, el EMV de $h(\boldsymbol{\theta})$ es $h(\hat{\boldsymbol{\theta}})$.

P2: Distribución asintótica del EMV. Si la familia de distribuciones o modelos de probabilidad $f(\mathbf{y}|\boldsymbol{\theta})_{\boldsymbol{\theta} \in \Theta}$ satisface las condiciones de regularidad y \mathbf{y} es una variable aleatoria con función de distribución $f(\mathbf{y}|\boldsymbol{\theta})$, con $\boldsymbol{\theta} \in \Theta$. Entonces, si el tamaño de muestra n es grande, el estimador de máxima verosimilitud para:

$$\hat{\boldsymbol{\theta}} \approx N(\boldsymbol{\theta}, F(\boldsymbol{\theta})^{-1}),$$

donde $F(\boldsymbol{\theta})$ es la matriz de información de Fisher y es dada por

$$F(\boldsymbol{\theta}) = E \left(\left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^2 \right) = -E \left(\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right)$$

Luego se estima $F(\boldsymbol{\theta})$ por $F(\hat{\boldsymbol{\theta}})$. En la práctica, cuando no se conoce $F(\boldsymbol{\theta})$, se utiliza la matriz de información observada $F_{OBS}(\hat{\boldsymbol{\theta}})$.

$$F_{OBS}(\hat{\boldsymbol{\theta}}) = \left(-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

A.2. Implementación computacional

Las implementaciones computacionales utilizadas para estimar los parámetros en el Capítulo 4 son rutinas que obtienen estimaciones de la máxima verosimilitud de a través de optimizaciones computacionales. Estas rutinas en vez de utilizar la matriz de información de Fisher o la matriz observada de información de Fisher, aplican métodos numéricos que permiten estimar la máxima verosimilitud sin tener que hallar estas expresiones.

Las tres implementaciones utilizan la función `nlnmb` de R la cual es una rutina que forma parte de la librería PORT (Fox, 2013). Esta función utiliza gradientes como parte del proceso de optimización. Los gradientes son calculados por métodos numéricos que aproximan la

primera y segunda derivada. Los detalles de cómo es implementada la función se dan en (Fox, 2013).

A.2.1. Funciones y paquetes en R

A continuación se listan las implementaciones computacionales para las propuestas de distribución skew-t.

Distribución	Paquete		
	gamlss	sn	fGarch
skew-t 1	gamlssML	selm	
skew-t 2	gamlssML		sstdFit
skew-t 3	gamlssML		

Cuadro A.1: Funciones y paquetes en R para la implementación computacional por EMV



Apéndice B

Código R para estimar la medida de asimetría de Patil

A continuación se presenta la función de R utilizada para estimar la medida de asimetría de [Patil et al. \(2014\)](#).

```
rm(list=ls())
library(sn)
library(ggplot2)
library(reshape2)
library(skewt)
library(gamlss.dist)
library(truncdist)

setwd("~/Documents/PUCP Maestr\ 'ia estad\ 'istica/Tesis/Tesis")

wd <- getwd()

#=====
# Funciones
#=====
#v_p: cortes de p a los cuales se va aplicar la estimaci\ 'on
# f_d: funci\ 'on de de densidad poblacional relacionada a la distribuci\ 'on
# f_p: funci\ 'on de de distribuci\ 'on acumulada relacionada a la distribuci\ 'on
# f_q: funci\ 'on de cuantiles
# f_r: funcio\ 'on de vector aleatorio relacionada a la distribuci\ 'on
# n_rand: numero de puntos aleatorios
# ...: parametros de las funciones f_d, f_p, f_q

f_eta <- function(v_p, f_d, f_p, f_q, f_r,n_rand ,...){

  if( any( v_p > 1 | v_p< 0.5 )){
    stop( "valor de v_p mayor que 1 o menor que 0.5")
  }
  v_p <- v_p[ !v_p %in% 1]
  v_xi <- f_q(v_p,...)
  v_xi_s= f_q(1-v_p, ...)
  v_rand <- f_r(n_rand, ... )
  v_rand <- v_rand[ !is.na(v_rand)]
}
```

