



CIMAT

Centro de Investigación en Matemáticas, A.C.

---

MEDICIÓN DE RIESGO EN PORTAFOLIOS  
APLICANDO UN MODELO  
GARCH-EVT-CÓPULA

T E S I S

Que para obtener el grado de

**Maestro en Ciencias**

con especialidad en

**Probabilidad y Estadística**

P R E S E N T A:

**Verónica Alarcón Márquez**

Director de Tesis:

Dr. Daniel Hernández Hernández

---

Autorización de la versión final

Guanajuato, Gto. 30, Agosto de 2019.

---

Acta No.: 152

Libro No.: 002

Foja No.: 152

En la Ciudad de Guanajuato, Gto., siendo las 15:00 horas del día 30 de agosto del año 2019, se reunieron los miembros del jurado integrado por los señores:

**DRA. GRACIELA MARÍA DE LOS DOLORES GONZÁLEZ FARÍAS** (CIMAT)

**DR. EHYTER MATÍAS MARTÍN GONZÁLEZ** (UG)

**DR. DANIEL HERNÁNDEZ HERNÁNDEZ** (CIMAT)

bajo la presidencia del primero y con carácter de secretario el segundo, para proceder a efectuar el examen que para obtener el grado de

**MAESTRO EN CIENCIAS  
CON ESPECIALIDAD EN PROBABILIDAD Y ESTADÍSTICA**

Sustenta

**VERÓNICA ALARCÓN MÁRQUEZ**

en cumplimiento con lo establecido en los reglamentos y lineamientos de estudios de posgrado del Centro de Investigación en Matemáticas, A.C., mediante la presentación de la tesis

**"MEDICIÓN DE RIESGO EN PORTAFOLIOS APLICANDO  
UN MODELO GARCH-EVT-CÓPULA "**

Los miembros del jurado examinaron alternadamente al (la) sustentante y después de deliberar entre sí resolvieron declararlo (a):

*Aprubada*

*Graciela Farías*

**DRA. GRACIELA MARÍA DE LOS DOLORES GONZÁLEZ FARÍAS**  
Presidente

*Ehyter Matías*

**DR. EHYTER MATÍAS MARTÍN GONZÁLEZ**  
Secretario

*Daniel Hernández*

**DR. DANIEL HERNÁNDEZ HERNÁNDEZ**  
Vocal



**CIMAT**  
DIRECCIÓN  
GENERAL

**DR. VÍCTOR MANUEL RIVERO MERCADO**  
Director General

---

*“Essentially, all models are wrong, but some are useful.”*

---

George Box

---

# Índice general

<b>Introducción</b>	<b>I</b>
<b>1. Preliminares</b>	<b>9</b>
1.1. Series de Tiempo . . . . .	9
1.1.1. Modelo ARMA . . . . .	10
1.1.2. Modelo GARCH . . . . .	11
1.1.3. Modelo GJR-GARCH . . . . .	12
1.2. Teoría de Valores Extremos (EVT) . . . . .	12
1.3. Cópulas . . . . .	14
1.3.1. Algunos tipos de funciones cópula . . . . .	15
1.3.2. Estimación de Parámetros . . . . .	17
1.3.3. Dependencia entre variables aleatorias . . . . .	19
1.4. Pruebas de bondad de Ajuste . . . . .	20
1.4.1. Backtesting . . . . .	22
<b>2. Modelo estadístico para el portafolio</b>	<b>25</b>
2.1. Medidas de Riesgo VaR y CVaR . . . . .	26
2.2. Modelo de series de tiempo para los retornos . . . . .	29
2.3. Distribuciones marginales de las divisas . . . . .	30
2.3.1. Ajuste de Valores Extremos por umbral . . . . .	31
2.3.2. Selección de Umbral para ajuste GPD . . . . .	33
2.4. Construcción de la Distribución Conjunta . . . . .	34
2.4.1. Dependencia a través de cópulas por pares . . . . .	35
2.5. Simulación Montecarlo para calcular el VaR . . . . .	38
<b>3. Implementación del modelo</b>	<b>39</b>
3.1. Estimación de parámetros del modelo de series de tiempo . . . . .	41
3.1.1. Residuos del modelo GJR-GARCH . . . . .	46
3.2. Ajuste de distribuciones Marginales . . . . .	47
3.2.1. Ajuste Distribución $t$ -Student . . . . .	48
3.2.2. Ajuste Distribución Marginal aplicando EVT . . . . .	49
3.3. Estimación de funciones cópula . . . . .	59

3.3.1. Aplicando ajuste $t$ -Student de las marginales y una función cópula 4-variada . . . . .	59
3.3.2. Aplicando valores extremos y PCC . . . . .	62
3.3.3. Vine Copulas . . . . .	62
3.4. Cálculo del VaR del portafolio . . . . .	69
3.4.1. Aplicando marginales $t$ -Student y cópula 4-variada . . . . .	69
3.4.2. Aplicando la estimación de cópulas por pares . . . . .	70
3.4.3. Evaluación de la estimación del VaR . . . . .	73
<b>4. Conclusiones y trabajo futuro</b>	<b>77</b>
<b>5. Apéndice: Código en R</b>	<b>79</b>



# Agradecimientos

Agradezco a todos quienes hicieron posible la realización de este trabajo, a mis padres, hermanas y hermanos por ser mi inspiración y el motor en mi vida personal pues me han impulsado a cumplir mis metas.

A mi asesor el Dr. Daniel Hernández por la confianza en mi trabajo, por su guía y consejos en todo el desarrollo y la exigencia en la elaboración de esta tesis. Gracias a mis sinodales, Dra. Graciela Gómez Farías y Dr. Ehyter Martín, por sus valiosos comentarios y observaciones que permitieron enriquecer este trabajo en cada una de las diferentes áreas.

A mis amigos de la maestría, por ser verdaderos compañeros y hacer una aventura de cada una de las tareas y retos durante estos dos años.

Gracias al Consejo Nacional de Ciencia y Tecnología (CONACYT), por el apoyo recibido durante mis estudios de maestría, por la beca número 488245 como parte del programa de becarios de posgrados de excelencia, lo que hizo posible la realización de esta tesis.

Quiero agradecer al Centro de Investigación en Matemáticas A. C. por la formación académica que recibí durante el programa de maestría y parte de mi licenciatura, los cursos, seminarios y eventos diversos que promueven una formación integral de nosotros como estudiantes.



# Resumen

*En este trabajo se aborda el problema de calcular medidas de riesgo de un portafolio de divisas, tales como el Valor en Riesgo (VaR) y el Valor en Riesgo Condicional (CVaR), por lo que se requiere estimar la distribución conjunta de los rendimientos del portafolio.*

*Para obtener dicha distribución, se aplica un modelo GARCH–EVT–Cópula en el cual, primero se ajusta un modelo de series de tiempo para describir la autocorrelación y heteroscedasticidad en las series de rendimientos para cada activo que conforma el portafolio; luego las distribuciones marginales de los residuos estandarizados se modelan aplicando teoría de valores extremos por umbral. Por último, la dependencia entre las divisas se describe por medio de funciones cópula bivariadas. Este desarrollo ha sido motivado principalmente por el trabajo de Omari et.al. (2018).*

*Finalmente, se presentan los resultados de la implementación del modelo con datos históricos de precios diarios de las divisas euro (EUR), dólar estadounidense (USD), libra esterlina (GBP) y yen japonés (JPY), las cuales son las más importantes en el mercado de divisas y se usa como numerario o moneda de referencia el peso mexicano (MXN).*



# Introducción

En este trabajo se aborda el problema de calcular medidas de riesgo de un portafolio de divisas aplicando un modelo que combina un modelo de Heterocedasticidad Condicional Autoregresiva Generalizada (GARCH), Teoría de Valores Extremos (EVT) y Funciones Cópula, nos vamos a referir en adelante al modelo completo como GARCH–EVT–Cópula, dicho modelo consiste en ajustar una serie de tiempo para describir la autocorrelación y heteroscedasticidad en las series de rendimientos para cada activo que conforma el portafolio; luego las distribuciones marginales de los residuos estandarizados se modelan aplicando teoría de valores extremos, para lo cual se eligen umbrales razonables en cada caso; finalmente la dependencia entre las divisas se describe por medio de funciones cópula. Con ello, se obtiene la distribución conjunta de los rendimientos del portafolio y se calcula el Valor en Riesgo (VaR) y el Valor en Riesgo Condicional (CVaR) como medidas de riesgo. Este desarrollo es motivado principalmente por el trabajo de *Omari et.al.* (2018).

La medición del riesgo en el estudio de portafolios juega un rol muy importante pues permite al inversionista dimensionar las posibles pérdidas para cierta estrategia de inversión; en la literatura se encuentran distintos planteamientos para resolver este problema, algunos de ellos tratan sobre el cálculo de la volatilidad estocástica, el cálculo del VaR y el CVaR. Aquí nos concentramos en estas dos últimas medidas, más adelante se discuten algunas de las ventajas y desventajas de las mismas.

El VaR se puede interpretar como la pérdida máxima esperada en un horizonte de tiempo establecido dado un nivel de confianza. Sin embargo el VaR no siempre se puede calcular cuando las distribuciones no son continuas, además que no satisface la propiedad de subaditividad, esto es, que no está acotado por la suma del VaR individual de los activos que conforman el portafolio. Por estas razones se hace un planteamiento para el cálculo del CVaR como una medida de riesgo alternativa.

En la literatura se pueden encontrar diferentes metodologías para el cálculo del VaR, algunos ejemplos son aplicando un modelo Delta-Normal, donde se supone que los rendimientos del portafolio siguen una distribución normal. Otra de ellas es el método de simulación histórica en la que se eligen representantes de datos históricos y se calcula el VaR de manera empírica considerando estas observaciones. La metodología

que aplicaremos en este trabajo es el método de simulación Montecarlo, para ello se requiere estimar la distribución conjunta de los rendimientos del portafolio y finalmente realizar una cantidad grande de simulaciones de dicha distribución y calcular a partir de estas el VaR de manera empírica.

Por lo tanto, para obtener las medidas de riesgo de interés en este trabajo se requiere un modelo que capture la distribución conjunta de los rendimientos, así cuando la dimensión del portafolio es grande, esta estimación representa un problema de alta complejidad.

Además, las series financieras presentan características importantes que no todos los modelos son capaces de capturar, tales como: no normalidad, agrupamiento de la volatilidad, colas pesadas y volatilidad condicional asimétrica.

En trabajos recientes, se ha motivado la aplicación de combinaciones de modelos para lograr capturar más propiedades de las series financieras. Esto se puede apreciar, por ejemplo, en el trabajo de *Rockinger & Jondeau* (2006) donde se introduce la combinación Cópula-GARCH para modelar la dependencia entre mercados de activos. En *Tang et.al.* (2015) se aplica el modelo GARCH-EVT-Cópula para estimar el riesgo a través del cálculo del VaR y CVaR para portafolios de gas natural. En *Omari et. al.* (2018) se aplica este mismo modelo para un portafolio de cuatro divisas y se calcula el VaR.

La primer etapa del modelo GARCH-EVT-Cópula, consiste en ajustar un modelo  $ARMA(m, n) - GJR - GARCH(p, q)$ ; donde el modelo  $ARMA$  (Autorregresivo de media móvil) filtra la correlación de la serie y luego el modelo  $GJR - GARCH$  hace diferencia entre los errores positivos y negativos, a través de una función indicadora que permite dar mayor peso a errores negativos, esto refleja la asimetría de la volatilidad que se observa en el mercado. Este modelo es una modificación de un modelo GARCH y ha sido introducido por Glosten, Jagannathan y Runkle (1993), de ahí su nombre.

Luego de ajustar el modelo de series de tiempo a los rendimientos y estimar sus parámetros, se obtienen los residuos estandarizados correspondientes a cada divisa, éstos se modelan con una distribución de valores extremos de la familia de distribuciones Pareto Generalizada. En *McNeil* (2000) se muestra que los residuos estandarizados presentan menos correlación que los rendimientos originales, por lo que se sugiere hacer el proceso de inferencia sobre las series de los residuos estandarizados y relacionar su distribución con la de los rendimientos.

En particular, para cada serie de residuos estandarizados se estima su distribución marginal de tal manera que las colas de la distribución son ajustadas con un elemento de la familia Pareto Generalizada y para el interior se realiza una estimación de la

densidad por kernel Gaussiano. Para este ajuste es necesario elegir umbrales  $u^L$  y  $u^R$  que determinan las colas inferior y superior de la distribución marginal, respectivamente. La elección de estos umbrales se hace con base en la función media de excesos y la estabilidad del parámetro de forma para la distribución Pareto Generalizada en el umbral adecuado.

La estructura de dependencia entre las divisas es modelada por medio de funciones cópula, para esta estimación consideramos dos casos, el primero tomando en cuenta una cópula multivariada de la dimensión del portafolio y en un segundo caso por medio de cópulas bivariadas para cada par de divisas. La estimación de los parámetros de las cópulas se hace por máxima verosimilitud, y luego se selecciona el mejor tipo de función cópula mediante criterios AIC y BIC.

En conjunto, la elección de distribuciones marginales y un modelo de cópula adecuados, capturan la distribución conjunta de los rendimientos del portafolio; esto hace posible una buena estimación del VaR y el CVaR como medidas de riesgo del portafolio.

Este trabajo está elaborado de la siguiente forma, en el primer capítulo se presentan los preliminares que versan sobre modelos de series de tiempo, en particular ARMA y GARCH; cópulas; distribuciones de valores extremos, y en particular la distribución Pareto Generalizada. En el segundo capítulo, se presenta a detalle el modelo propuesto por *Omari et.al.*(2018) y se plantea el problema para nuestro conjunto de datos. En éste capítulo también se presenta el método *pair-copula-construction*, el cual permite construir cualquier distribución conjunta por medio de cópulas bivariadas y sus distribuciones marginales.

Por último, en el capítulo 3 se presentan los resultados de la implementación de este modelo para obtener las medidas de riesgo de interés para un portafolio de las divisas: euro (EUR), dólar estadounidense (USD), libra esterlina (GBP) y yen japonés (JPY). Los datos de precios históricos que se van a considerar están dentro del periodo de tiempo del 03 de enero de 2000 al 17 de enero de 2019 y han sido descargados del portal financiero *investing.com*.

Además, se presentan los resultados de las pruebas de validación del modelo, conocidas como *backtesting*, tales como la prueba de Kupiec y la de Christoffersen las cuales prueban la hipótesis que la proporción de veces en las que la estimación del VaR fue menor a la observada es igual al nivel elegido para el cálculo del VaR.





# Capítulo 1

## Preliminares

En este capítulo damos una breve introducción a los temas que serán de importancia en el desarrollo de este trabajo. En la sección (1.1) presentamos modelos de series de tiempo, tales como modelos ARMA y modelos GARCH, estos modelos serán aplicados para modelar los rendimientos de un portafolio de divisas en el siguiente capítulo. También se presenta en la sección (1.2) los fundamentos de la teoría de valores extremos por umbral, lo cual se aplicará para modelar las colas de las distribuciones marginales de activos financieros en un portafolio.

En la sección (1.3) hacemos una introducción al tema cópulas, las cuales se aplican para describir la dependencia entre variables aleatorias y en este caso de los instrumentos financieros que conforman un portafolio. En la última parte de este capítulo, mencionamos algunas pruebas de hipótesis que se aplicarán como referente del ajuste y selección del modelo; así como las pruebas de backtesting para la evaluación del mismo.

### 1.1. Series de Tiempo

Vamos a considerar una serie de tiempo como un conjunto de variables aleatorias indexadas, en este caso el índice corresponderá con el tiempo  $t$  dentro de un periodo establecido; así podemos pensar en una serie de tiempo como el resultado de observar una variable aleatoria  $X_t$  en un periodo de tiempo establecido. Por la naturaleza de los datos, los modelos de series de tiempo son frecuentemente aplicados para modelar datos financieros. En este trabajo vamos a modelar los rendimientos de divisas con un modelo de series de tiempo que combina un modelo ARMA y un modelo GARCH, los cuales vamos a discutir más adelante.

**Definición 1.1.1.** Sea  $\{X_t\}$  una serie de tiempo con  $\mathbb{E}[X_t^2] < \infty$ . Se define la *función de media* de  $\{X_t\}$  como

$$\mu_X(t) = \mathbb{E}[X_t].$$

Y la *función de covarianza* de  $\{X_t\}$  como

$$\gamma_X(r, s) = Cov(X_r, X_s) = \mathbb{E}[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

para cualesquiera enteros  $r$  y  $s$ .

Estas funciones permiten definir la propiedad de estacionariedad de una serie de tiempo, la cual es una de las propiedades deseables para su estudio.

**Definición 1.1.2.** Una serie de tiempo  $\{X_t\}$  se dice que es *estacionaria* si satisface las siguientes dos condiciones:

- i.  $\mu_X(t)$  es independiente de  $t$ .
- ii.  $\gamma_X(t+h, t)$  es independiente de  $t$  para cada  $h$ .

Algunos autores se refieren a esta propiedad como débilmente estacionaria, para hacer diferencia con la propiedad fuerte, aquí nos vamos a referir a este término únicamente como serie de tiempo estacionaria.

Un caso particular de una serie de tiempo muy utilizada es el caso de un proceso estocástico de ruido blanco, la característica de esta serie es que cualesquiera dos observaciones a través del tiempo no están correlacionadas.

**Definición 1.1.3.** Una serie de tiempo  $(X_t, t \in \mathbb{Z})$  se dice que es un ruido blanco si cumple que:

1.  $\mathbb{E}(X_t) = 0$ ,
2.  $Var(X_t) = \sigma^2$ ,
3.  $\forall k \neq 0, Cov(X_t, X_{t+k}) = 0$

Además, se dice que  $X_t$  es un ruido blanco Gaussiano si además cumple que

$$X_t \sim N(0, \sigma^2).$$

### 1.1.1. Modelo ARMA

Los modelos autoregresivos de promedio móvil (ARMA) son modelos de series de tiempo en los que cada observación actual depende de manera lineal, con ciertos coeficientes, de las observaciones anteriores y también de los promedios de las variables anteriores.

**Definición 1.1.4.** Un proceso estocástico  $r_t$  sigue un modelo  $ARMA(n, m)$  si es un proceso estacionario y se puede escribir de la siguiente forma,

$$r_t = \phi_0 + \sum_{i=1}^m \phi_i r_{t-i} + \sum_{i=1}^n \theta_i a_{t-i} + a_t,$$

donde  $a_t$  es un proceso de ruido blanco y los parámetros  $m, n$  son enteros no negativos que se conocen como los órdenes autoregresivo y de promedio móvil, respectivamente.