APÉNDICE B. CÓDIGO R ESTIMACIÓN DE ASIMETRÍA

```

v_rand <- v_rand[ !is.infinite(v_rand)]

v_rand_d <- f_d( v_rand,...)
v_rand_p <- f_p( v_rand,...)
df_long <- data.frame(rand= v_rand,rand_d= v_rand_d, rand_p = v_rand_p )

k_corr_1 <- cor( df_long$rand_d , df_long$rand_p)
v_corr_p <- numeric(length(v_p))

for(i in 1:length(v_p)) {

  v_dens <- rep(0, length(v_rand))
  v_acum <- rep(1, length(v_rand))

  v_dens[df_long$rand <= v_xi[i] ] <- df_long$rand_d[ df_long$rand <= v_xi[i] ] /v_p[i]
  v_acum[df_long$rand <= v_xi[i] ] <- df_long$rand_p[ df_long$rand <= v_xi[i] ] /v_p[i]

  v_corr_p[i] <- cor( v_dens,v_acum)
}
v_corr_p <- c( v_corr_p, k_corr_1)

v_corr_p_s <- numeric(length(v_p))

for(i in 1:length(v_p)) {

  v_dens <- rep(0, length(v_rand))
  v_acum <- rep(0, length(v_rand))

  v_dens[df_long$rand >= v_xi_s[i] ] <- df_long$rand_d[df_long$rand >= v_xi_s[i] ] /v_p[i]
  v_acum[df_long$rand >= v_xi_s[i] ] <- 1-(1- df_long$rand_p[ df_long$rand >=
  v_xi_s[i] ] )/v_p[i]

  v_corr_p_s[i] <- cor( v_dens,v_acum)
}
v_corr_p_s <- c( v_corr_p_s, k_corr_1)

df_rho <- data.frame(p=c(v_p,1),corr_p= v_corr_p,corr_p_s=v_corr_p_s ,
  corr_sum= abs(v_corr_p +v_corr_p_s))
eta <- -0.5* sign(k_corr_1) * max( df_rho$corr_sum ,na.rm = T)

ls_fin <- list(df_rho=df_rho, eta =eta , parametros= ...)

return(ls_fin)
}

```

Esta función es utilizada para estimar la medida de asimetría de [Patil et al. \(2014\)](#) para las tres propuestas de distribución skew-t. A continuación se presenta como ejemplo su aplicación

APÉNDICE B. CÓDIGO R ESTIMACIÓN DE ASIMETRÍA

en la distribución skew-t 1.

```

=====
# Ejemplo de Aplicaci\on txT
=====

v_p <- seq(from=0.5,to=1,by=0.01)

ls_eta_txT_v1 <- list()
v_alphas <- -20:20

system.time(
  for(i in 1:length(v_alphas) ){ #length(v_alpha)
    ls_eta_txT_v1[[i]] <- f_eta(v_p= v_p, f_d = dst, f_p= pst, f_q = qst,
                              f_r= rst,n_rand= 200000 ,alpha=v_alphas[i],nu=1 )
  }
)

save(ls_eta_txT_v1, file= "ls_eta_txT_v1.RData" )

v_p <- seq(from=0.5,to=1,by=0.01)

ls_eta_txT_v3 <- list()
v_alphas <- -20:20

system.time(
  for(i in 1:length(v_alphas) ){ #length(v_alpha)
    ls_eta_txT_v3[[i]] <- f_eta(v_p= v_p, f_d = dst, f_p= pst, f_q = qst,
                              f_r= rst,n_rand= 200000 ,alpha=v_alphas[i],nu=3 )
  }
)

save(ls_eta_txT_v3, file= "ls_eta_txT_v3.RData" )

v_p <- seq(from=0.5,to=1,by=0.01)

v_alphas <- -20:20

ls_eta_txT_v10 <- list()
system.time(
  for(i in 38:41 ){ #length(v_alpha)
    ls_eta_txT_v10[[i]] <- f_eta(v_p= v_p, f_d = dst, f_p= pst, f_q = qst,
                              f_r= rst,n_rand= 200000 ,alpha=v_alphas[i],nu=10 )
  }
)