El primer sumando corresponde a la parte de un modelo autoregresivo, este modelo parte de la idea de que  $m$  observaciones anteriores pueden explicar la observación actual.

### 1.1.2. Modelo GARCH

Una característica de las series de tiempo de datos financieros es que no satisfacen la condición de homoscedasticidad, es decir, la volatilidad no es constante a lo largo del tiempo. Engle (1982) introduce una nueva clase de procesos estocásticos llamados modelos ARCH, en estos modelos la varianza condicional no es constante y depende del cuadrado de las perturbaciones pasadas.

**Definición 1.1.5.** Un proceso  $a_t$  sigue un modelo  $ARCH(q)$ , sí y sólo si,

$$\begin{aligned} a_t &= \sigma_t \eta_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i a_{t-i}^2 \end{aligned}$$

donde  $\alpha_0 > 0$  y  $\alpha_i \geq 0$ ,  $i = 1, \dots, q$ , para garantizar que la varianza sea positiva. Además se requiere que  $\sum_{i=1}^q \alpha_i < 1$  para que el proceso  $\sigma_t^2$  sea estacionario.

Luego en 1986, Bollerslev presentó los modelos  $GARCH$  como una generalización del modelo  $ARCH$ , en los cuales la varianza condicional no sólo depende de los cuadrados de las perturbaciones anteriores, como en los trabajos de Engle, sino que además depende de las varianzas condicionales de periodos anteriores.

**Definición 1.1.6.** Sea  $r_t$  una serie de tiempo con función de media  $\mu_t$  y definimos  $a_t = r_t - \mu_t$ . Entonces el proceso  $a_t$  sigue un modelo  $GARCH(p, q)$  si

$$a_t = \sigma_t \eta_t, \quad \eta_t \sim i.i.d.N(0, 1)$$

donde,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{j=1}^q \alpha_j a_{t-j}^2$$

donde,  $\alpha_0 > 0$  y  $\alpha_j \geq 0$ ,  $\beta_i \geq 0$ ,  $i = 1, \dots, p$  y  $j = 1, \dots, q$ . Además se pide que  $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ , para que la varianza sea positiva y existan los momentos de orden superior.

Para observar el efecto en la varianza de este modelo, notemos que en el caso de un modelo  $GARCH(1, 1)$ , donde la volatilidad está dada por,

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 a_{t-1}^2$$

es claro que valores altos de  $a_{t-1}^2$  o de  $\sigma_{t-1}^2$  provocan un valor grande de  $\sigma_t^2$ . Este efecto se conoce como agrupamiento de la volatilidad, pues un shock grande  $a_{t-1}^2$  tiende a ser seguido por otro shock grande  $a_t$ , esta propiedad se observa en algunos datos financieros.

### 1.1.3. Modelo GJR-GARCH

Otra característica importante de las series de datos financieros es que la volatilidad condicional es asimétrica, es por ello que en Glosten, Jagannathan y Runkle (1993) introdujeron una modificación a un modelo GARCH, en la cual, consideran que los choques positivos y negativos tienen diferente impacto en la serie de tiempo.

El modelo referido es conocido como  $GJR-GARCH$ , por los nombres de los autores, el cual podemos escribir de la siguiente manera:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \epsilon_t^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_i I_{t-i} \epsilon_{t-i}^2$$

donde:

$$I_{t-i} = \begin{cases} 1, & \text{si } \epsilon_{t-i} < 0 \\ 0, & \text{si } \epsilon_{t-i} \geq 0 \end{cases} \quad (1.1)$$

Notemos que a diferencia de un modelo GARCH original, que sólo considera el término al cuadrado de la volatilidad, en este caso la función indicadora permite diferenciar entre errores positivos y negativos. El coeficiente  $\gamma_i$  representa un peso adicional a errores negativos de la serie.

El modelo particular que se va a estudiar en este trabajo, combina un modelo ARMA y un modelo GJR-GARCH, donde este último se aplica para modelar el efecto de la volatilidad asimétrica. La expresión analítica del modelo completo se presenta en el siguiente capítulo.

## 1.2. Teoría de Valores Extremos (EVT)

Sea  $X : \Omega \rightarrow \mathbb{R}_+$  una variable aleatoria no negativa definida en un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$  con distribución  $F$ .

En la teoría de valores extremos hay dos enfoques principales, el primero consiste en considerar los valores máximos por bloques de la distribución de la variable  $X$ , es decir,

se está interesado en describir el comportamiento estadístico de la variable aleatoria

$$M_n = \max\{X_1, \dots, X_n\}$$

donde  $X_1, \dots, X_n$  es una secuencia de variables aleatorias independientes con distribución  $F$  y  $M_n$  puede ser interpretado como el máximo de las observaciones en un bloque de valores fijado previamente.

El segundo enfoque, el cual consideraremos en este trabajo, consiste en seleccionar un umbral  $u \in \mathbb{R}_+$  tal que determina el *exceso de  $X$  bajo el umbral  $u$*  como la variable aleatoria,

$$X - u$$

para los valores de  $X$  tales que  $X > u$ , estos valores son conocidos como la excedencia de  $X$  sobre  $u$ . La función de distribución condicional de los excesos de  $X$  bajo  $u$  está dada por,

$$F_u(y) = \mathbb{P}(X - u \leq y | X > u) = \frac{F(u + y) - F(u)}{1 - F(u)} \quad (1.2)$$

Notemos que la función  $F_u$  describe la distribución del valor en exceso por encima del umbral cuando el umbral  $u$  es sobrepasado.

Sean  $(X_1, \dots, X_n)$  una sucesión de variables aleatorias independientes e idénticamente distribuidas y  $F_u$  su función de distribución condicional del exceso. De acuerdo con el teorema de *Pickands (1975)*, *Balkema y de Haan (1974)* se tiene que para la clase de distribuciones  $F$  que pertenecen a algún *dominio de atracción maximal*, la función de distribución condicional de los excesos  $F_u(y)$ , para un umbral  $u$  adecuado se satisface que,

$$F_u(y) \approx G_{\xi, \beta}(y), \quad u \rightarrow \infty \quad (1.3)$$

donde  $G_{\xi, \beta}(y)$  es la distribución Pareto Generalizada. Ésta función está definida como,

$$G_{\xi, \beta}(y) = \begin{cases} 1 - (1 + \xi y / \beta)^{-1/\xi}, & \text{si } \xi \neq 0, \\ 1 - \exp(-y/\beta), & \text{si } \xi = 0, \end{cases} \quad (1.4)$$

donde  $\beta > 0$  y el soporte de la distribución es  $[0, \infty)$  cuando  $\xi \geq 0$  y  $[0, -\beta/\xi]$  cuando  $\xi < 0$ .

De manera más general, este resultado establece que para una clase grande de funciones de distribuciones  $F$  es posible encontrar una función medible  $\beta(u)$  tal que

$$\lim_{u \rightarrow x_0^-} \sup_{0 \leq y < x_0 - u} |F_u(y) - G_{\xi, \beta(u)}(y)| = 0$$

Este resultado puede revisarse en el teorema (3.4.13) de *Embrechts et.al.(1997)*. Es importante mencionar que no todas las distribuciones pueden aproximarse por una

distribución Pareto Generalizada, para ello es necesario que la distribución pertenezca a algún dominio de atracción de las distribuciones de valores extremos: Gumbel, Weibull o Fréchet; este es un problema muy estudiado en la práctica.

A continuación definimos la función media de excesos de la variable  $X$ , la cual será de utilidad para seleccionar un umbral  $u$  adecuado, con base en el resultado límite anterior, tal que determine la cola de la distribución y sea razonable estimarla con una función de distribución Pareto Generalizada.

**Definición 1.2.1.** Sea  $X$  una variable aleatoria no negativa y  $F$  su función de distribución. Para un umbral  $u > 0$  tal que  $\mathbb{P}[X > u] > 0$ , se define la *función media de excesos* de la variable  $X$ , como la función  $e_F : [u, \infty) \rightarrow [0, \infty)$  dada por,

$$e_F(u) = \mathbb{E}[X - u | X > u].$$

En la teoría de valores extremos es de gran importancia estudiar el comportamiento de ésta función; tanto para determinar el tipo de cola de la distribución como el dominio de atracción con base en las observaciones, como también para la selección del umbral adecuado.

### 1.3. Cópulas

Una función cópula relaciona las densidades marginales a partir de una distribución multivariada. Una cópula puede ser interpretada como una distribución de probabilidad multivariada en la que la distribución de probabilidad marginal de cada variable puede ser "separada" de la estructura de dependencia entre las variables aleatorias. Formalmente,

**Definición 1.3.1.** Una cópula es una función  $C : [0, 1]^d \rightarrow [0, 1]$  que satisface las siguientes propiedades,

1.  $C$  es no decreciente en cada argumento.
2.  $C(1, 1, \dots, 1, u_k, 1, \dots, 1) = u_k$  para cada  $k \leq d$  y  $\forall u_k \in [0, 1]$  y para cualquier  $k \leq d$ ,

$$C(u_1, \dots, 0_k, \dots, u_d) = 0.$$

3.  $C$  es  $n$ -creciente, es decir, para todo  $a = (a_1, \dots, a_n)$ ,  $b = (b_1, \dots, b_n) \in [0, 1]^n$  con  $a_i \leq b_i$ ,  $i = 1, \dots, n$  se tiene que,

$$\sum_{j_1=1}^2 \dots \sum_{j_n=1}^2 (-1)^{j_1+\dots+j_d} C(u_{u_{1j_1}}, \dots, u_{u_{dj_d}}) \geq 0,$$

donde  $u_{ij_1} = a_i$ ,  $u_{ij_2} = b_i$ , para todo  $i = 1, \dots, d$ .

El resultado más importante en esta teoría es el Teorema de Sklar, el cual establece que cualquier distribución conjunta  $F_{\mathbf{X}}$  puede ser representada como una función cópula  $C(\cdot)$  de sus distribuciones marginales univariadas, esta relación entre las distribuciones marginales y la distribución conjunta es de gran utilidad en la teoría de probabilidad y estadística.

**Teorema 1.3.2** (Teorema de Sklar, 1959). *Sea  $F_{\mathbf{X}}$  una función de distribución conjunta con marginales  $F_1, \dots, F_d$ . Entonces existe una cópula  $C$  tal que para cualquier vector  $(x_1, \dots, x_d)$  en el dominio de  $F_{\mathbf{X}}$ , se tiene que*

$$F_{\mathbf{X}}(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (1.5)$$

*Si las distribuciones marginales son continuas,  $C(\cdot)$  está únicamente determinada. Recíprocamente, si  $C$  es una cópula y  $F_1(x_1), \dots, F_d(x_d)$  son funciones de distribución, entonces la función  $C$  definida en (1.5) es una función de distribución conjunta.*

Por lo tanto, cualquier función cópula que toma un conjunto de cualesquiera distribuciones marginales univariadas  $\{F_i(x_i)\}$  como sus argumentos define una distribución de probabilidad conjunta con marginales  $\{F_i(x_i)\}$ .

Supongamos que la función  $C$  tiene derivadas parciales de orden  $d$ , entonces la función de densidad conjunta puede ser derivada de la función cópula usando la regla de la cadena y se obtiene que,

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\partial^d C(F_1(x_1), \dots, F_d(x_d))}{\partial F_1(x_1), \dots, \partial F_d(x_d)} \prod_i^d f_i(x_i) = c(F_1(x_1), \dots, F_d(x_d)) \prod_{i=1}^d f_i(x_i) \quad (1.6)$$

donde  $c(\cdot)$  es llamada la *densidad de la cópula*.

En la práctica, se aplican distintos tipos de funciones cópulas, mencionaremos aquí algunos que vamos a usar más adelante para el estudio de la dependencia entre variables aleatorias.

### 1.3.1. Algunos tipos de funciones cópula

En la literatura no hay una clasificación concreta de los tipos de funciones cópula, sin embargo, a continuación se presentan algunos tipos de éstas en función de su relación de dependencia (McNeil *et.al.* 2001).

#### Cópulas Elípticas

La característica más importante de las cópulas elípticas es que representan relaciones de dependencia simétricas, sin importar que se analice la cola izquierda o derecha de las distribuciones implicadas. La función de distribución de éstas cópulas es una función elíptica.

**Definición 1.3.3.** Si  $\mathbf{X}$  es un vector aleatorio  $n$ -dimensional y para algún  $\mu \in \mathbb{R}^n$  y alguna matriz  $\Sigma$  de dimensión  $n \times n$ , simétrica y no negativa definida tales que la función característica  $\phi_{\mathbf{X}-\mu}(t)$  de  $\mathbf{X} - \mu$  es una función de la forma cuadrática  $t^T \Sigma t$ , es decir,  $\phi_{\mathbf{X}-\mu}(t) = \psi(t^T \Sigma t)$ . Decimos que  $\mathbf{X}$  tiene distribución elíptica con parámetros  $\mu$ ,  $\Sigma$ , y  $\psi$ , escribimos  $\mathbf{X} \sim E_n(\mu, \Sigma, \psi)$ .

Las cópulas Normales y Student son ejemplos de cópulas elípticas, las cuales son simétricas y su aplicación es relativamente simple, ya que se conocen bien las distribuciones que están asociadas a ellas.

### Cópula Gaussiana

La cópula Gaussiana o Normal es la función de dependencia asociada a la distribución normal multidimensional. Para el caso de dos variables la función de distribución tiene la forma siguiente,

$$\begin{aligned} C(u_1, u_2) &= \Psi_\rho(\Psi^{-1}(u_1), \Psi^{-1}(u_2)) \\ &= \int_{-\infty}^{\Psi^{-1}(u_1)} \int_{-\infty}^{\Psi^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{-1/2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dx dy \end{aligned} \quad (1.7)$$

donde  $\Psi_\rho$  es la distribución normal bivariada estándar, con coeficiente de correlación lineal  $\rho$  entre las variables  $X$  y  $Y$ .

### Cópula t-student

De manera análoga a la anterior, la cópula  $t$  - Student con  $\nu$  grados es la función de dependencia asociada a la distribución  $t$  multidimensional. Por ejemplo, en el caso bivariado:

$$\begin{aligned} C(u_1, u_2; \rho, \nu) &= t_{\nu, \rho}(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)) \\ &= \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)}\right)^{-\frac{(\nu+2)}{2}} dx dy \end{aligned} \quad (1.8)$$

donde  $t_{\nu, \rho}$  es la distribución  $t$ -Student con  $\nu$  grados de libertad,  $\rho$  es el coeficiente de correlación de Pearson.

### Cópulas arquimedianas

Este tipo de funciones admite una gran diversidad de funciones cópula, a diferencia de las elípticas (simétricas), las cópulas arquimedianas permiten recoger muchos tipos de estructuras de dependencia adicionales. Además estas funciones pueden ser construidas fácilmente.



### Cópula Clayton

La cópula Clayton es un tipo de cópula arquimediana asimétrica, donde la cola izquierda es más pesada. La función de distribución está dada por,

$$C(u_1, u_2; \theta) = \max \left\{ (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, 0 \right\} \quad (1.9)$$

donde  $\theta$  es el parámetro de la cópula, cuando este parámetro tiende a infinito se muestra que hay una mayor dependencia; mientras que si es cercano a cero muestra menor dependencia entre las variables.

### Cópula Frank

Es un tipo de cópula arquimediana simétrica, cuya función de distribución bivariada está dada por:

$$C(u_1, u_2) = -\frac{1}{\theta} \ln \left( 1 + \frac{(\exp(-\theta u_1) - 1)(\exp(-\theta u_2) - 1)}{\exp(-\theta) - 1} \right) \quad (1.10)$$

El comportamiento en las colas de esta función es parecido a la cópula Gaussiana, sin embargo, muestra mayor dependencia entre los valores centrales.

En la práctica, para modelar la dependencia entre variables aleatorias, dada una muestra de observaciones de ellas, primero se selecciona una función cópula adecuada y con base en las observaciones se estiman los parámetros de la función cópula elegida. Algunos métodos para estimar los parámetros de la cópula así como de sus distribuciones marginales usan máxima verosimilitud o método de momentos.

En la siguiente sección se presenta un método para la estimación de los parámetros de una cópula, el cual simultáneamente estima los parámetros de las distribuciones marginales.

#### 1.3.2. Estimación de Parámetros

En esta sección explicamos el método *Inference Function for Margin* (IFM) para la estimación de los parámetros de una cópula así como de los parámetros de las distribuciones marginales de manera independiente, este método fue presentado por Joe & Xu (1996).

Consideremos  $\mathbf{X} = (X_1, \dots, X_d)$  un vector aleatorio con distribución conjunta  $F$  y distribuciones marginales dadas por  $F_1(x_1; \alpha_1), \dots, F_d(x_d; \alpha_d)$ ; notemos que  $\alpha_j$  representa el vector de parámetros para cada una de las distribuciones marginales  $F_j$  con  $j = 1, \dots, d$ .

Supongamos que se tiene una muestra de este vector aleatorio, de tamaño  $n$ , dada por  $\{\mathbf{x}_i = (x_{i1}, \dots, x_{id})\}_{i=1}^n$ . Luego, con el objetivo de describir la dependencia entre las variables aleatorias  $X_1, \dots, X_d$  y aplicando el teorema de Sklar, tenemos que existe una cópula  $C$  tal que podemos escribir la función de distribución multivariada  $F$  del vector  $\mathbf{X} = (X_1, \dots, X_d)$  en términos de sus distribuciones marginales, como

$$F(\mathbf{x}; \alpha_1, \alpha_2, \dots, \alpha_d, \theta) = C(F_1(x_1; \alpha_1), F_2(x_2; \alpha_2), \dots, F_d(x_d; \alpha_d); \theta),$$

donde  $\theta$  representa los parámetros de un determinado tipo de función cópula, como los que mencionamos en la sección anterior.

Además se tiene que si las distribuciones marginales  $F_j$  son continuas, entonces la cópula está determinada de manera única por,

$$C(u_1, u_2, \dots, u_d) = F(F_1^{-1}(x_1), F_2^{-1}(x_2), \dots, F_d^{-1}(x_d))$$

Para seleccionar una función cópula adecuada, primero se selecciona una familia paramétrica de cópulas,

$$C(\mathbf{u}|\theta), \quad \mathbf{u} \equiv (u_1, u_2, \dots, u_d) \in [0, 1]^d$$

parametrizada por un vector finito de parámetros,  $\theta \in \Theta \subset \mathbb{R}^p$  con  $p \geq 1$ .