```

APÉNDICE B. CÓDIGO R ESTIMACIÓN DE ASIMETRÍA

```
}  
)  
save(ls_eta_txT_v10, file= "ls_eta_txT_v10.RData" )  
  
v_p <- seq(from=0.5,to=1,by=0.01)  
  
#ls_eta_txT_v10 <- list()  
v_alphas <- -20:20  
  
ls_eta_txT_v50 <- list()  
system.time(  
  for(i in 1:length(v_alphas) ){ #length(v_alpha)  
    ls_eta_txT_v50[[i]] <- f_eta(v_p= v_p, f_d = dst, f_p= pst, f_q = qst,  
                               f_r= rst,n_rand= 200000 ,alpha=v_alphas[i],nu=50 )  
    print(i)  
  }  
)
```



Apéndice C

Código R para estimar parámetros a conjunto de datos

A continuación se presenta el cálculo de retorno logarítmico de la tasa de cambio PEN a USD.

```
rm(list = ls())

library(data.table)
library(lubridate)
setwd("~/Documents/PUCP Maestr\ 'ia estad\ 'istica/Tesis")

dt_sol <- data.table(read.csv(file = "tasaCambioSol.csv" ,stringsAsFactors=F))

setnames(dt_sol,c("date","value"))

dt_sol$date <- gsub("Ene", "Jan",dt_sol$date)
dt_sol$date <- gsub("Abr", "Apr",dt_sol$date)
dt_sol$date <- gsub("Ago", "Aug",dt_sol$date)
dt_sol$date <- gsub("Set", "Sep",dt_sol$date)
dt_sol$date <- gsub("Dic", "Dec",dt_sol$date)

dt_sol$date <- as.Date (dt_sol$date , "%d%b%y")

dt_sol <- dt_sol[ value!="n.d."]
dt_sol[ ,value:=as.numeric(value)]

dt_sol_time_series <- dt_sol

prepDatos <- function(dt_1, v_mes){
  dt_1 <- rbind(dt_1, data.frame( date=as.Date(NA), value=as.numeric(NA)))
  dt_1$value_mas_un_dia <- c(NA, dt_1$value[1:(length(dt_1$value)-1)])
  dt_1$change_un_dia <- dt_1$value_mas_un_dia/dt_1$value-1
  dt_1$log_change_un_dia <- log(dt_1$value_mas_un_dia/dt_1$value)

  dt_1$date_mas_un_anho <- dt_1$date+years(1)
  dt_1$date_mas_un_anho[which(dt_1$date_mas_un_anho>
                             max(dt_1$date,na.rm = T) )] <- NA
  k_cond <- any(! dt_1$date_mas_un_anho %in% dt_1$date)
  while(k_cond){
```

```

dt_1$date_mas_un_anho[
  which(!(dt_1$date_mas_un_anho %in%
    dt_1$date))] <- dt_1$date_mas_un_anho[
    which(!(dt_1$date_mas_un_anho %in% dt_1$date))] +1

k_cond <- any(! dt_1$date_mas_un_anho %in% dt_1$date)
}

dt_2 <- dt_1[ , .(date,value)]
setnames(dt_2,c( "date_mas_un_anho", "value_mas_un_anho"))

setkey(dt_2,date_mas_un_anho )

setkey(dt_1,date_mas_un_anho )

dt_2 <- unique(dt_2)
setkey(dt_2,date_mas_un_anho )
dt_1 <- merge(dt_1,dt_2,all.x=T)

dt_1$change_un_anho <- dt_1$value_mas_un_anho/dt_1$value-1
dt_1$log_change_un_anho <- log(dt_1$value_mas_un_anho/dt_1$value)

#=====meses
for(k_mes in v_mes){
  date_mas_x_mes <- paste0("date_mas_",k_mes, "_mes")
  value_mas_x_mes <- paste0("value_mas_",k_mes, "_mes")
  change_x_mes <- paste0("change_",k_mes, "_mes")
  log_change_x_mes <- paste0("log_change_",k_mes, "_mes")
  dt_1[ ,eval(expr = parse(text = date_mas_x_mes)) := date+months(k_mes)]
  dt_1[ eval(expr = parse(text = date_mas_x_mes)) >
    max(date,na.rm = T) ,
    eval(expr = parse(text = date_mas_x_mes)) := NA ]
  k_cond <- any(! dt_1[ ,
    eval(expr = parse(text = date_mas_x_mes)) ] %in% dt_1$date)
  while(k_cond){
    index <- which(! dt_1[ ,
    eval(expr = parse(text = date_mas_x_mes)) ] %in% dt_1$date)
    dt_1[ index ,
    eval(expr = parse(text = date_mas_x_mes)) :=
    eval(expr = parse(text = date_mas_x_mes)) +1 ]

    k_cond <- any(! dt_1[ ,
    eval(expr = parse(text = date_mas_x_mes)) ] %in% dt_1$date)
  }

  dt_3 <- dt_1[ , .(date,value)]
  setnames(dt_3,c( date_mas_x_mes,value_mas_x_mes))

  setkeyv(dt_3,date_mas_x_mes )