Dado que si la función  $C$  tiene derivadas de orden  $d$ , entonces mediante regla de la cadena, podemos escribir la densidad conjunta  $f(\mathbf{x})$  como,

$$f(\mathbf{x}; \alpha_1, \alpha_2, \dots, \alpha_d, \theta) = \frac{\partial^d C(F_1(x_1; \alpha_1), \dots, F_d(x_d; \alpha_d); \theta)}{\partial F_1(x_1; \alpha_1) \partial F_2(x_2; \alpha_2) \dots \partial F_d(x_d; \alpha_d)} \prod_{j=1}^d f_j(x_j; \alpha_j)$$

Luego, para cada variable  $X_j$  con  $j = 1, \dots, d$ ; podemos considerar las observaciones  $x_{1j}, \dots, x_{nj}$  como una muestra de tamaño  $n$  para dicha variable y de este modo, se tienen  $d$  funciones de log-verosimilitud para las distribuciones marginales univariadas de la forma,

$$L_j(\alpha_j) = \sum_{i=1}^n \log f_j(x_{i,j}; \alpha_j), \quad j = 1, \dots, d.$$

De manera análoga, suponiendo observaciones independientes entonces la función de verosimilitud es de la forma,

$$L(\theta; \alpha_1, \alpha_2, \dots, \alpha_d) = \sum_{i=1}^n \log f(\mathbf{x}_i; \alpha_1, \alpha_2, \dots, \alpha_d, \theta) \tag{1.11}$$

Una manera de estimar los parámetros, es encontrar  $\hat{\theta}, \hat{\alpha}_1, \dots, \hat{\alpha}_d$  tales que maximicen la función  $L(\cdot)$ ; sin embargo, realizar ésta optimización puede ser complicada

incluso numéricamente.

Una alternativa a éste procedimiento es el método IFM, el cual consiste en realizar las  $d$  optimizaciones de las funciones de log-verosimilitud marginales univariadas y luego realizar la optimización de la función de verosimilitud conjunta como función del parámetro  $\theta$ . Este algoritmo se puede resumir en los siguientes pasos:

- a. Maximizar cada función de log-verosimilitud marginal  $L_j(\alpha_j)$ , para  $j = 1, \dots, d$ . Así, obtener estimadores  $\tilde{\alpha}_1, \dots, \tilde{\alpha}_d$ .
- b. Considerar los estimadores  $\tilde{\alpha}_1, \dots, \tilde{\alpha}_d$  del paso anterior, y luego, maximizar la función  $L(\theta; \tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_d)$  como función de  $\theta$  para encontrar el estimador  $\tilde{\theta}$ .

Este procedimiento es computacionalmente más simple que estimar todos los parámetros  $\theta, \alpha_1, \alpha_2, \dots, \alpha_d$  simultáneamente de la función de verosimilitud en (1.11).

En el trabajo de Patton (2006) se prueba que el estimador  $\tilde{\theta} = (\tilde{\theta}, \tilde{\alpha}_1, \dots, \tilde{\alpha}_d)$  obtenido con éste método satisface, bajo ciertas condiciones de regularidad, la propiedad de ser asintóticamente normal.

#### 1.3.3. Dependencia entre variables aleatorias

Una de las ventajas de las cópulas es que éstas funciones han servido para definir nuevas medidas de dependencia entre variables aleatorias, como alternativa al coeficiente de correlación lineal de Pearson, que definimos más adelante. Del mismo modo, se definen algunas de éstas medidas de dependencia entre variables aleatorias basadas en cópulas.

##### Coeficiente de correlación de Pearson

Consideremos  $X$  y  $Y$  variables aleatorias tales que su media está dada por  $\mu_X$  y  $\mu_Y$ , respectivamente. Se define el coeficiente de correlación de Pearson como,

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\mathbb{E}[(X - \mu_X)^2] \mathbb{E}[(Y - \mu_Y)^2]}$$

donde  $\sigma_{XY}$  representa la covarianza de  $(X, Y)$  y  $\sigma_X, \sigma_Y$  son las desviaciones estándar de  $X$  e  $Y$ , respectivamente.

Se tiene que mientras más cercano a 1 sea el valor de  $|\rho_{X,Y}|$  mayor será la dependencia lineal en el siguiente sentido: si  $\rho_{X,Y}$  es positivo, entonces  $X$  crece conforme crece  $Y$  y si  $\rho_{X,Y}$  es negativo entonces las variables se mueven de forma lineal en sentido opuesto. Sin embargo, si ocurre que  $\rho_{X,Y} = 0$ , esto sólo indica que no hay una relación de dependencia lineal entre  $X$  y  $Y$ , pero no quiere decir que no exista otro tipo de

relación (no lineal).

En este sentido, dado que las funciones cópula permiten extraer la dependencia entre variables aleatorias, donde la dependencia no necesariamente es lineal, entonces parece razonable establecer medidas de dependencia basadas en cópulas. Una de estas medidas es la que definimos a continuación.

### $\tau$ de Kendall

Consideremos  $X_1, X_2$  variables aleatorias; la  $\tau$  de Kendall es una medida de dependencia la cual puede deducirse a través de la cópula asociada a estas variables como,

$$\tau = 4 \int_0^1 \int_0^1 C_\theta(u_1, u_2) dC_\theta(u_1, u_2) - 1$$

Este valor puede estimarse a partir de una muestra  $(x_{1i}, x_{2i})$ ,  $i = 1, \dots, N$  de las variables  $(X_1, X_2)$ . Aquí se considera el orden en que se observaron los elementos de la muestra, así para cada par de observaciones consecutivas  $(x_{1i}, x_{2i})$  y  $(x_{1j}, x_{2j})$  se observa si  $x_{1i} \geq x_{1j}$  y  $x_{2i} \geq x_{2j}$  o si  $x_{1i} \leq x_{1j}$  y  $x_{2i} \leq x_{2j}$ , en estos casos se dice que las observaciones son concordantes y en caso contrario se dice que son discordantes, excepto cuando  $x_{1i} = x_{1j}$  y  $x_{2i} = x_{2j}$  pues en este caso no son de ningún tipo. De este modo, la  $\tau$  de Kendall puede estimarse como,

$$\hat{\tau} = \frac{\#\text{pares concordantes} - \#\text{pares discordantes}}{\frac{1}{2}N(N-1)}$$

Vamos a utilizar ésta medida de dependencia para elegir los pares de variables aleatorias que representan una mayor dependencia para la construcción de una distribución multivariada por medio de cópulas por pares en el siguiente capítulo.

## 1.4. Pruebas de bondad de Ajuste

En esta sección presentamos los elementos de algunas pruebas para evaluar la bondad de ajuste del modelo propuesto y para su evaluación.

### Ljung-Box test

La prueba Box-Ljung (1978) se aplica para comprobar el ajuste de un modelo de serie de tiempo. Esta prueba se aplica a los residuos de la serie después de ajustar el modelo.

La prueba examina  $m$  autocorrelaciones de los residuos. Si las autocorrelaciones son muy pequeñas, se concluye que el modelo no presenta una falta de ajuste significativa. Podemos definir esta prueba como sigue:

- $\mathcal{H}_0$ : El modelo no presenta falta de ajuste.
- $\mathcal{H}_a$ : El modelo presenta falta de ajuste.
- *Estadística de prueba*: Dada una serie de tiempo  $(Y_t)$  de longitud  $n$ , se define la estadística de prueba como:

$$Q = n(n + 2) \sum_{k=1}^m \frac{\hat{r}_k^2}{n - k}$$

donde  $\hat{r}_k^2$  es la autocorrelación estimada de la serie en el  $k$ -ésimo retardo y  $m$  es el número de retrasos considerados.

- *Región crítica*: La prueba Box-Ljung rechaza la hipótesis nula si,

$$Q > \chi_{1-\alpha, h}^2$$

donde  $\chi_{1-\alpha, h}^2$  es la distribución Xi-cuadrada con  $h$  grados de libertad y nivel de significancia  $\alpha$ .

En el caso de aplicar un modelo  $ARMA(p, q)$  y dado que se consideran los residuales, entonces  $h = m - p - q$ .

### Jarque-Bera test

La prueba Jarque-Bera es una prueba de bondad de ajuste para comprobar si los datos de una muestra siguen una distribución normal, se basa en las características de simetría y curtosis de esta distribución.

Podemos definir esta prueba como sigue:

- $\mathcal{H}_0$ : Los datos siguen una distribución normal.
- $\mathcal{H}_a$ : Los datos siguen una distribución normal.
- *Estadística de prueba*:

$$JB = \frac{n - k + 1}{6} \left( S^2 + \frac{1}{4}(C - 3)^2 \right)$$

donde  $n$  es el número de observaciones,  $k$  es el número de retardos considerados,  $S$  es la asimetría de la muestra,  $C$  es la curtosis de la muestra; estas últimas se calculan como sigue:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}},$$

$$C = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2},$$

$\hat{\sigma}^2$ ,  $\hat{\mu}_3$  y  $\hat{\mu}_4$  representan las estimaciones de los momentos centrales segundo, tercero y cuarto, respectivamente.

El estadístico de Jarque-Bera se distribuye asintóticamente como una distribución  $\chi^2$  con dos grados de libertad.

## Durbin-Watson test

La prueba de hipótesis Durbin-Watson es la más usada para detectar la presencia de autocorrelación en modelos de series de tiempo. Esta prueba permite contrastar la hipótesis de no autocorrelación frente a la hipótesis alternativa de autocorrelación de primer orden bajo un esquema autoregresivo  $AR(1)$ , es decir

$$x_t = c + \phi_1 x_{t-1} + \epsilon_t$$

De manera analítica, estas hipótesis están dadas por

- $\mathcal{H}_0: \phi_1 = 0$ .
- $\mathcal{H}_a: 0 < |\phi_1| < 1$ .
- *Estadística de prueba:*

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

Donde  $n$  es el número de observaciones. Puesto que  $d$  es aproximadamente igual a  $2(1 - r)$ , donde  $r$  es la autocorrelación de la muestra de los residuos entonces  $d = 2$  indica que no hay autocorrelación.

### 1.4.1. Backtesting

Dado que el VaR no es observable, son necesarias algunas herramientas estadísticas que permitan determinar si la estimación de esta medida de riesgo es razonable. En finanzas, una validación de este tipo se conoce como *Backtesting*, este procedimiento incluye principalmente dos pruebas de hipótesis, la prueba de Kupiec y la de Christoffersen, las cuales presentamos a continuación.

Una manera intuitiva para comprobar la bondad del modelo propuesto es observar cuál es la proporción de períodos de la muestra en que se observa una pérdida superior a la predicción del modelo; es decir, cuando la pérdida observada con los valores en la muestra es superior al VaR. Dicha proporción debería ser en promedio igual al nivel de significancia considerado. Para ello se considera la siguiente función indicadora,

$$I_{t+1} = \begin{cases} 1, & \text{si } L_{t+1} < VaR(q) \\ 0, & \text{si } L_{t+1} \geq VaR(q) \end{cases} \quad (1.12)$$

donde  $L_{t+1}$  es la pérdida del portafolio observada al tiempo  $t + 1$ . Se define  $K = \sum_{i=1}^T I_t$ , el número de excedencias en un periodo de tiempo establecido donde la pérdida observada del capital excede la estimación del VaR.

### Prueba no condicional de Kupiec

La prueba de Kupiec (1995) evalúa la hipótesis nula  $H_0 : \hat{q} = \alpha$  donde  $\hat{q} = \frac{K}{T}$  es la proporción de excesos mayores al VaR en los periodos  $\{1, \dots, T\}$ , a los cuales nos referiremos como excepciones del VaR. La estadística de prueba de Kupiec está dada por,

$$t_\nu = \frac{\hat{q} - \alpha}{\sqrt{\hat{q}(1 - \hat{q})/T}}$$

ésta estadística se distribuye asintóticamente como una distribución  $t$  con  $T - 1$  grados de libertad [Kupiec (1995)].

### Prueba condicional de Christoffersen

Christoffersen propone otro método para evaluar las estimaciones del VaR, en su trabajo de 1998 introduce una prueba condicional, esta prueba parte de observar que si una aproximación para la estimación del VaR captura de manera precisa la distribución condicional de los retornos y sus propiedades dinámicas, entonces, las excepciones deben ser impredecibles.

Como antes, consideremos  $\hat{q}$  como la proporción de excepciones del VaR en los periodos  $\{1, \dots, T\}$ . Supongamos que  $K$  es el número de estas excepciones, entonces la variable  $K$  es una suma de las variables aleatorias Bernoulli  $I_{t+1}$  de la ecuación (1.12), por lo tanto,  $K$  tiene distribución binomial dada por

$$P(K) = \binom{T}{K} q^K (1 - q)^{T-K}$$

Entonces por una parte se quiere probar la hipótesis  $H_0 : \hat{q} = \alpha$ , para ello se considera como estadística de prueba la función de verosimilitud siguiente,

$$LR_{UC} = 2 \log \left[ \left(1 - \frac{K}{T}\right)^{T-K} \left(\frac{K}{T}\right)^K \right] - 2 \log [(1 - \alpha)^{T-K} \alpha^K]$$

Esta estadística se distribuye asintóticamente como una distribución Chi-cuadrada con un grado de libertad. Esta parte es equivalente a la prueba de Kupiec, sin embargo la prueba de Christoffersen considera además la independencia del modelo por medio del siguiente estadístico

$$LR_{IND} = -2 \log [(1 - \pi_0)^{n_{00} + n_{11}}] + 2 \log [(1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}}]$$

donde  $n_{ij}$  es el número de días en que el estado  $i$ -ésimo es seguido del estado  $j$ -ésimo, con  $i, j \in 0, 1$ , consideramos cero el estado en el cual la pérdida del portafolio es menor que el VaR estimado, y 1 cuando el retorno actual es mayor que el VaR estimado. Este estadístico sigue una distribución Chi-cuadrado con un grado de libertad. Por lo tanto, en la prueba de Christoffersen se considera tanto la condición de cobertura como la de independencia, el estadístico de esta prueba es

$$\begin{aligned} LR_{CC} &= LR_{UC} + LR_{IND} \\ &= 2 \log [(1 - \pi_0)^{n_{00}} \pi_0^{n_{01}} (1 - \pi_1)^{n_{10}} \pi_1^{n_{11}}] - 2 \log [(1 - \alpha)^{T-K} \alpha^K] \quad (1.13) \end{aligned}$$

donde  $\pi_i$  denota la probabilidad que la excepción ocurra al tiempo  $t + 1$  condicionando sobre el estado  $i$  al tiempo  $t$ , entonces

$$\pi_0 = \frac{n_{01}}{n_{00} + n_{01}} \quad \text{y} \quad \pi_1 = \frac{n_{11}}{n_{10} + n_{11}}$$

Por su definición, la estadística  $LR_{CC}$  tiene una distribución chi-cuadrada con dos grados de libertad.



# Capítulo 2

## Modelo estadístico para el portafolio

El objetivo de este capítulo es plantear un modelo estadístico para la distribución conjunta de los rendimientos de un portafolio de divisas, para después establecer y calcular medidas de riesgo del portafolio con base en dicha distribución.

Consideremos un mercado con  $d$  activos cuyos rendimientos al tiempo  $t$  están dados por  $r_t = (r_t^1, \dots, r_t^d)$ , supongamos que  $F_{r_t}$  es la distribución de este vector aleatorio. Sea  $\omega$  un portafolio para estos activos dado por,

$$\omega = (\omega_1, \dots, \omega_d) \text{ tal que, } \sum_{i=1}^d \omega_i = 1, \quad 0 \leq \omega_i \leq 1, \quad i = 1, \dots, d,$$

donde  $\omega_i$  representa la proporción del capital a invertir en el  $i$ -ésimo activo durante el  $t$ -ésimo periodo. Supongamos que el capital inicial está dado por  $x_0 = x$ , entonces el capital al tiempo  $t$  está dado por la variable aleatoria,

$$X_t = x_{t-1} \cdot \langle r_t, \omega \rangle = x_{t-1} \cdot \sum_{i=1}^d r_t^i \cdot \omega_i, \quad (2.1)$$

donde  $x_t$  es el capital observado al tiempo  $t$ .

En este trabajo se va a considerar un portafolio formado por cuatro divisas: euro (EUR), dólar estadounidense (USD), yen japonés (JPY) y libra esterlina (GBP). Se han seleccionado estas monedas por ser las más valoradas en el mercado de divisas y además son usadas como monedas de reserva, es decir, son usadas en grandes cantidades por instituciones o gobiernos como parte de sus reservas internacionales. Para la uniformidad de los datos, la moneda que usaremos como numerario será el peso mexicano (MXN).

El trabajo que motiva este desarrollo es el presentado por *Omari et.al.* (2018), en el cual se calcula el VaR de un portafolio de divisas (euro, dólar estadounidense, libra esterlina y rand sudafricano). La metodología seguida por los autores de dicho trabajo, consiste primero en ajustar un modelo ARMA–GJR–GARCH para cada divisa, y con base en éste, se obtiene una serie de residuos estandarizados. Luego, a cada una de las series de residuos se le ajusta una distribución de valores extremos en las colas y en el interior un ajuste de densidad por kernel, lo que da lugar a las distribuciones marginales. La obtención de la distribución conjunta de los rendimientos se hace a través de funciones cópula.

En este capítulo primero describimos algunas de las propiedades de las medidas del riesgo que vamos a estudiar. Luego, en la sección (2.3) se discute el ajuste de las distribuciones marginales para los residuos estandarizados. Finalmente en (2.4) se construye la distribución conjunta del portafolio, para la cual se aplica un enfoque de construcción basado en cópulas por pares en la Sección 2.4.1.

## 2.1. Medidas de Riesgo VaR y CVaR

La cuantificación del riesgo es un problema central para la toma de decisiones en inversiones y otras operaciones financieras. El *riesgo financiero* hace referencia a la posibilidad de pérdida o ganancia, esto debido a cambios en los factores que afectan el valor de un activo. En finanzas hay diferentes tipos de riesgo, algunos de ellos son el *riesgo de mercado*, el *riesgo de crédito* y el *riesgo operativo*; aquí nos concentramos en el riesgo de mercado, el cual, como su nombre lo indica, se origina precisamente por las fluctuaciones del mercado. Éste tipo de riesgo se define como la posibilidad de tener pérdidas derivadas de los movimientos en los precios de mercado, incluyendo tasas de interés, tasas de cambio y cotización de las acciones. En todos los casos es de gran interés establecer y calcular medidas de riesgo.