```

APÉNDICE C. CÓDIGO R ESTIMACIÓN DE PARÁMETROS

```

dt_3 <- unique(dt_3)
setkeyv(dt_3,date_mas_x_mes )

setkeyv(dt_1, date_mas_x_mes )

dt_1 <- merge(dt_1,dt_3,all.x=T)
dt_1[,
  eval(expr = parse(text = change_x_mes)) :=
  eval(expr = parse(text = value_mas_x_mes)) /value-1]
dt_1[, eval(expr = parse(text = log_change_x_mes)):=
  log(eval(expr = parse(text = value_mas_x_mes)) /value)]

}

setkey(dt_1,NULL)
dt_1
}

dt_sol <- prepDatos(dt_sol,v_mes=c(1:12,24,36,48))
detach("package:lubridate", unload=TRUE)

```

La estimación de de los parámetros por por máxima verosimilitud y métodos numéricos se da con las siguientes funciones.

```

library(data.table)

library(sn)

library(stringr)
library(gamlss)
library(ggplot2)
library(xtable)
library(reshape2)
library(MASS)
library(fGarch)
library(grid)
library(Hmisc)
f_matcoef <- function(object,vpart=F){
  coef <- object$Allpar
  se.coef <- vcov(object, "se")
  tval <- coef/se.coef

  v_links <- c(object$mu.link, object$sigma.link, object$nu.link, object$tau.link)
  v_links <- ifelse(v_links=="identity","",paste("", v_links, " "))
  if( vpart){
    v_var <- paste0(v_links , c("$\\mu$", "$\\sigma$", "$q$", "$p$"))
  } else {
    v_var <- paste0(v_links , c("$\\mu$", "$\\sigma$", "$\\alpha$", "$v$"))
  }
}

```

APÉNDICE C. CÓDIGO R ESTIMACIÓN DE PARÁMETROS

```

v_sig <- 2*(1-pnorm(abs(tval)))
v_sig <- as.character(iffelse(v_sig==0,"< 2.22e-16",v_sig))
matcoef <- data.frame(v_var,coef, se.coef, tval, v_sig)
row.names( matcoef) <- NULL
names(matcoef) <- c("Par\\'ametro","Estimado", "Error std.", "t valor", "Pr(>|t|)")
return(matcoef)
}

f_matcoef1 <- function(object){
coef <- object@param$dp
se.coef <- sqrt(diag(object@param.var$dp))
tval <- coef/se.coef
v_var <- c("$\\mu$", "$\\sigma$","$\\alpha$","$v$")
v_sig <- 2*(1-pnorm(abs(tval)))
v_sig <- as.character(iffelse(v_sig==0,"< 2.22e-16",v_sig))
matcoef <- data.frame(v_var,coef, se.coef, tval, v_sig)
row.names( matcoef) <- NULL
names(matcoef) <- c("Parametro","Estimado", "Error std.", "t valor", "Pr(>|t|)")
return(matcoef)
}

f_aplicacion <- function(v_datos){

v_datos <- v_datos[ !is.na(v_datos) & !is.infinite(v_datos) ]
k_n <- length(v_datos)

ls_df_res <- list()

#=====
fit_t <- fitdistr( v_datos, "t")
ls_df_res[[1]] <- data.frame(aic= 6-2*fit_t$loglik,
                           mu= fit_t$estimate[1],
                           sigma= fit_t$estimate[2],
                           alpha= NA, v= fit_t$estimate[3],dist="t-student")
#=====

#=====
fit_norm <- fitdistr( v_datos, "normal")
ls_df_res[[2]] <- data.frame(aic= 6-2*fit_norm$loglik,
                           mu= fit_norm$estimate[1],
                           sigma= fit_norm$estimate[2],
                           alpha= NA, v= NA, dist="Normal")
#=====

#=====
fit_txT <- selm(formula=v_datos~1,family="ST" )#,fixed.param=list(nu=3))
#fit_txT <- gamlssML(y=v_datos,family="ST2",tau.fix=1.3)
str(fit_txT)
fit_txT@logL
#fit_asa@ opt.method$par
fit_txT@param$dp

```

APÉNDICE C. CÓDIGO R ESTIMACIÓN DE PARÁMETROS

```

ls_df_res[[3]]<- data.frame(aic= 8-2*fit_txT@logL,
                           mu= fit_txT@param$dp[1],
                           sigma= fit_txT@param$dp[2],
                           alpha= fit_txT@param$dp[3], v= fit_txT@param$dp[4],
                           dist="ST t$\times$T")

#=====

#fit_ttrunc <- sstdFit(v_datos)
fit_ttrunc <- gamlssML(y=v_datos,family="ST3")
sum_fit_ttrunc <- summary(fit_ttrunc)

sum_fit_ttrunc[c("aic", "mu","sigma","nu" ,"tau" )]
ls_df_res[[4]]<- data.frame(aic= sum_fit_ttrunc$aic,
                           mu= sum_fit_ttrunc$mu,
                           sigma=sum_fit_ttrunc$sigma,
                           alpha= sum_fit_ttrunc$nu, v= sum_fit_ttrunc$tau,
                           dist="ST t-truncada")

#=====

#fit_vpart <- gamlssML(y=v_datos,family="ST5")
sum_fit_vpart <- summary(fit_vpart)

sum_fit_vpart[c("aic", "mu","sigma","nu" ,"tau" )]

df_vpart <- data.frame(aic= sum_fit_vpart$aic,
                      mu= sum_fit_vpart$mu,
                      sigma=sum_fit_vpart$sigma,
                      q= sum_fit_vpart$nu, p= sum_fit_vpart$tau,
                      dist="ST v-partida")

df_vpart$a <- 1/df_vpart$p + df_vpart$q/(df_vpart$p*sqrt(2*df_vpart$p+df_vpart$q^2))
df_vpart$b <- 1/df_vpart$p - df_vpart$q/(df_vpart$p*sqrt(2*df_vpart$p+df_vpart$q^2))
df_vpart$alpha <- (df_vpart$a-df_vpart$b)/(df_vpart$a+df_vpart$b)
if ( df_vpart$alpha>0){
  df_vpart$v <- 2*df_vpart$b
} else {
  df_vpart$v <- 2*df_vpart$a
}

ls_df_res[[5]]<- data.frame(aic= df_vpart$aic,
                          mu= df_vpart$mu,

```

```

        sigma=df_vpart$sigma,
        alpha= df_vpart$alpha, v= df_vpart$v,
        dist="ST v-partida")

#=====

#=====
fit_phixPhi <- selm(formula=v_datos~1,family="SN")

ls_df_res[[6]]<- data.frame(aic= 6-2*fit_phixPhi@logL,
                          mu= fit_phixPhi@param$dp[1],
                          sigma= fit_phixPhi@param$dp[2],
                          alpha= fit_phixPhi@param$dp[3], v= NA,
                          dist="SN  $\phi$  times  $\Phi$ ")

#=====

df_res <- do.call(rbind, ls_df_res)
df_res <- df_res[ order( df_res$aic),]
row.names(df_res) <- NULL
names(df_res) <- c("AIC",
                  " $\mu$ ", " $\sigma$ ",
                  " $\alpha$ ", " $v$ ", "Distribuci'on")
df_res <- df_res[, c("Distribuci'on", "AIC",
                    " $\mu$ ", " $\sigma$ ", " $\alpha$ ", " $v$ ")]

df_fit_txT <- f_matcoef1( fit_txT)
df_fit_ttrunc <- f_matcoef( fit_ttrunc)
df_fit_vpart <- f_matcoef( fit_vpart,vpart=T)
ls_final <- list(df_res=df_res, df_vpart=df_vpart,
                df_fit_txT=df_fit_txT,df_fit_ttrunc=df_fit_ttrunc,
                df_fit_vpart=df_fit_vpart,k_n=k_n,v_datos=v_datos )

ls_final
}