Sea  $X$  una variable aleatoria que representa la pérdida o ganancia de una posición financiera, si consideramos  $\Omega$  como el espacio de los posibles escenarios para esta variable aleatoria entonces para  $\omega \in \Omega$ ,  $X(\omega) > 0$  representa una ganancia y  $X(\omega) < 0$  una pérdida. Denotemos por  $\mathcal{X}$  al conjunto de variables aleatorias de las pérdidas. Definimos una *medida de riesgo* como una función dada por,

$$\mathcal{R} : \mathcal{X} \rightarrow [0, \infty].$$

En la literatura hay varias medidas de riesgo, en algunos trabajos se ha estudiado propiedades deseables en estas funciones, por ejemplo, en el trabajo de *Arztner et.al.*(1997,1999) se introduce el concepto de medidas de riesgo coherentes por medio de axiomas que éstas deben satisfacer.

**Definición 2.1.1.** Una medida de riesgo  $\mathcal{R}$  es *coherente* si satisface las siguientes propiedades:

1. **Invarianza traslacional:** Para todo  $X \in \mathcal{X}$  y cualquier constante  $c \in \mathbb{R}$  se cumple que  $\mathcal{R}(X + c) = \mathcal{R}(X) - c$ . Esta propiedad dice que si se invierte una cantidad adicional en el portafolio, y se invierte prudentemente, entonces su riesgo se reduce proporcionalmente en esa cantidad.
2. **Homogeneidad positiva:** Sea  $c > 0$  una constante y  $X \in \mathcal{X}$ , entonces  $\mathcal{R}(c \cdot X) = c \cdot \mathcal{R}(X)$ . Esta propiedad se refiere a que si se aumenta la posición en un portafolio, o en alguno de los activos que lo componen, el riesgo debe incrementarse proporcionalmente.
3. **Monotonicidad:** Si  $X, Y \in \mathcal{X}$ , son tales que  $F_X(x) \geq F_Y(x)$  para todo  $x$ , entonces  $\mathcal{R}(X) \leq \mathcal{R}(Y)$ . Esta propiedad indica que a mayor rentabilidad debe corresponder un mayor riesgo.
4. **Subaditividad:** Para cualesquiera  $X$  y  $Y$  tales que  $X, Y, X + Y \in \mathcal{X}$ , se tiene que  $\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$ . En el trabajo de *Arztner* comentan esta propiedad como “*una fusión no crea un riesgo extra*”; es decir, el riesgo global de un portafolio formado por dos o más activos es menor o igual que la suma de los riesgos individuales. Como consecuencia de esta propiedad se tiene que la diversificación del portafolio no aumenta el riesgo.

Respecto al riesgo de mercado, algunas de las medidas más populares en la industria financiera, son el VaR (Valor en Riesgo) y CVaR (Valor en Riesgo Condicional). Supongamos que el capital al tiempo  $t$  está dado por la variable aleatoria  $X_t$  dada en la ecuación (2.1), estamos interesados en calcular estas medidas de riesgo para la variable  $X_t$ , las cuales definimos a continuación.

**Definición 2.1.2.** Dado un nivel de confianza  $\alpha \in (0, 1)$  se definen,

$$VaR_{t+1|t}(\alpha) = - \inf \{ l \in \mathbb{R} : F_{X_{t+1}}(l) \geq \alpha \},$$

$$CVaR_{t+1|t}(\alpha) = -\mathbb{E} [X_{t+1} \mid X_{t+1} \leq VaR_{t+1|t}(\alpha)]$$

Notemos que para calcular tales medidas de riesgo para el tiempo  $t + 1$  es necesario tener una estimación de la distribución del capital  $X_{t+1}$ .

En este trabajo estimaremos el VaR por el método de simulación Montecarlo para lo cual necesitamos estimar la distribución conjunta de los rendimientos del portafolio y con base en esta estimación generar una gran cantidad de simulaciones de los rendimientos y finalmente de manera empírica calcular el VaR y CVaR. Por lo tanto, nos

vamos a concentrar en el problema de estimar la distribución conjunta de los rendimientos del portafolio.

En el trabajo de *Artzner* (1999), sobre medidas de riesgo coherentes, se prueba que el VaR no es una medida de riesgo coherente (página 209) pues no satisface la propiedad de subaditividad, mientras que el CVaR sí es coherente; *Acerbi & Tasche* (2002) demuestran que el CVaR satisface las propiedades anteriores y por lo tanto, ésta es una medida de riesgo coherente.

En el trabajo de *Richtárik* (2015), página 19, se presenta el siguiente ejemplo para mostrar la no subaditividad del VaR:

*Ejemplo 2.1.3.* Considere dos activos  $A$  y  $B$ , cuyas pérdidas y probabilidades respectivas se muestran en la tabla de abajo,

	$\xi_1$	$\xi_2$	$\xi_3$
$P(\xi_i)$	0.04	0.04	0.92
$A$	1000	0	0
$B$	0	1000	0

Cuadro 2.1: Pérdidas de los activos  $A$  y  $B$  en tres diferentes escenarios.

De la definición del VaR y considerando que  $X$  es la variable aleatoria correspondiente a las pérdidas de cierta posición financiera, se tiene que para un nivel  $0 < \alpha < 1$ , lo podemos escribir como:

$$VaR_\alpha = \min\{c : P(X \leq c) \geq \alpha\}$$

Entonces, si calculamos el VaR con un nivel de confianza del 95 % para  $A$ ,  $B$  y  $A + B$  tenemos que,

$$\begin{aligned} VaR_{0.95}(A) &= \min\{c : P(A \leq c) \geq 0.95\} = 0, \text{ pues, } P(A \leq 0) = 0.96 \\ VaR_{0.95}(B) &= \min\{c : P(B \leq c) \geq 0.95\} = 0, \text{ pues, } P(B \leq 0) = 0.96 \\ VaR_{0.95}(A + B) &= \min\{c : P(A + B \leq c) \geq 0.95\} = 1000 \end{aligned}$$

Por lo tanto, aquí se tiene que  $VaR_{0.95}(A + B) \not\leq VaR_{0.95}(A) + VaR_{0.95}(B)$ , lo que muestra que el VaR no satisface la propiedad de subaditividad.

Para las medidas de riesgo que satisfacen esta propiedad, se puede tener la garantía de que el riesgo conjunto está acotado por la suma de los riesgos individuales. Por lo tanto, en el caso del VaR no es suficiente hacer una estimación de las distribuciones marginales de cada divisa y calcular su riesgo para tener esta cota, sino que además necesitamos estimar la distribución conjunta de los rendimientos del portafolio para calcular dicha medida de riesgo.

## 2.2. Modelo de series de tiempo para los retornos

Por simplicidad para la notación y el procedimiento que se presenta a continuación, se va a realizar la estimación de la distribución de rendimientos para un solo activo, es decir, tomamos el caso en que  $d = 1$ . Volveremos a considerar  $d > 1$  hasta la sección (2.4) en donde se hace el planteamiento para la distribución conjunta de los rendimientos cuando se consideran más de un activo.

Supongamos que el precio de una divisa al tiempo  $t$  está dado por  $p_t$ , entonces consideramos como retorno o rendimiento de dicha divisa al tiempo  $t$ , al logaritmo de la tasa de cambio en el precio,

$$r_t = \log \left( \frac{p_t}{p_{t-1}} \right). \quad (2.2)$$

De acuerdo con las medidas de riesgo que discutimos en la sección anterior, para su cálculo cuando  $d = 1$  es necesario contar con un modelo que capture la distribución de los rendimientos. Cuando se considera  $d > 1$  entonces se debe estimar la distribución del capital del portafolio como en la ecuación (2.1), lo que implica estimar la distribución conjunta de los rendimientos. En cualquier caso, se necesita estimar las distribuciones marginales de los activos, en este caso de divisas, que conforman el portafolio.

Con el fin de estimar dichas distribuciones, quisiéramos considerar los rendimientos históricos de cada serie como observaciones independientes; sin embargo, en algunos trabajos de investigación como en McNeil(2000) se muestra que los residuos estandarizados obtenidos después de ajustar un modelo de series de tiempo, presenta menos correlación que los rendimientos originales.

Por lo tanto, para cada una de las divisas seleccionadas, se va a ajustar un modelo de series de tiempo  $ARMA-GJR-GARCH$ . En el modelo  $ARMA$  (Sección 1.1.1) primero se filtra la correlación de la serie y luego el modelo  $GJR - GARCH$  (Sección 1.1.3) explica la volatilidad, haciendo diferencia entre los errores positivos y negativos, para compensar la asimetría de la volatilidad que se espera observar en estos datos.

El modelo completo  $ARMA(m, n) - GJR - GARCH(p, q)$  tiene la siguiente forma:

$$r_t = c + \sum_{i=1}^m \phi_i r_{t-i} + \sum_{j=1}^n \theta_j \epsilon_{t-j} + \epsilon_t \quad (2.3)$$

$$\epsilon_t = \sigma_t z_t \quad (2.4)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p (\alpha_i \epsilon_{t-i}^2 + \gamma_i I_{t-i} \epsilon_{t-i}^2) + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (2.5)$$

donde:

$$I_{t-i} = \begin{cases} 1, & \text{si } \epsilon_{t-i} < 0 \\ 0, & \text{si } \epsilon_{t-i} \geq 0 \end{cases} \quad (2.6)$$

Notemos que con base en este modelo podemos reescribir la serie de rendimientos  $r_t$  como,

$$r_t = \mu_t + z_t \sigma_t, \quad (2.7)$$

donde  $\sigma_t$  está dada como en (2.5) y

$$\mu_t = c + \sum_{i=1}^m \phi_i r_{t-i} + \sum_{j=1}^n \theta_j \epsilon_{t-j}. \quad (2.8)$$

Al considerar una secuencia de observaciones, para cada una de las divisas seleccionadas, de los rendimientos  $(r_1, r_2, \dots, r_t)$  como una serie de tiempo podemos realizar el ajuste del modelo anterior que consiste en estimar sus coeficientes mediante máxima verosimilitud. De este modo, primero se eligen arbitrariamente los parámetros  $(m, n, p, q)$  y, con éstos fijos, se estiman coeficientes  $\hat{\phi}_1, \dots, \hat{\phi}_m, \hat{\theta}_1, \dots, \hat{\theta}_n, \hat{\alpha}_1, \dots, \hat{\alpha}_p, \hat{\gamma}_1, \dots, \hat{\gamma}_p, \hat{\beta}_1, \dots, \hat{\beta}_q$ . Luego, con la estimación de estos parámetros se obtiene la verosimilitud y mediante el Criterio de Información de Akaike (AIC), el cual definimos en el primer capítulo, se elige alguna de las combinaciones de los parámetros  $(m, n, p, q)$  dentro de las que fueron evaluadas.

Una vez elegidos los parámetros  $m, n, p, q$  mediante el criterio AIC, y teniendo la estimación de los coeficientes del modelo, se obtienen los coeficientes  $\hat{\mu}_t$  y  $\hat{\sigma}_t$  como en la ecuaciones (2.8) y (2.5), respectivamente. Entonces, los residuos estandarizados del modelo están dados por

$$z_t = \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t}. \quad (2.9)$$

Con este procedimiento, obtenemos una serie de residuos estandarizados para cada una de las series de retornos de divisas con base en el modelo establecido para cada una de ellas. Dicho procedimiento para calcular los residuos estandarizados es el presentado por *Mc.Neil & Frey (2000)*.

## 2.3. Distribuciones marginales de las divisas

Recordando que el objetivo principal es calcular el  $VaR(\alpha)$  de un portafolio de divisas, hemos discutido que se necesita estimar la distribución conjunta de los rendimientos del portafolio. Para realizar una estimación de dicha distribución quisiéramos tener observaciones independientes de los rendimientos; sin embargo, en *Mc.Neil (2000)* se discute que los retornos presentan mayor correlación que la de los residuos estandarizados; obtenidos después de aplicar un modelo de series de tiempo como el estudiado en la sección anterior. Por ello, se va a ajustar un modelo ARMA–GJR–GARCH adecuado para cada divisa, y luego se obtienen las correspondientes series de residuos estandarizados. Los coeficientes del modelo de series de tiempo permiten que la variabilidad de las observaciones se describan mediante los residuos estandarizados, por lo

tanto, es de interés estimar la distribución conjunta de los residuos estandarizados de las series. Dicha distribución se construye con base en las distribuciones marginales y funciones cópula.

El ajuste de las distribuciones marginales para cada serie de residuos estandarizados representa un problema interesante desde el punto de vista de inferencia estadística. En muchos casos, la teoría de valores extremos es una herramienta adecuada para modelar el comportamiento de las colas en las series de datos financieros. Por lo tanto, aquí ajustaremos dos diferentes distribuciones, una distribución  $t$ -Student y una distribución cuyas colas son estimadas con una distribución Pareto Generalizada, que discutiremos más adelante.

Dadas las observaciones de rendimientos para cada una de las divisas, como observaciones a través del tiempo de una variable aleatoria  $r_t$  en un espacio de probabilidad inducido  $(\Omega, \mathcal{F}, \mathbb{P})$ . Así, podemos considerar a cada una de éstas series de observaciones como una serie de tiempo  $(r_t, t \in \mathbb{Z})$  y estimar los parámetros del modelo anterior. De acuerdo con la ecuación (2.9), tenemos que podemos considerar la dinámica de éste modelo dada por

$$r_t = \mu_t + \sigma_t Z_t. \quad (2.10)$$

donde la variable  $Z_t$  representa los residuos estandarizados del modelo; suponemos que éstas son independientes e idénticamente distribuidas (i.i.d.) con media cero y varianza unitaria y denotamos por  $F_Z(z)$  a su función de distribución. Además, suponemos que  $\mu_t$  y  $\sigma_t$  son medibles respecto a  $\mathcal{F}_{t-1}$ , la información disponible del proceso al tiempo  $t - 1$  y consideramos que  $Z_t$  es independiente de  $\mathcal{F}_{t-1}$ .

De acuerdo con esta relación, nos ocuparemos de estimar una distribución para las series de residuos estandarizados correspondientes a cada una de las series de divisas. Para ello, en la parte práctica aplicaremos dos ajustes: uno ajustando una distribución  $t$ -Student y un segundo ajuste con una distribución en la que se considera una estimación de valores extremos por umbral para la estimación de las colas de dicha distribución.

### 2.3.1. Ajuste de Valores Extremos por umbral

Al considerar las observaciones anteriores  $Z_t$  de los residuales estandarizados, se usará teoría de valores extremos para modelar la cola de la distribución  $F_Z$ . El ajuste para cola derecha e izquierda es análogo, por lo que aquí vamos a detallar solamente el análisis para la estimación de la cola derecha de la distribución.

Partimos del supuesto que se ha fijado un umbral  $u$  que determina las observaciones excedentes con las cuales se estimará la distribución para la cola derecha. En la siguiente sección se discute una manera de elegir este umbral. Suponemos además que las observaciones de residuos estandarizados pertenecen a algún dominio de atracción

de las distribuciones de valores extremos, por lo que se satisface el teorema de *Pickands* (1975) – *Balkema & de Haan* (1974), el cual fue introducido en la Sección 1.2; para mayores detalles respecto a los dominios de atracción de las distribuciones, se sugiere consultar dichos trabajos.

Denotemos por  $F_u(y)$  a la distribución del exceso condicional de  $Z_t$  bajo el umbral  $u$ , como en la ecuación (1.2); esta función está dada por

$$F_u(y) = P(Z - u \leq y | Z > u) = \frac{F(y + u) - F(u)}{1 - F(u)}, \quad (2.11)$$

para  $0 \leq y \leq z_{(n)} - u$  y  $z_n$  representa el valor más grande observado en una muestra de residuos estandarizados con la que se realiza la estimación. Notemos que para los puntos  $z > u$ , al evaluar  $F_u(z - u)$  y despejar la ecuación (2.11), obtenemos la relación

$$1 - F(z) = (1 - F(u))(1 - F_u(z - u)). \quad (2.12)$$

Suponemos que para una muestra de tamaño  $N$ , un número aleatorio de puntos  $N_u > 0$  exceden el umbral  $u$ . De este modo, si estimamos el lado derecho de la ecuación anterior usando la proporción aleatoria  $N_u/N$  de datos en la cola de la distribución y por otra parte estimamos el lado derecho de la ecuación usando la distribución Pareto Generalizada, cuya expresión fue dada en la ecuación (1.4), y reemplazando sus parámetros por estimadores de máxima verosimilitud, entonces obtenemos el estimador de la cola de la distribución como:

$$F(\hat{x}) = 1 - \frac{N_u}{N} \left( 1 - \hat{\xi} \frac{x - u}{\hat{\beta}} \right)^{-1/\hat{\xi}}, \quad x > u$$

En Smith (1987), en su Teorema 5.1, se prueba que los estimadores de máxima verosimilitud satisfacen las propiedades de consistencia y normalidad asintótica.

Con base en lo anterior, los residuos estandarizados se pueden modelar con una distribución Pareto Generalizada, y para el interior se considera una estimación de densidad por kernel Gaussiano. Por lo tanto, eligiendo umbrales  $u^L$  y  $u^R$  para las colas inferior y superior de la distribución marginal, respectivamente; se tiene que la distribución marginal para cada divisa se puede escribir como:

$$F(z) = \begin{cases} \frac{N_{u^L}}{N} \left( 1 + \xi^L \frac{u^L - z}{\beta^L} \right)^{-1/\xi^L}, & \text{si } z < u^L \\ \phi(z), & \text{si } u^L < z < u^R \\ 1 - \frac{N_{u^R}}{N} \left( 1 + \xi^R \frac{u^R - z}{\beta^R} \right)^{-1/\xi^R}, & \text{si } z > u^R \end{cases}$$

Donde  $\beta, \xi$  son parámetros de escala y forma, respectivamente;  $\phi(z)$  es la distribución estimada mediante un kernel Gaussiano en el intervalo  $[u^L, u^R]$ ; por último,  $N_{u^L}$  (análogamente  $N_{u^R}$ ) es la cantidad de datos que están por debajo de  $u^L$ .



### 2.3.2. Selección de Umbral para ajuste GPD

En la sección anterior comentamos sobre el ajuste de la distribución Pareto Generalizada (GPD), el cual para llevarse a cabo se requiere hacer una elección de umbrales  $u^R$  y  $u^L$  para determinar las colas derecha e izquierda, respectivamente, de la distribución de residuos estandarizados.

Dada una muestra de observaciones  $\{x_1, \dots, x_n\}$  y fijando un umbral  $u$ , definimos como *eventos extremos* a las observaciones  $\{x_i : x_i > u\}$ ; al ordenar estas observaciones las podemos etiquetar como  $x_{(1)}, x_{(2)}, \dots, x_{(k)}$ . Definimos como *excesos del umbral  $u$*  a las observaciones  $y_{(j)} = x_j - u$ . Suponemos que estas observaciones siguen una distribución Pareto Generalizada.

En lo siguiente nos centraremos en la elección del umbral que determina la cola derecha de la distribución; para el caso izquierdo se procede de manera análoga. Para la elección del umbral es necesario considerar que un umbral muy bajo aumentaría el sesgo del modelo, mientras que la elección de un umbral muy alto provocaría que la cantidad de excesos fueran pocas para la estimación de los parámetros de la distribución. Un método para elegir este umbral se basa en la función media de excesos.

Se sabe que si  $Y$  sigue una distribución Pareto Generalizada de parámetros  $\xi$  y  $\sigma$ , entonces la media de  $Y$  es,

$$\mathbb{E}[Y] = \frac{\sigma}{1 - \xi}, \quad \xi > 1.$$

Supongamos que la distribución Pareto Generalizada es adecuada para modelar los excesos de la variable aleatoria  $X$  bajo el umbral  $u_0$ , entonces

$$\mathbb{E}[X - u_0 | X > u_0] = \frac{\sigma_{u_0}}{1 - \xi}, \quad \xi > 1,$$

donde  $\sigma_{u_0}$  es el parámetro de escala de la distribución Pareto generalizada correspondiente a los excesos del umbral  $u_0$ . De aquí, si la distribución es válida para los excesos del umbral  $u_0$ , entonces será válida también al considerar un umbral  $u$  tal que  $u > u_0$ ,

$$\mathbb{E}[X - u | X > u] = \frac{\sigma_u}{1 - \xi} = \frac{\sigma_{u_0} + \xi u}{1 - \xi} = \frac{\sigma_{u_0}}{1 - \xi} + \frac{\xi}{1 - \xi} u.$$

Por lo tanto, de esta expresión vemos que se espera que esta función dependa linealmente del parámetro  $u$ , así como para elegir un umbral adecuado podemos explorar la función media de excesos empírica al variar  $u$ , la cuál está dada por:

$$\frac{1}{N_u} \sum_{i=1}^{N_u} (x_{(i)} - u). \quad (2.13)$$

Por lo tanto, al tener la gráfica de esta función variando el parámetro  $u$  se elige el valor más grande de  $u$  tal que el gráfico muestre una tendencia lineal.

### Elección de umbral conforme a la estabilidad del parámetro de forma de la distribución Pareto Generalizada

Otro método para seleccionar un umbral adecuado consiste en ajustar la distribución Pareto Generalizada en un rango de umbrales posibles, luego se elige el más razonable según la estabilidad del parámetro de forma estimado en cada caso.

La familia de distribuciones Pareto Generalizada tiene la siguiente característica: si dicha distribución es útil para modelar los excesos bajo un umbral  $u_0$ , entonces la excedencia de cualquier umbral  $u$  más alto (en el caso del ajuste de la cola derecha) también sigue una distribución Pareto Generalizada. Además los parámetros de escala de estas distribuciones satisfacen la siguiente relación,

$$\sigma_u = \sigma_{u_0} + \xi(u - u_0), \quad \xi \neq 0. \quad (2.14)$$

En el caso que  $\xi = 0$ , se considera la reparametrización  $\sigma^* = \sigma_u - \xi u$ .

Luego, para cada  $u$  dentro de cierto rango de posibles umbrales, se estiman los parámetros  $\hat{\sigma}$  y  $\hat{\xi}$  mediante el método de máxima verosimilitud. Se espera que la estimación del parámetro de forma sea constante por encima del umbral  $u_0$  en el caso que este umbral sea razonable para que los excesos correspondientes sigan una distribución Pareto Generalizada, esto nos proporciona una herramienta para la selección de umbrales adecuados.

En el siguiente capítulo se aplican estas técnicas para la selección de umbrales para los residuos estandarizados de cada divisa y así estimar las distribuciones marginales para cada una.

## 2.4. Distribución conjunta de los rendimientos del portafolio

Para construir la distribución conjunta de los rendimientos de las divisas que componen el portafolio, vamos a hacer uso de la estimación de distribuciones marginales que discutimos en la sección anterior. Luego, usando estas funciones marginales y un ajuste de funciones cópula se construye la distribución conjunta de los rendimientos del portafolio.

En esta sección consideramos que para cada serie de residuos, se han seleccionado umbrales  $u^L$  y  $u^R$  tales que determinan las colas izquierda y derecha de la distribución de residuos estandarizados, respectivamente. Luego se obtiene el ajuste de los parámetros de la Distribución Pareto Generalizada,  $\xi^L$  y  $\sigma^L$  para la cola izquierda y los parámetros  $\xi^U$  y  $\sigma^U$  para la cola derecha de esta distribución. Con base en éstos umbrales, la función de distribución marginal para cada divisa está dada como en la ecuación (2.13).

Considerando las distribuciones marginales que se tienen para cada divisa, se procede a modelar la dependencia con funciones cópula. Al aplicar funciones cópula es necesario transformar nuestras observaciones para que se encuentren dentro del intervalo  $(0, 1)$ , es por ello que hacemos referencia al siguiente teorema,

**Teorema 2.4.1** (Teorema de la Transformada Inversa). *Sea  $X$  una variable aleatoria con distribución  $F$ , continua e invertible, y sea  $F^{-1}$  su función inversa. Entonces la variable aleatoria  $U = F(X)$  tiene distribución uniforme en  $(0, 1)$ . Conversamente, si  $U$  es una variable aleatoria con distribución uniforme en  $(0, 1)$  entonces la variable aleatoria  $X = F^{-1}(U)$  se distribuye como  $F$ .*

Por lo tanto, si consideramos los residuos estandarizados  $z_t^i$ , para la  $i$ -ésima divisa y al considerar la transformación  $u_t^i = F^i(z_t^i)$  entonces los datos  $u_t^i$  tienen distribución uniforme en  $(0, 1)$ . Ahora queremos construir la distribución conjunta a partir del método de construcción basado en cópulas bivariadas que discutiremos en la siguiente sección.

### 2.4.1. Dependencia a través de cópulas por pares

El problema de estimar de funciones cópula para describir la dependencia entre variables aleatorias ha sido ampliamente estudiado. Sin embargo, la estimación de funciones cópulas es más sencilla para el caso bivariado y se tienen más herramientas de exploración en ése caso. De este modo, para aprovechar éstas herramientas y dado que representa mayor complejidad estimar los parámetros de cópulas multivariadas en general, se propone aplicar un método de construcción para la distribución conjunta basado en cópulas bivariadas y en las distribuciones marginales.

Joe (1996) fue el primero en dar una construcción probabilista de una función de distribución multivariada basada en cópulas bivariadas, luego Aas et. al. (2009) establecen el método PCC (pair-copula-construction) desde un enfoque de inferencia. En esta sección explicamos el procedimiento de construcción para un vector aleatorio de dimensión cuatro, de acuerdo con la dimensión del portafolio.

Recordemos que como en la ecuación (1.6), para el caso bivariado la función  $c_{1,2}(\cdot)$  representa la densidad de la cópula bivariada  $C_{1,2}(\cdot)$  dada por,

$$c_{1,2}(u_1, u_2) = \frac{\partial^2 C_{1,2}(u_1, u_2)}{\partial u_1 \partial u_2}$$

Entonces, por diferenciación se obtiene que la densidad conjunta bivariada de las variables aleatorias  $X_1$  y  $X_2$  se puede escribir como

$$f(x_1, x_2) = c_{1,2}(F(x_1), F(x_2)) \cdot f(x_1) \cdot f(x_2).$$

Para mostrar la construcción de la distribución conjunta, veamos primero el caso para un vector aleatorio  $X = (x_1, x_2, x_3)$  con distribución  $F$  y distribuciones marginales  $F_1, F_2, F_3$ , pues en este caso podemos escribir su función de densidad,

$$f(x_1, x_2, x_3) = f(x_1)f(x_2|x_1)f(x_3|x_1, x_2).$$

Luego, aplicando el Teorema de Sklar tenemos que,

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{c_{1,2}(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2)}{f_1(x_1)} = c_{1,2}(F_1(x_1), F_2(x_2)) f_2(x_2).$$

Por otra parte,

$$\begin{aligned} f(x_3|x_1, x_2) &= \frac{f(x_2, x_3|x_1)}{f(x_2|x_1)} \\ &= \frac{c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) f(x_2|x_1) f(x_3|x_1)}{f(x_2|x_1)} \\ &= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) f(x_3|x_1) \\ &= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) c_{1,3}(F_1(x_1), F_3(x_3)) f_3(x_3). \end{aligned} \quad (2.15)$$

Por lo tanto, tenemos que

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \\ &\quad \cdot c_{1,2}(F_1(x_1), F_2(x_2)) c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) \\ &\quad \cdot c_{1,3}(F_1(x_1), F_3(x_3)) \end{aligned} \quad (2.16)$$

Para el caso de un vector aleatorio de dimensión cuatro, en el apéndice del trabajo de (*Aas, 2016*) se llega a que la descomposición canónica de la función de densidad conjunta, la cual se obtiene de la siguiente forma:

Recordemos que las densidades condicionales se pueden escribir como

$$\begin{aligned} f(x_2|x_1) &= c_{1,2}(F_1(x_1), F_2(x_2)) f_2(x_2) \\ f(x_3|x_1, x_2) &= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) c_{1,3}(F_1(x_1), F_3(x_3)) f_3(x_3). \end{aligned}$$

De manera análoga se obtiene

$$\begin{aligned} f(x_4|x_1, x_2, x_3) &= \frac{f(x_3, x_4|x_1, x_2)}{f(x_3|x_1, x_2)} \\ &= \frac{c_{3,4|1,2}(F(x_3|x_1, x_2), F(x_4|x_1, x_2)) f(x_3|x_1, x_2) f(x_4|x_1, x_2)}{f(x_3|x_1, x_2)} \\ &= c_{3,4|1,2}(F(x_3|x_1, x_2), F(x_4|x_1, x_2)) f(x_4|x_1, x_2) \\ &= c_{3,4|1,2}(F(x_3|x_1, x_2), F(x_4|x_1, x_2)) \frac{f(x_2, x_4|x_1)}{f(x_2|x_1)}, \end{aligned}$$

donde,

$$\begin{aligned}
 \frac{f(x_2, x_4|x_1)}{f(x_2|x_1)} &= \frac{c_{2,4|1}(F(x_2|x_1), F(x_4|x_1)) f(x_2|x_1) f(x_4|x_1)}{f(x_2|x_1)} \\
 &= c_{2,4|1}(F(x_2|x_1), F(x_4|x_1)) f(x_4|x_1) \\
 &= c_{2,4|1}(F(x_2|x_1), F(x_4|x_1)) c_{1,4}(F_1(x_1), F_4(x_4)) f_4(x_4). \quad (2.17)
 \end{aligned}$$

Aplicando probabilidad condicional, podemos descomponer la densidad conjunta como

$$f(x_1, x_2, x_3, x_4) = f(x_1) f(x_2|x_1) f(x_3|x_1, x_2) f(x_4|x_1, x_3, x_2)$$

Luego, al sustituir las expresiones anteriores en esta última, obtenemos

$$\begin{aligned}
 f(x_1, x_2, x_3, x_4) &= f(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\
 &\quad \cdot c_{1,2}(F_1(x_1), F_2(x_2)) c_{1,3}(F_1(x_1), F_3(x_3)) \\
 &\quad \cdot c_{1,4}(F_1(x_1), F_4(x_4)) \\
 &\quad \cdot c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)) \cdot c_{2,4|1}(F(x_2|x_1), F(x_4|x_1)) \\
 &\quad \cdot c_{3,4|1,2}(F(x_3|x_1, x_2), F(x_4|x_1, x_2)) \quad (2.18)
 \end{aligned}$$

Por lo tanto, tenemos que a partir de la ecuación anterior es posible construir la densidad conjunta de un vector aleatorio de dimensión cuatro, ésta misma descomposición la aplicamos en particular para el vector aleatorio de los residuos estandarizados de las series que estamos considerando. La función de distribución para dicho vector aleatorio, la obtenemos mediante integración a partir de la ecuación (2.18).

Notemos que para la descomposición anterior se requieren las distribuciones marginales, que en el caso de los residuos estandarizados en la Sección 3.2 planteamos cómo estimarlas, y además se requiere estimar las funciones cópula bivariadas. Para ello se elige una función cópula de las presentadas en la Sección 1.3.1 y se estiman sus parámetros aplicando el método *Inference Function for Margin* presentado en el capítulo anterior. Ésto se hace para cada uno de los tipos de funciones cópula presentados y se elige la cópula que mejor describa los datos con base en su función de máxima verosimilitud.

En conclusión, esta descomposición nos permite, a través de cópulas por pares y sus distribuciones marginales, construir la distribución conjunta de los residuos estandarizados de los activos que conforman el portafolio.

## 2.5. Simulación Montecarlo para calcular el VaR

Supongamos que podemos obtener un número grande de simulaciones de los residuos estandarizados con base en la estimación de la función de distribución conjunta  $F_Z$  que construimos en la sección anterior. Cada una de estas simulaciones es un vector  $\tilde{z}_{t+1}^j$  de la forma

$$\tilde{z}_{t+1}^j = (z_{(t+1),1}^j, z_{(t+1),2}^j, z_{(t+1),3}^j, z_{(t+1),4}^j)$$

Luego, de acuerdo con los coeficientes estimados de la series de tiempo para cada serie de rendimientos al tiempo  $t + 1$ , obtenemos una simulación en términos de los rendimientos de los activos dados por,

$$\tilde{r}_{t+1}^j = \hat{\mu}_t + \tilde{z}_t^j * \hat{\sigma}_t \quad (2.19)$$

A partir de la ecuación del capital al tiempo  $t + 1$ , podemos calcular los posibles escenarios para el capital con base en los datos simulados.

$$\tilde{x}_{t+1}^j = x_t \cdot \sum_{i=1}^d \tilde{r}_{t+1}^i \cdot \omega_i,$$

donde  $x_t$  es el capital observado al tiempo  $t$ .

Por lo tanto, a partir de la definición 2.1.2 se calculan las medidas del riesgo del portafolio VaR y CVaR considerando la distribución empírica de los datos

$$\{x_{t+1}^1, \dots, x_{t+1}^{nSim}\}$$

Este procedimiento completo de estimación y simulación se repite en cada tiempo  $t$  dentro del periodo en que se desea obtener las medidas de riesgo. En la práctica, esto representa un problema de alta complejidad computacional.

# Capítulo 3

## Implementación del modelo

En este capítulo se presentan los resultados de implementar el modelo estudiado en el capítulo anterior para estimar la distribución conjunta de los rendimientos de un portafolio de divisas, con el fin de calcular el VaR como una medida de riesgo. Se va a considerar el portafolio que da el mismo peso a cada uno de los activos en cada tiempo y en este caso, está formado por las divisas euro (EUR), dólar estadounidense (USD), libra esterlina (GBP) y yen japonés (JPY); dichas monedas juegan un rol muy importante en el mercado de divisas.

Se consideran datos de precios diarios dentro del periodo del 03 de Enero de 2000 al 17 de enero de 2019, las gráficas de la evolución de los precios de las monedas de interés se observan en la siguiente gráfica.

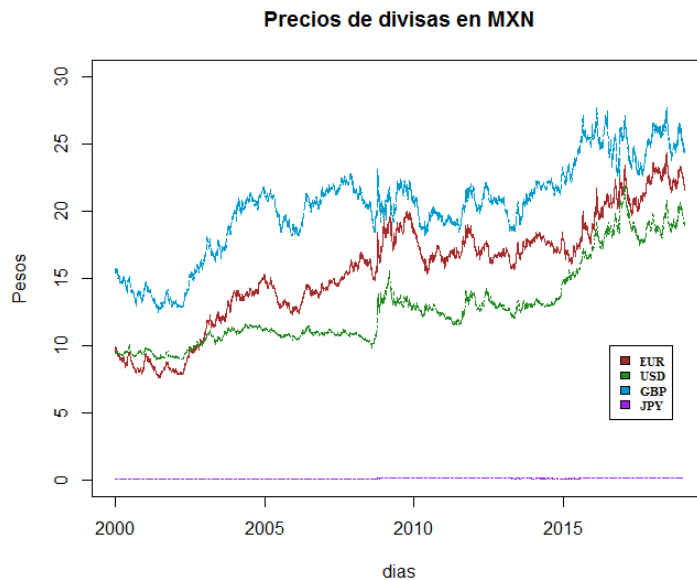


Figura 3.1: *Datos históricos de precios diarios de divisas respecto al peso mexicano.*

A partir de estas series de precios se obtienen las series de rendimientos para cada divisa de acuerdo con la ecuación (2.2). Vamos a considerar un capital inicial  $x_0 = 1$  y para cada tiempo el portafolio estará dado por  $\omega = (0.25, 0.25, 0.25, 0.25)$ ; recordemos que el capital al tiempo  $t$  está dado por

$$X_t = X_{t-1} \cdot \sum_{i=1}^d r_t^i, \omega_i = X_{t-1} \cdot \langle r_t, \omega \rangle.$$

Estamos interesados en calcular el VaR del capital  $X_t$ , para ello necesitamos estimar la distribución conjunta de los rendimientos del portafolio, como discutimos en el capítulo anterior; para la estimación de dicha distribución en cada tiempo se consideran para la estimación 1000 rendimientos históricos.

Vamos a mostrar en este capítulo la estimación a detalle realizada para el cálculo del VaR del capital para un periodo de tiempo, y en la última sección mostramos los resultados de repetir este procedimiento en una ventana de tiempo más grande. Para esto último, se necesita repetir la estimación que vamos a comentar aquí en cada tiempo dentro del rango que se evalúa, lo cual implica un tiempo computacional alto. Considerando ésto, vamos a presentar el desarrollo del modelo usando los 1000 rendimientos diarios anteriores al 01 de Enero de 2008, los rendimientos de las divisas observados para este periodo se presentan en la siguiente gráfica.