```

Bibliografía

- Aas, K. y Haff, I. (2006). The generalized hyperbolic skew student's t -distribution, *Journal of Financial Econometrics* **4(2)**: 275–309.
- Arellano-Valle, R. B., Gómez, H. W. y Quintina, F. A. (2005b). Statistical inference for a general class of asymmetric distributions, *J. Statist. Plann. Inference* **128**: 427–443.
- Azzalini, A. (1985). A class of distributions wich includes the normal ones, *Scandinavian Journal of Statistics* **12**: 171–178.
- Azzalini, A. (2014). *The R sn package : The skew-normal and skew- t distributions (version 1.1-0)*, Università di Padova, Italia.
URL: <http://azzalini.stat.unipd.it/SN>
- Azzalini, A. y Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution, *Journal of the Royal Statistical Society Ser. B* **65**: 367–389.
- BCRP (2015). Tipo de cambio nominal.
URL: <https://estadisticas.bcrp.gob.pe/estadisticas/series/diarias/tipo-de-cambio-nominal>
- Branco, M. D. y Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis* **79**: 99–114.
- Fahrmeir, L., Kneib, T. y Lang, S. (2013). *Regression: Models, Methods, and Applications.*, Springer.
- Fernández, C. y Steel, M. F. J. (1998). On bayesian modelling of fat tails and skewness, *Journal of The American Statistical Association* **93**: 359–371.
- Fox, P. (2013). *The Port Matemathical Subroutine Library Version 3*, AT and T Bell Laboratories Murry Hill, NJ.
- Jones, M. y Faddy, M. (2003). A skew extension of the t distribution, with applications, *Journal of the Royal Statistical Society Series B*, **65**: 159–174.
- King, R., y with contributions from Emily Anderson. (2012). *skewt: The Skewed Student- t Distribution*. R package version 0.1.
URL: <http://CRAN.R-project.org/package=skewt>
- Li, X. y Morris, J. M. (1991). On measuring asymmetry and the reliability of the skewness measure, *Stat. Probab. Lett.* **12**: 267–271.
- Marchenko, Y. V. y Genton, M. G. (2010). Multivariate log-skew-elliptical distributions with applications to precipitation data, *Enviromentrics* **21**: 318–340.
- McNeil, A. J., Rudiger, F. y Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton University Press.

- Patil, P. N., Bagkovos, D. y Wood, A. (2014). A measure of asymmetry based on a new necessary and sufficient condition for symmetry, *The Indian Journal of Statistics* **76-A**: 123–145.
- Patil, P. N., Patil, P. P. y Bagkovos, D. (2012). A measure of symmetry, *Statistics Papers* **52**: 971–985.
- Prentice, R. L. (1975). Discrimination among some parametric models, *Biometrika* **62**: 607–614.
- Rigby, R. A. y Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape,(with discussion), *Applied Statistics* **54**: 507–554.
- Sahu, S. K., Dey, D. K. y Branco, M. D. (2003). A new class of multivariate skew distributions with applications to bayesian regression models, *The Canadian Journal of Statistics* **31(2)**: 129–150.
- Stasinopoulos, M., with contributions from Calliope Akantziliotou, B. R., Heller, G., Ospina, R., Motpan, N., McElduff, F., Voudouris, V., Djennad, M., Enea, M. y Ghalanos., A. (2014). *gamlss.dist: Distributions to be used for GAMLSS modelling*. R package version 4.3-1.
URL: <http://CRAN.R-project.org/package=gamlss.dist>
- Walls, W. D. (2005). Modeling heavy tails and skewness in film returns, *Applied Financial Economics* **15**: 1181–1188.
- Zhau, D. y Galbraith, J. W. (2010). A generalized asymmetric student-t distribution with application to financial econometrics, *Journal of Econometrics* **157(2)**: 297–305.