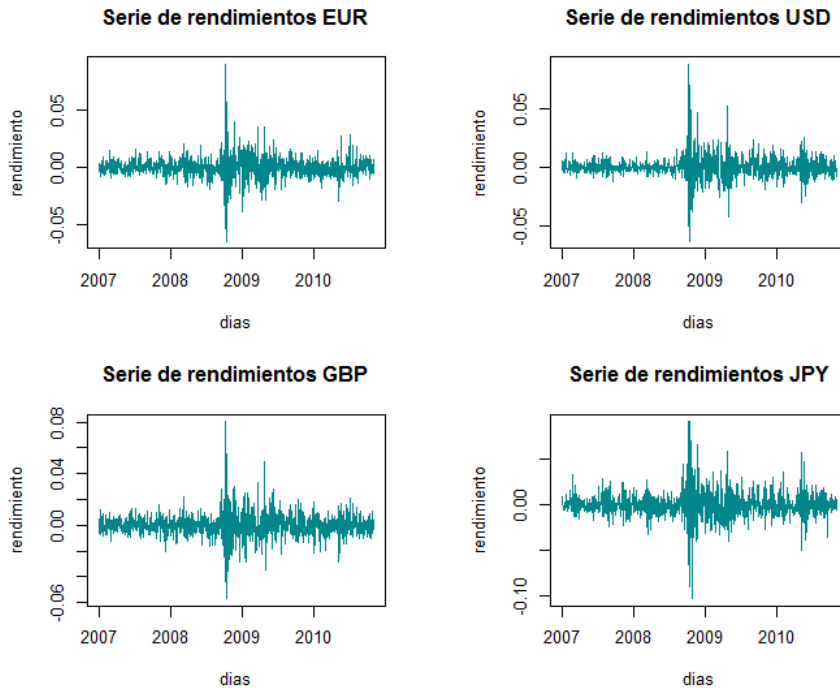


Figura 3.2: Rendimientos diarios en pesos mexicanos.



### 3.1. ESTIMACIÓN DE PARÁMETROS DEL MODELO DE SERIES DE TIEMPO

De acuerdo con el modelo propuesto y considerando las series de rendimientos, primero se aplica un modelo  $ARMA-GJR-GARCH$  para cada una de ellas; la estimación de los parámetros de este modelo aparece en la primera sección de este capítulo. Luego, con base en el modelo estimado se calculan los residuos estandarizados en cada caso y lo siguiente es ajustar la distribución conjunta de dichos residuales, aprovechando la relación entre la distribución de rendimientos y la distribución de residuos estandarizados que se comentó en la ecuación (2.19).

La distribución conjunta de los residuos estandarizados, se obtendrá a partir de la construcción discutida en la Sección 2.4, para la cual, se estiman las distribuciones marginales en la Sección 3.2; para esto se presentan dos ajustes importantes: el primero es estimar los parámetros de distribuciones  $t$ -Student y el segundo es estimar los valores extremos con distribuciones Pareto Generalizada. Luego se analiza la dependencia de los residuos mediante el ajuste de funciones cópula en la Sección 3.3.

Finalmente en la Sección 3.4, se presenta el cálculo del VaR de nuestro portafolio evaluado en algunos periodos de tiempo, en los distintos casos de estimación que consideramos.

Los rendimientos considerados antes presentan la siguiente matriz de correlación, y observando estos valores tenemos que es adecuado incluir en el modelo la dependencia de estas divisas, pues presentan una correlación alta.

	<b>EUR</b>	<b>USD</b>	<b>GBP</b>	<b>JPY</b>
<b>EUR</b>	1.0000000	0.8838746	0.9207660	0.8947897
<b>USD</b>	0.8838746	1.0000000	0.8411800	0.8591616
<b>GBP</b>	0.9207660	0.8411800	1.0000000	0.7421641
<b>JPY</b>	0.8947897	0.8591616	0.7421641	1.0000000

### 3.1. Estimación de parámetros del modelo $ARMA-GJR-GARCH$ para los rendimientos

A cada serie de residuos estandarizados correspondiente a cada divisa, se le ajusta un modelo de series de tiempo que presentamos en la Sección 2.2; adelante presentamos para diferentes selecciones de órdenes del modelo los valores del Criterio de Información de Akaike (AIC), con el fin de seleccionar un modelo adecuado.

Primero se estableció un rango de valores para los parámetros  $m, n$  correspondientes al modelo  $ARMA$  y  $p, q$  correspondientes al modelo de la volatilidad; el rango de valores para cada uno fue  $\{0, 1, 2, 3\}$  y se tomaron en cuenta todas las combinaciones de los parámetros. Para cada combinación de valores  $m, n, p, q$  se estimaron los coeficientes del modelo por máxima verosimilitud, y con ello se calculó el Criterio de Información de Akaike para seleccionar cuál de estas combinaciones de estos parámetros era la más adecuada. A continuación mostramos algunos de éstos resultados.

En el caso de la serie de rendimientos del Euro, tenemos lo siguiente:

m	n	Verosimilitud	AIC	p_opt	q_opt
0	0	17412.59	<b>-6.96203</b>	2	3
1	0	17339.69	-6.933662	2	3
0	1	17339.69	-6.933663	2	3
1	1	17339.97	-6.933299	2	3
1	2	17339.97	-6.932960	2	3
1	3	17340.02	-6.932525	2	3
3	1	17340.01	-6.932519	2	3
3	3	17354.54	-6.937604	2	3
3	4	17355.61	-6.937632	2	3

Para este activo, notamos que considerando los valores de AIC, podemos elegir un modelo  $ARMA(0,0)$  como los valores de AIC son cercanos, vamos a preferir el modelo más sencillo para evitar la sobre parametrización del modelo, en éste caso seleccionamos un modelo  $ARMA(0,0) - GJR - GARCH(2,3)$ , los coeficientes de dicho modelo son estimados por máxima verosimilitud haciendo uso del paquete *rugarch* de R.

Considerando también que el paquete *rugarch* tiene opciones en la distribución para el cálculo de la verosimilitud, considerando un modelo GJR-GARCH(2,3) y al considerar algunas de estas opciones tenemos que,

m	n	Verosimilitud	AIC (norm)	AIC (std)	AIC (sged)
0	0	17412.59	-6.933552	-6.96474	-6.96203
1	0	17339.69	-6.933662	-6.964475	-6.962033
0	1	17339.69	-6.933663	-6.964476	-6.962037
1	1	17339.97	-6.933299	-6.964086	-6.96166

Cuadro 3.1: Comparación del ajuste del modelo GJR-GARCH eligiendo diferentes distribuciones para el cálculo de la verosimilitud.

Tenemos que no se observa una variación significativa en el criterio AIC, es por ello que a continuación se muestran los resultados del criterio AICC introducido por Brockwell & Davis (2006) como una modificación al criterio AIC, el cual está dado por

$$AIC(\beta) = -2L_X(\beta)/nObs + 2(nPar)/nObs$$

donde  $nObs$  es el número de observaciones,  $nPar$  es el número de observaciones y  $L_X(\beta)$  la verosimilitud del modelo.

La estadística AICC es adecuada para los modelos de la volatilidad condicional, esta estadística está dada por,

$$AICC(\beta) = \frac{2(p+q+1)nObs}{nObs-p-q-2} - 2L_X(\beta)/nObs$$

### 3.1. ESTIMACIÓN DE PARÁMETROS DEL MODELO DE SERIES DE TIEMPO

De manera análoga al criterio AIC, un modelo elegido de acuerdo con la estadística AICC es tal que sus parámetros minimizan dicha estadística. En la siguiente tabla se muestra una comparación de dicha estadística en diferentes modelos para la volatilidad condicional.

<b>p</b>	<b>q</b>	<b>AIC (norm)</b>	<b>AICC (norm)</b>	<b>AICC (std)</b>	<b>AICC (sged)</b>
1	1	-6.932717	6.0009	6.000898	6.000898
1	2	-6.932636	8.004105	8.004103	8.004103
2	2	-6.933243	10.00811	10.00811	10.00811
2	3	-6.933946	12.01292	12.01292	12.01292

Cuadro 3.2: Comparación de las estadísticas AICC en el modelo GJR-GARCH( $p, q$ ) eligiendo diferentes valores de  $p$  y  $q$ , así como diferentes distribuciones para el cálculo de la verosimilitud.

Tenemos que sí hay una variación en la estadística cuando se cambian los valores de los parámetros  $p$  y  $q$ , sin embargo, para valores iguales de estos parámetros, la estadística no varía de manera significativa al cambiar el tipo de la distribución.

Procediendo de manera análoga, en el caso de la serie de retornos del dólar USD tenemos que:

<b>m</b>	<b>n</b>	<b>Verosimilitud</b>	<b>AIC</b>	<b>p_opt</b>	<b>q_opt</b>
0	0	18491.93	-7.394652	2	3
1	0	18500.85	<b>-7.397818</b>	2	3
0	1	18500.76	-7.397783	2	3
1	1	18491.93	-7.394652	2	3
1	2	18503.37	<b>-7.398318</b>	2	3
1	3	18501.52	-7.397200	2	3
3	1	18503.17	-7.397946	2	3
3	3	18504.53	-7.397690	2	3
3	4	18504.73	-7.397289	2	3

Seleccionamos el modelo  $ARMA(1,0) - GJR - GARCH(2,3)$  para la serie de rendimientos correspondiente al dólar estadounidense, cuyo valor de AIC es más chico. Luego considerando fijos  $m$  y  $n$  se varían  $p, q$  en las posibles combinaciones y se eligen los valores  $p_{opt}$  y  $q_{opt}$  tales que hacen que el criterio AIC sea menor.

En otro caso, para la serie de retornos de la libra GBP, obtenemos:

<b>m</b>	<b>n</b>	<b>Verosimilitud</b>	<b>AIC</b>	<b>p_opt</b>	<b>q_opt</b>
0	0	17553.29	-7.019118	2	3
1	0	17556.12	-7.019853	2	3
0	1	17556.2	<b>-7.019885</b>	2	3
1	1	17556.54	-7.019618	2	3
1	2	17560.18	-7.020274	3	3
1	3	17560.20	-7.019885	3	3
3	1	17560.21	-7.019887	3	3
3	3	17564.75	<b>-7.020905</b>	3	3
3	4	17563.16	-7.020002	1	1

De manera análoga al primer caso elegimos el modelo  $ARMA(0, 1) - GJR - GARCH(2, 3)$  por ser uno cuyo valor de AIC es menor y por ser el modelo más sencillo.

Por último, para la elección del modelo de los retornos del yen (JPY) se tiene

<b>m</b>	<b>n</b>	<b>Verosimilitud</b>	<b>AIC</b>	<b>p_opt</b>	<b>q_opt</b>
0	0	16556.37	-6.620273	2	3
1	0	16557.29	-6.620239	2	3
0	1	16557.28	-6.620238	2	3
1	1	16557.88	<b>-6.620510</b>	1	3
1	2	16557.91	-6.620113	1	3
1	3	16558.23	-6.619853	1	3
3	1	16558.26	-6.619862	1	3
3	3	16561.32	-6.620359	1	3
3	4	16561.32	-6.619854	1	3
4	4	16564.96	-6.620108	3	3
4	5	16578.29	<b>-6.625593</b>	2	3

En este caso es fácil notar que, en general, mientras más parámetros consideramos, el valor de AIC disminuye, sin embargo no es adecuado un modelo con muchos parámetros pues puede haber un mayor error en la estimación de los coeficientes del modelo. Por esto, elegimos el modelo  $ARMA(1, 1) - GJR - GARCH(1, 3)$  para la serie de rendimientos del yen japonés.

En resumen, con base en el criterio de información de Akaike y con objeto de no sobreparametrizar los modelos, los parámetros que vamos a considerar para los modelos de los rendimientos de cada tasa de cambio, los presentamos en el siguiente cuadro.

### 3.1. ESTIMACIÓN DE PARÁMETROS DEL MODELO DE SERIES DE TIEMPO

---

Cuadro 3.3: Parámetros del modelo ARMA(m,n)-GJR-GARCH(p,q)

<i>Divisa\Parametros</i>	<b>m</b>	<b>n</b>	<b>p</b>	<b>q</b>
<b>EUR</b>	0	0	2	3
<b>USD</b>	1	0	2	3
<b>GBP</b>	0	1	2	3
<b>JPY</b>	1	1	1	3

Después de elegir valores fijos de los parámetros  $m, n, p, q$ , luego los coeficientes del modelo ARMA(m,n)-GJR-GARCH(p,q) fueron estimados por máxima verosimilitud con la función *ugarchfit* del paquete *rugarch* disponible en la paquetería de R, estos coeficientes son  $\hat{\phi}_1, \dots, \hat{\phi}_m, \hat{\theta}_1, \dots, \hat{\theta}_n, \hat{\alpha}_1, \dots, \hat{\alpha}_p, \hat{\gamma}_1, \dots, \hat{\gamma}_p, \hat{\beta}_1, \dots, \hat{\beta}_q$ ; de acuerdo con las ecuaciones (2.3)– (2.6).

### 3.1.1. Residuos del modelo GJR-GARCH

Después de ajustar un modelo ARMA(m,n)-GJR-GARCH(p,q) para cada serie de rendimientos de las divisas, obtenemos los residuos estandarizados correspondientes dados en la ecuación (2.9). Como mencionamos en el capítulo anterior, las series de rendimientos presentan mayor correlación que las series de residuos estandarizados resultantes, lo cual podemos observar en el caso de los datos correspondientes al dólar estadounidense, en la siguiente gráfica mostramos los correlogramas correspondientes.

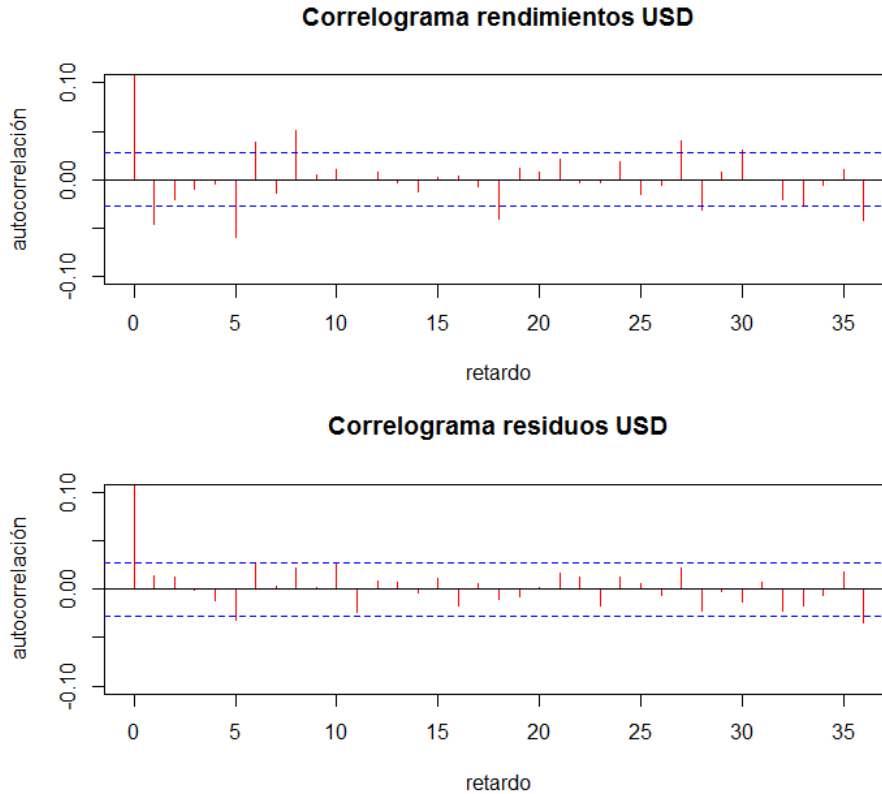


Figura 3.3: Comparación de la correlación de los rendimientos con la de los residuos estandarizados en el caso de la serie para el dólar.

De acuerdo con estos correlogramas, aunque ambos presentan correlaciones bajas notamos que efectivamente los residuos estandarizados presentan menos correlación que los rendimientos originales de la serie. Pues en este caso los estadísticos de la prueba Durbin-Watson (Sección 1.4) no rebasan las bandas del 95 % de confianza de esta prueba, de aquí se interpreta que no hay evidencia en contra de la hipótesis de que los residuos estandarizados son no correlacionados.

Por lo tanto, es razonable estimar la distribución de los residuos, considerando éstos como observaciones independientes y luego con base en dicha distribución obtener la de los rendimientos, como se discutió en el capítulo anterior.

## 3.2. AJUSTE DE DISTRIBUCIONES MARGINALES

---

Para realizar el ajuste de las distribuciones marginales de los residuos estandarizados para cada divisa se presentan a continuación los histogramas de estos residuos en cada caso.

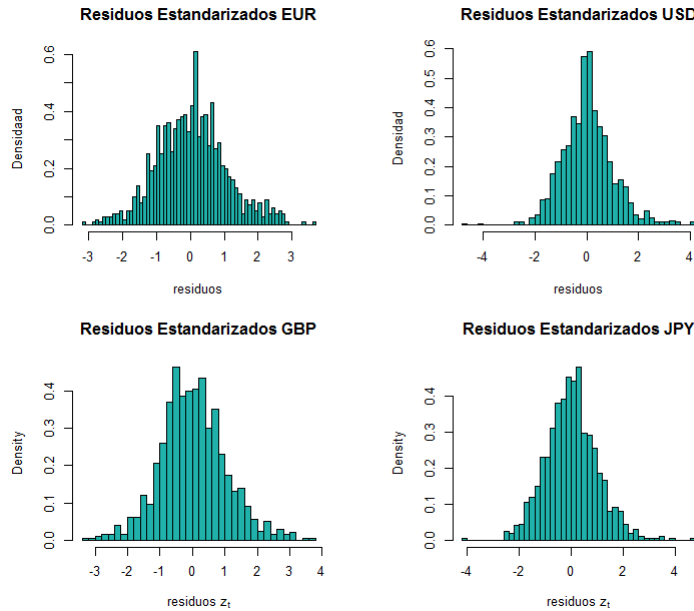


Figura 3.4: *Residuos estandarizados calculados a partir del ajuste de un modelo  $ARMA(m,n)$ -GJR-GARCH( $p,q$ ) para cada una de las divisas seleccionadas.*

Se puede observar en los histogramas que, en general, la cola derecha de dichas distribuciones es más pesada que la izquierda, lo cual es razonable dado que el modelo propuesto de series de tiempo penaliza más los errores negativos que positivos.

## 3.2. Ajuste de distribuciones Marginales

Para la obtención de la distribución conjunta de los rendimientos del portafolio, primero vamos a estimar las distribuciones marginales para cada una de las series de residuos estandarizados del modelo anterior.

Como un ajuste preliminar, hemos ajustado una distribución  $t$ -Student que presentamos en la siguiente sección. Además, también se realiza la estimación de la densidad aplicando teoría de valores extremos, para la cual explicamos también el procedimiento para la elección de umbrales adecuados y el ajuste de distribuciones Pareto Generalizadas en las colas de la distribución.

### 3.2.1. Ajuste Distribución $t$ -Student

Como un primer ajuste para las distribuciones marginales, proponemos ajustar una distribución  $t$ -Student para cada serie de residuos estandarizados. En cada caso, estimamos los grados de libertad mediante el método de máxima verosimilitud.

En el caso de cada serie de residuos se consideran observaciones  $(z_1, z_2, \dots, z_n)$  independientes de los residuos estandarizados; luego la función de densidad de una variable aleatoria  $Z$  con distribución  $t$ -student con  $\nu > 0$  grados de libertad, está dada por,

$$f(z) = \frac{\Gamma((\nu + 1)/2)}{\sqrt{\nu\pi} \Gamma(\nu/2)} (1 + z^2/\nu)^{-(\nu+1)/2}.$$

Para cada serie de residuos se calculó el parámetro de los grados de libertad para ajustar una  $t$ -Student adecuada; se obtuvieron las siguientes estimaciones de dichos parámetros en cada caso.

Divisa	$\hat{\nu}$
EUR	7.8
USD	7.95
GBP	8.69
JPY	7.45

Cuadro 3.4: *Estimación de parámetros para distribuciones marginales  $t$ -Student.*

Las densidades  $t$ -Student estimadas se muestran en las siguientes gráficas.

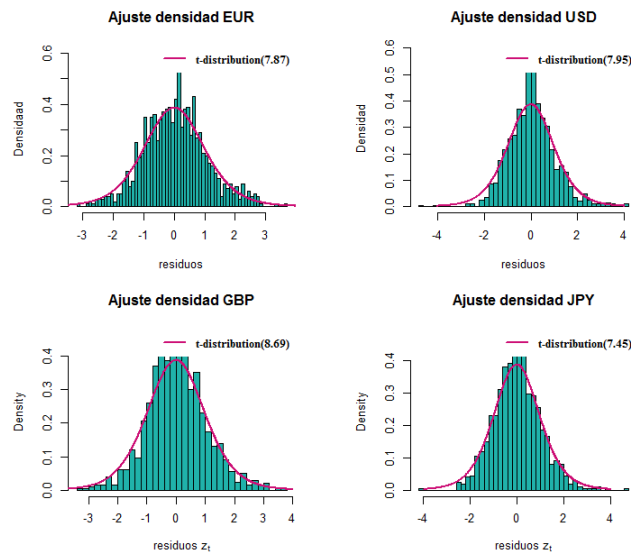


Figura 3.5: *Ajuste de densidades  $t$ -Student para cada serie de residuos estandarizados.*



Aunque el ajuste de la densidad  $t$ -Student parece razonable, es de gran importancia obtener un buen ajuste en las colas derecha e izquierda de la densidad de residuos pues estos representan errores positivos o negativos del modelo de serie de tiempo, lo que se traduce en pérdidas o ganancias de las monedas que conforman el portafolio.

Por otra parte, observamos que la densidad de residuos estandarizados no es simétrica, como muestran los histogramas del euro y libra esterlina. Observamos que la cola derecha parece un poco más pesada, aquí podemos ver el efecto de que el modelo  $GJR - GARCH$  pone mayor peso en los errores negativos del modelo.

En la última sección de este capítulo, sobre el cálculo de VaR en un horizonte de tiempo, se presentan resultados con ambos ajustes de las marginales, el ajuste de las marginales aplicando valores extremos lo explicamos en la siguiente sección.

### 3.2.2. Ajuste Distribución Marginal aplicando EVT

Para tratar de hacer un mejor ajuste en las colas de la distribución, modelaremos éstas con distribuciones pertenecientes a la familia Pareto Generalizada. Como discutimos en el capítulo anterior, antes de hacer la estimación de dichas distribuciones es necesario elegir umbrales tales que determinan los valores extremos de la densidad para las colas derecha e izquierda.

En cada caso aplicamos los métodos de selección de umbral presentados en la Sección 2.3.2, ambos son herramientas exploratorias que proporcionan elementos para decidir qué valor del umbral parece más adecuado. El primero de estos métodos trata de encontrar el valor más grande para el cual la función media de excesos tiene una tendencia lineal. Para ello, se grafica la función media de excesos empírica, dada en la ecuación (2.13).

Aquí hemos aplicado la función `mrlplot` del paquete `evmix` para obtener estas gráficas. Presentamos a continuación el resultado para el umbral derecho e izquierdo en el caso de la serie del euro.

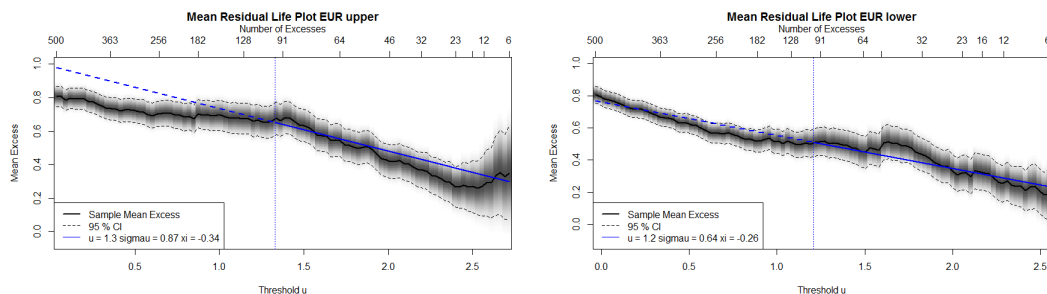


Figura 3.6: Selección de umbrales basada en la linealidad de la función media de excesos empírica para las colas derecha e izquierda de la serie de residuos del euro.

En estas gráficas la línea punteada muestra el mayor valor para el que es razonable considerar una tendencia lineal en la función media de excesos empírica.

Observemos que para el umbral de la cola izquierda consideramos el negativo de los residuos estandarizados pues la función *mrlplot* de R proporciona los resultados para la cola derecha de la distribución. En ambas gráficas se muestra también la estimación de los parámetros de la distribución Pareto Generalizada estimada y el número de excedencias en el eje superior.

### Selección de umbral con base en la estabilidad del parámetro de forma de la distribución Pareto Generalizada

Para complementar el método anterior, y dado que es más común la selección de umbral con base en la estabilidad del parámetro de forma de la distribución Pareto Generalizada, aquí aplicamos el segundo método discutido en la Sección 2.3.2. En el caso del euro obtenemos las siguientes gráficas.

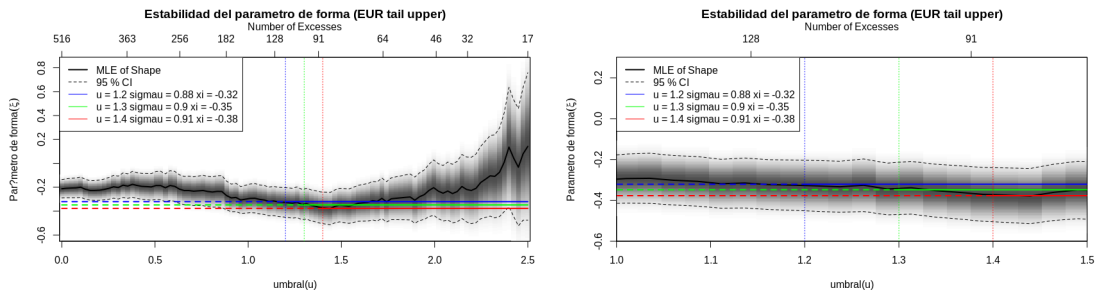
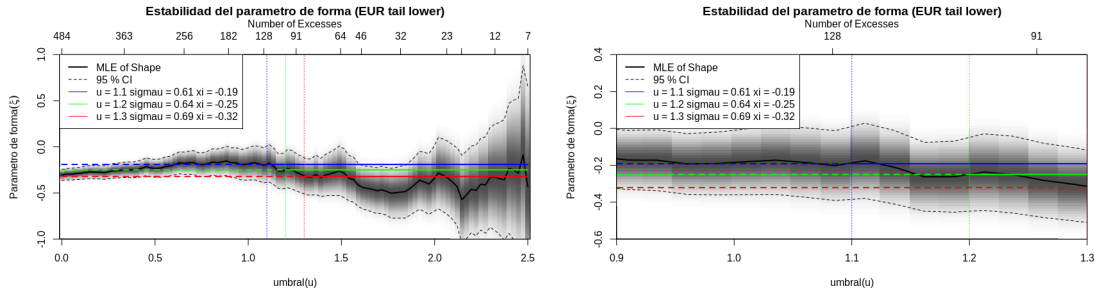


Figura 3.7: Selección de umbral basada en la estabilidad del parámetro de forma, en la gráfica de la derecha se observa con detalle el intervalo donde fue elegido el parámetro de forma de la distribución Pareto Generalizada para la cola derecha de la serie de residuos del euro.

Por lo tanto, en el caso del umbral para la cola derecha, se eligió  $u^U = 1.3$  y los estimadores de la distribución Pareto Generalizada, con base en máxima verosimilitud son  $\sigma = 0.9$  y el parámetro de forma  $\xi = -0.35$ .

De manera análoga para la seleccionar el umbral  $u^L$  que determina la cola izquierda de la distribución de residuos estandarizados para el euro, consideramos la estabilidad del parámetro de forma, en este caso tenemos las gráficas de su relación con el umbral.

### 3.2. AJUSTE DE DISTRIBUCIONES MARGINALES



En este caso hemos elegido el umbral  $u^L = -1.1$ , pues notamos que en este punto se tiene estabilidad en el parámetros de forma estimado. Además que en dichas gráficas se comparan algunos umbrales, notamos que los valores 1.2 y 1.3 presentan una estabilidad muy parecida respecto al parámetro de forma, sin embargo, para la elección del umbral también consideramos el número de excedencias para dicho umbral.

Por lo tanto, de acuerdo con estos métodos de selección de umbral hemos elegido umbrales  $u^L$  y  $u^R$  para el caso de la serie de residuos del euro; los parámetros para cada distribución Pareto Generalizada se estimaron por máxima verosimilitud, el ajuste con las densidades estimadas se muestra en la siguiente gráfica. Además se muestra la comparación con el ajuste de la densidad  $t$ -Student anterior.

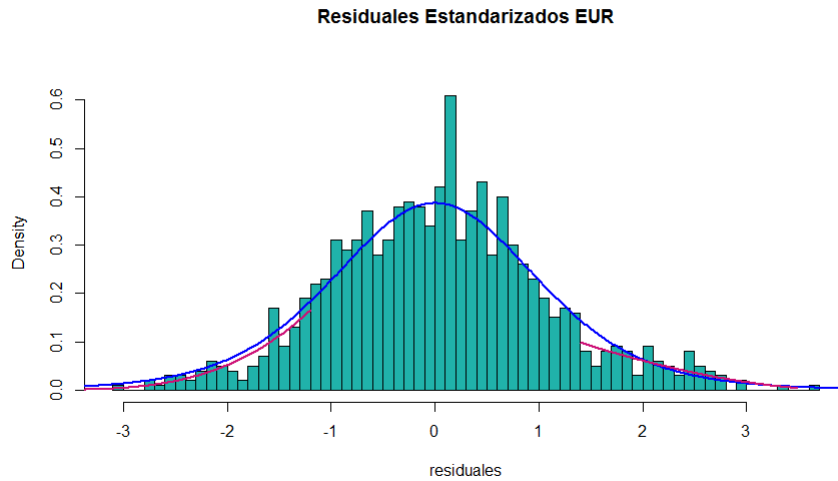


Figura 3.8: Comparación del ajuste de la densidad  $t$ -Student y la estimación en las colas con la distribución Pareto Generalizada para la serie de residuos del euro.

Observamos que en la cola derecha de la densidad hay un mejor ajuste con la distribución Pareto Generalizada que con la distribución  $t$ -Student. Además estimando

las colas de la densidad con dos distribuciones de diferentes parámetros se captura el efecto de la volatilidad asimétrica.

Al tener el ajuste en las colas con distribuciones Pareto Generalizada, el interior se estima mediante una estimación de densidad por kernel Gaussiano, de esta manera, la densidad que ajustamos se muestra en la siguiente gráfica.

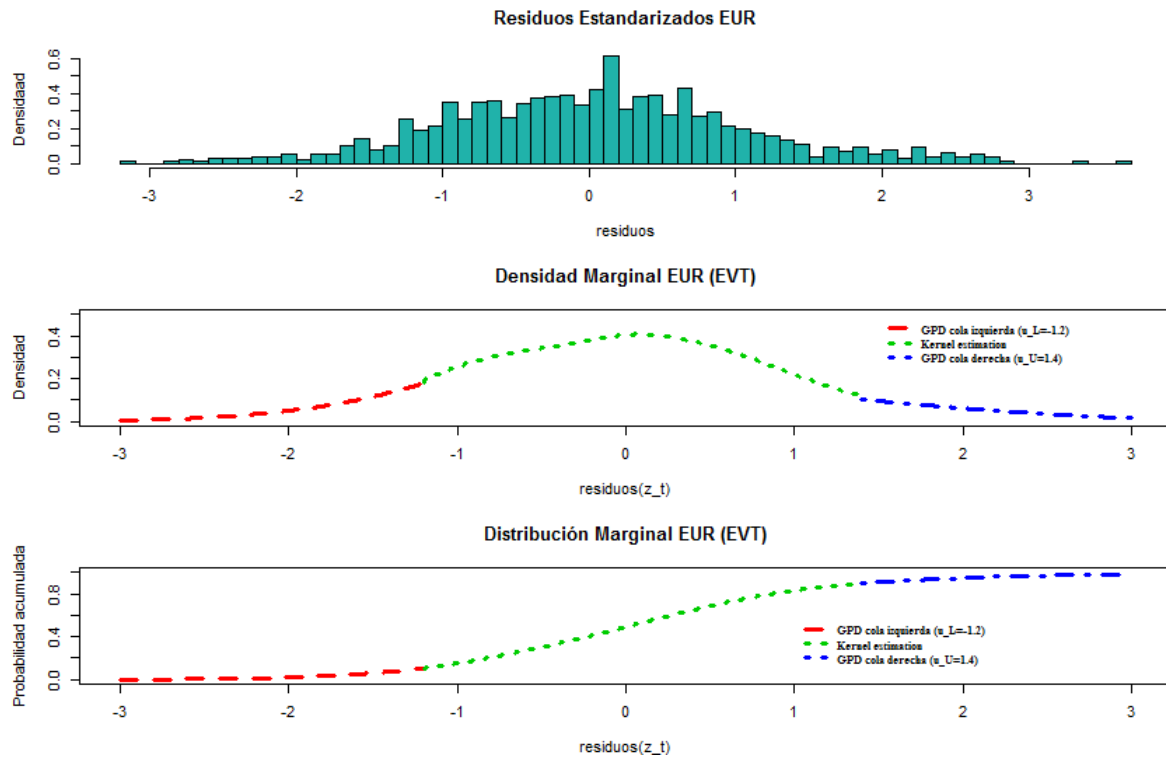


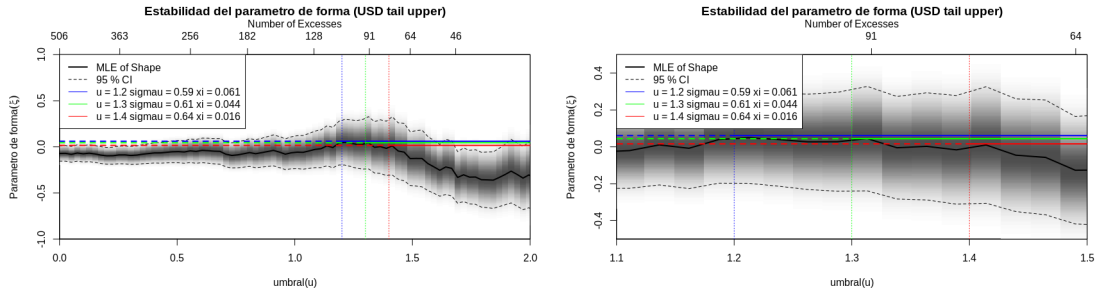
Figura 3.9: *En la primera gráfica se presenta el histograma de residuos estandarizados de la serie del euro con mil datos históricos; en la segunda gráfica se muestra la densidad estimada con colas de densidades Pareto Generalizada y el interior mediante kernel Gaussiano; en la tercer gráfica la función de distribución correspondiente a la densidad anterior, dada en la ecuación (2.4).*

Vamos a proceder de manera análoga con las series de residuos estandarizados de las otras divisas, para obtener sus distribuciones marginales correspondientes.

## Distribución marginal residuos estandarizados USD

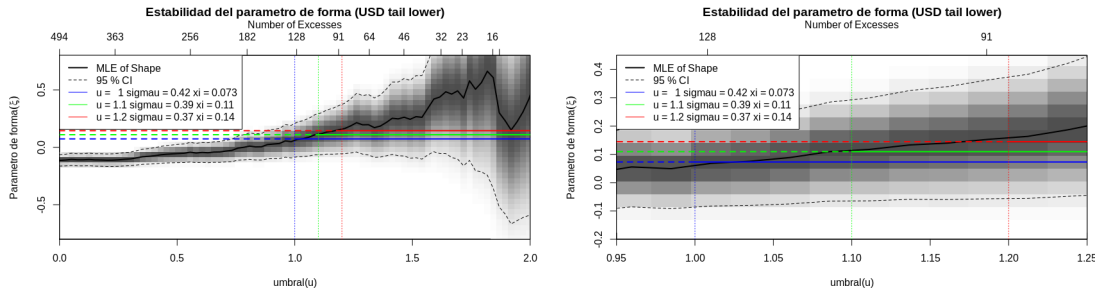
En el caso de la serie de residuos estandarizados del dólar, ajustamos una distribución  $t$ -Student, el estimador máximo verosímil para los grados de libertad es  $\hat{\nu} = 7.8$ .

Para el ajuste de las distribuciones Pareto Generalizada, seleccionamos umbrales para la cola derecha e izquierda de la distribución, con base en las gráficas de la estabilidad del parámetro de forma al variar el umbral en un rango de posibles valores. Para la cola derecha tenemos las gráficas,



En este caso, seleccionamos el umbral  $u^U = 1.3$ , pues notamos que para éste valor la relación entre el umbral y el parámetro de forma es constante y el número de excedencias parece adecuado para la estimación de parámetros.

Luego, en el caso de la elección del umbral  $u^L$ , las gráficas para observar la estabilidad del parámetro de forma para la cola izquierda son las siguientes,



Aquí seleccionamos el umbral  $u^L = -1.1$  con el mismo criterio que en el caso anterior. Al considerar estos umbrales obtenemos una estimación de los parámetros de las distribuciones Pareto Generalizada. Con base en éstos y realizando una estimación de densidad por kernel para los valores que no exceden ninguno de los umbrales derecho o izquierdo, tenemos que la estimación de la densidad completa es,

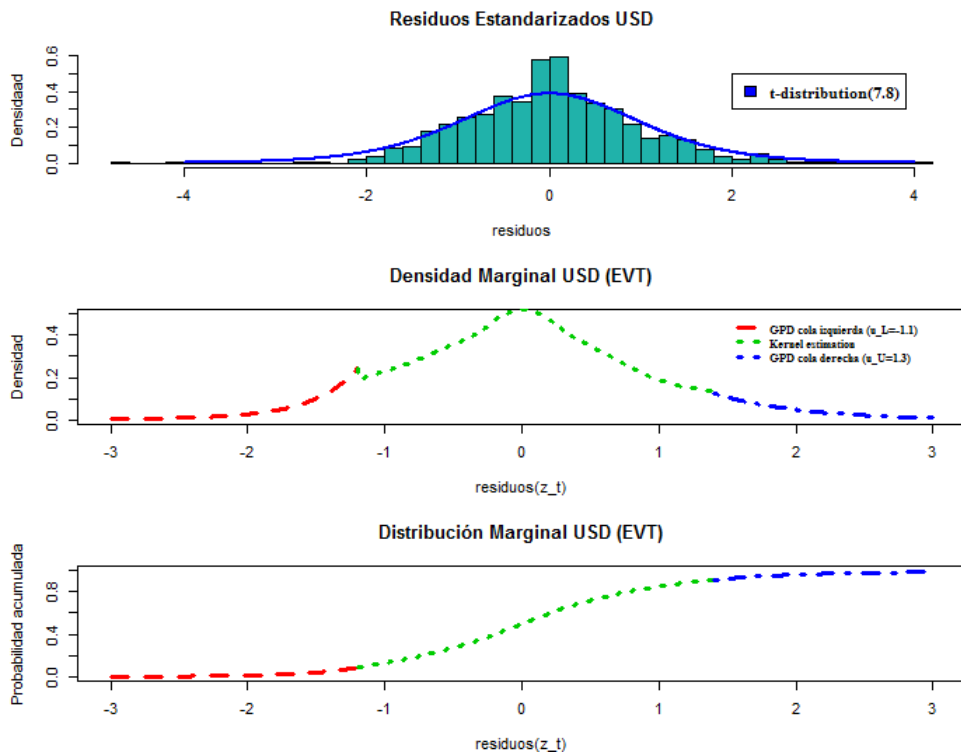


Figura 3.10: En la primera gráfica se presenta el histograma de residuos estandarizados de la serie del dólar para mil datos históricos; en la segunda gráfica se muestra la densidad estimada con colas de densidades Pareto Generalizada; en la tercer gráfica la función de distribución correspondiente a la densidad anterior, dada en la ecuación (2.4).

### Distribución marginal residuos estandarizados GBP

En el caso de la serie de residuos estandarizados para la libra esterlina (GBP) y procediendo de manera análoga a lo anterior, obtuvimos que el estimador máximo verosímil de la distribución  $t$ -Student es  $\hat{\nu} = 8.69$ . Y los umbrales seleccionados a través de los métodos discutidos antes son  $u^L = -1.1$  y  $u^U = 1.3$ .

### 3.2. AJUSTE DE DISTRIBUCIONES MARGINALES

La estabilidad de los parámetros de forma para las distribuciones Pareto Generalizada para las colas izquierda y derecha se muestra a continuación. Para la selección del umbral de la cola derecha para la serie de residuos de la libra esterlina, tenemos,

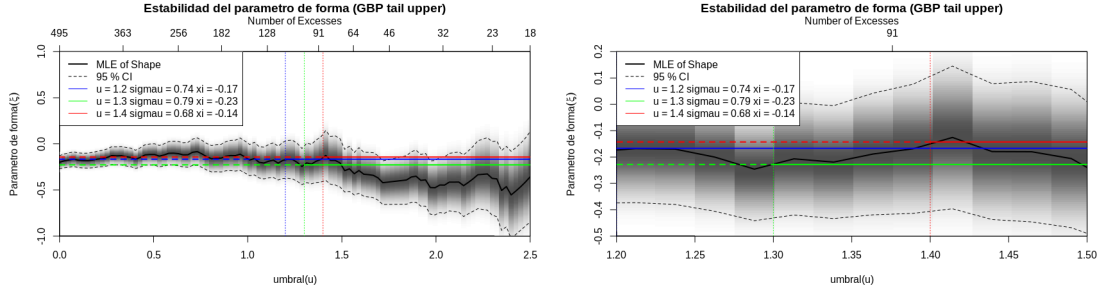


Figura 3.11: Selección de umbral para la distribución de la cola derecha con base en la estabilidad del parámetro de forma de la GPD. En la derecha se puede apreciar mejor la gráfica cerca del umbral seleccionado.

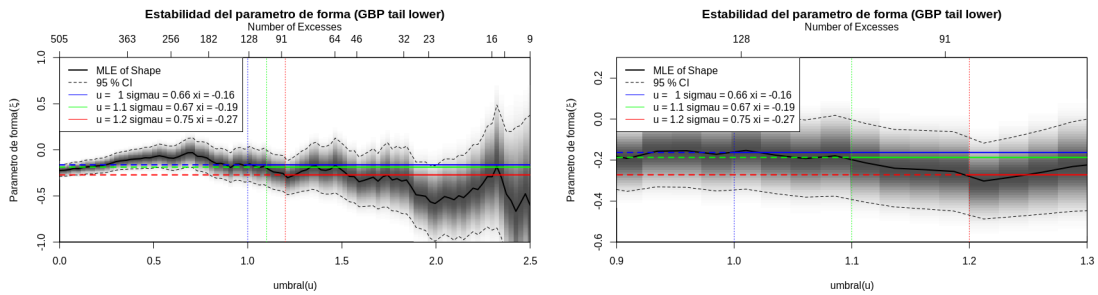


Figura 3.12: Selección de umbral para la distribución de la cola izquierda con base en la estabilidad del parámetro de forma de la GPD.

Observemos que en el caso de los valores del parámetro de forma para los umbrales  $u^L = 1$  y  $u^L = 1.1$  tenemos que la estabilidad que se presenta en estos valores es similar, sin embargo, el número de excedencias cambia considerablemente.

De este modo la densidad y función de distribución correspondiente, tienen la siguiente forma.

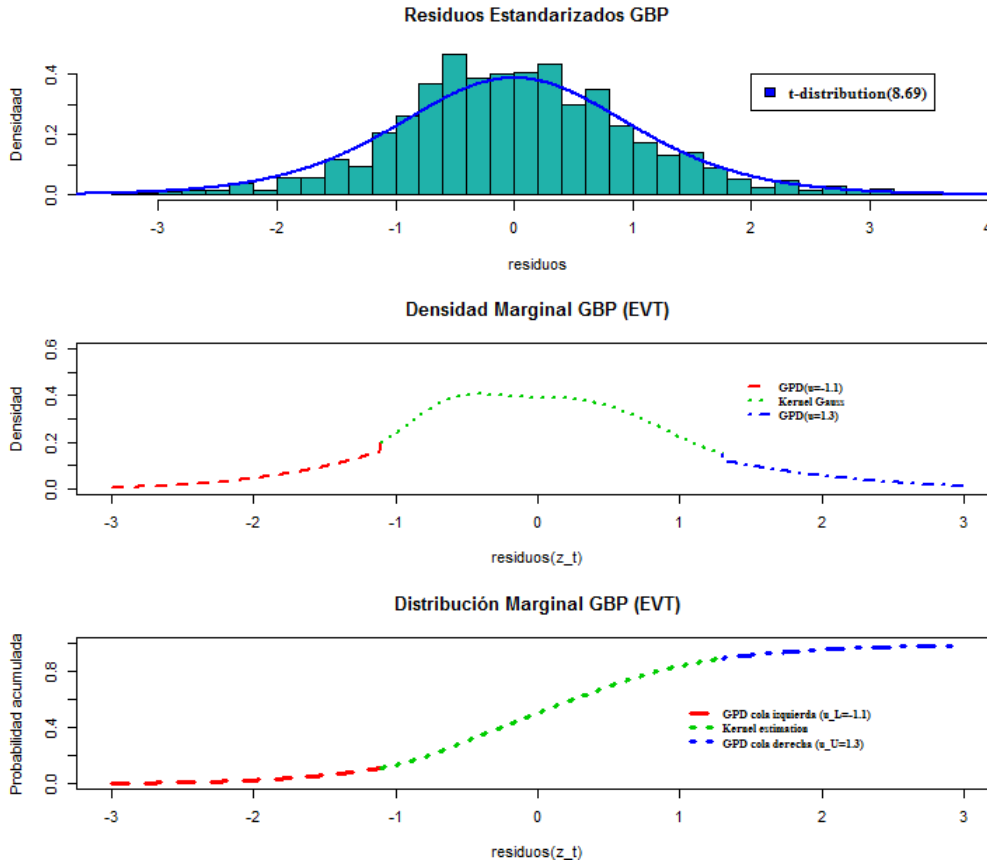


Figura 3.13: En la primera gráfica se presenta el histograma de residuos estandarizados de la serie de la libra para mil datos históricos; en la segunda gráfica se muestra la densidad estimada con colas de densidades Pareto Generalizada; en la tercer gráfica la función de distribución correspondiente a la densidad anterior, dada en la ecuación (2.4).

### Distribución marginal residuos estandarizados JPY

Por último, para el caso de la serie de residuos estandarizados del yen japonés, tenemos que el estimador máximo verosímil del parámetro de la distribución  $t$ -Student es  $\hat{\nu} = 7.45$ . Por otra parte, para la estimación de las colas de la distribución aplicando valores extremos nos basamos en el método de la estabilidad del parámetro de forma de la distribución Pareto Generalizada.

En el caso de la cola derecha de la distribución de residuos del yen japonés, tenemos,



### 3.2. AJUSTE DE DISTRIBUCIONES MARGINALES

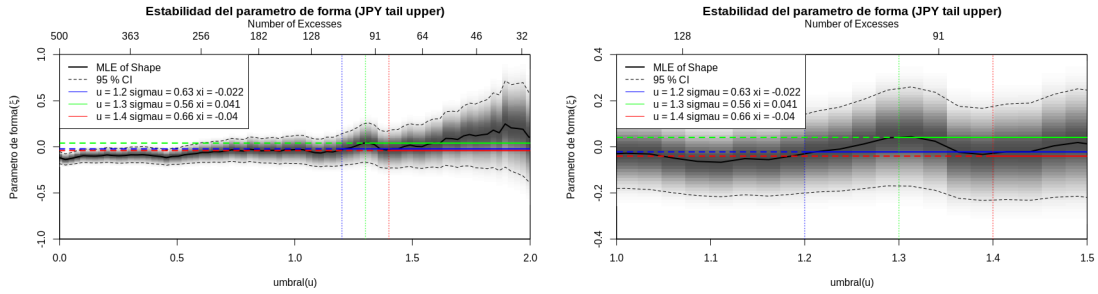


Figura 3.14: Selección de umbral para la cola derecha con base en la estabilidad del parámetro de forma de la GPD.

De manera análoga para el caso de la cola izquierda, observamos las siguientes gráficas para la selección del umbral.

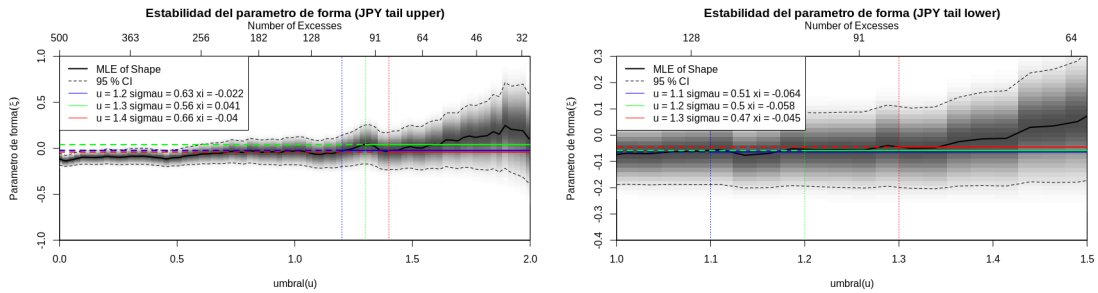


Figura 3.15: Selección de umbral para la cola izquierda con base en la estabilidad del parámetro de forma de la GPD. La gráfica de la derecha muestra un zoom cerca del umbral seleccionado.

Podemos observar que la estabilidad del parámetro de forma es parecida entre los umbrales que elegimos, pero el número de excedencias cambia considerablemente, ésto también fue un criterio de decisión. En este caso, se seleccionaron los umbrales  $u^L = -1.2$  y  $u^U = 1.3$ , con base en esto las funciones de densidad estimadas se muestran a continuación.

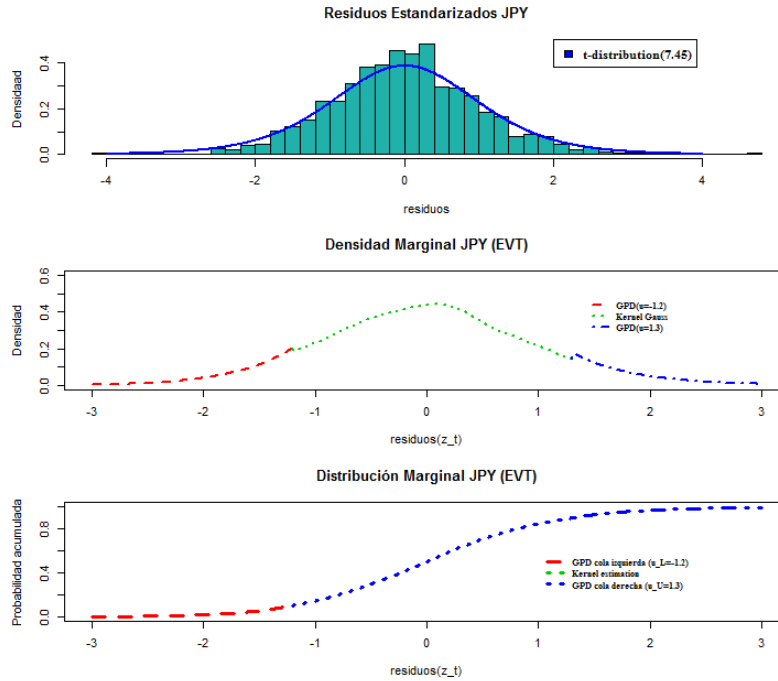


Figura 3.16: En la primera gráfica se presenta el histograma de residuos estandarizados de la serie del yen japonés para mil datos históricos; en la segunda gráfica se muestra la densidad estimada con colas de densidades Pareto Generalizada; en la tercer gráfica la función de distribución correspondiente a la densidad anterior, dada en la ecuación (2.4).

En resumen, se han obtenido dos ajustes para las distribuciones marginales: el primero con distribuciones  $t$ -Student, para el cual los parámetros estimados están dados en el Cuadro 3.2.1. En el segundo ajuste se aplica teoría de valores extremos por umbral, para éste se han elegido umbrales y estimado los parámetros de la distribución Pareto Generalizada en cada caso, dichos datos se concentran en el siguiente cuadro.

Divisa	$u^L$	$\sigma^L$	$\xi^L$	$u^U$	$\sigma^U$	$\xi^U$
EUR	-1.2	0.59	-0.18	1.4	0.87	-0.35
USD	-1.1	0.36	0.15	1.3	0.63	0.01
GBP	-1.1	0.67	-0.18	1.3	0.79	-0.22
JPY	-1.2	0.49	-0.05	1.3	0.55	0.04

Cuadro 3.5: Estimación de parámetros para las distribuciones Pareto Generalizada:  $u^L$  y  $u^U$  representan los umbrales inferior y superior, respectivamente.  $\xi$  representa el parámetro de forma y  $\sigma$  el parámetro de escala de la distribución Pareto Generalizada, usamos superíndices  $U$  y  $L$  para diferenciar los parámetros en los casos de la cola derecha e izquierda en cada caso.

En la evaluación del VaR en la última sección vamos a aplicar ambos ajustes para hacer una comparación de dichas estimaciones. De aquí, tenemos la estimación de las distribuciones marginales de los residuos estandarizados para cada serie de las cuatro divisas.

La estimación que se ha presentado aquí es para el cálculo del VaR para un único periodo de tiempo, si queremos evaluar el VaR por día en una ventana de tiempo mayor es necesario repetir este procedimiento de estimación de parámetros en cada tiempo.

### 3.3. Estimación de funciones cópula

Realizaremos el ajuste de las funciones cópula para describir la dependencia entre las divisas, para ello se van a considerar los dos ajustes presentados para las distribuciones marginales de las divisas.

En el capítulo 1 presentamos algunas de las funciones cópula más utilizadas, para los datos que tenemos de los residuos estandarizados, se calcularon los parámetros de este catálogo de funciones y en todos los casos se obtuvo que el tipo de cópula que mejor los describía era la cópula del tipo  $t$ -Student, el ajuste en este caso se muestra a continuación.

#### 3.3.1. Aplicando ajuste $t$ -Student de las marginales y una función cópula 4-variada

En esta sección se muestran los resultados de estimar los parámetros de una  $t$ -cópula de dimensión cuatro para las series de residuos para cada divisa, así como las estimaciones de las distribuciones  $t$ -Student de la sección anterior. Esto como un primer acercamiento a evaluar el modelo completo para el cálculo del VaR y como una simplificación del mismo pues la estimación de las distribuciones de valores extremos así como la construcción de cópulas por pares se presentan en la siguiente sección.

Para el ajuste de la cópula de dimensión 4, usamos la función  $tCopula$  del paquete *Copula* de R. Se seleccionó este tipo de cópula pues es la que mejor describe los datos observados de los residuos, para esto, se aplicó la función  $fitCopula$  del mismo paquete, la cual hace una evaluación de las funciones cópula más utilizadas y mediante el criterio AIC, se selecciona la más adecuada. Los parámetros estimados de la  $t$ -Copula son:

$$\hat{\nu} = 8.4760618 \quad \hat{\rho} = \begin{pmatrix} 1 & 0.5018728 & 0.6623015 & 0.5960071 \\ 0.5018728 & 1 & 0.7685974 & 0.5690953 \\ 0.6623015 & 0.7685974 & 1 & 0.5690953 \\ 0.5960071 & 0.5690953 & 0.5690953 & 1 \end{pmatrix}$$

Se utilizó la función  $rMvdc$  para simular 5000 muestras de ésta cópula con los parámetros estimados y las funciones  $dMvdc$  y  $pMvdc$  para calcular su probabilidad y

distribución, respectivamente. Estas simulaciones se usan para calcular una estimación del VaR vía Montecarlo; las simulaciones se muestran en la gráfica abajo.

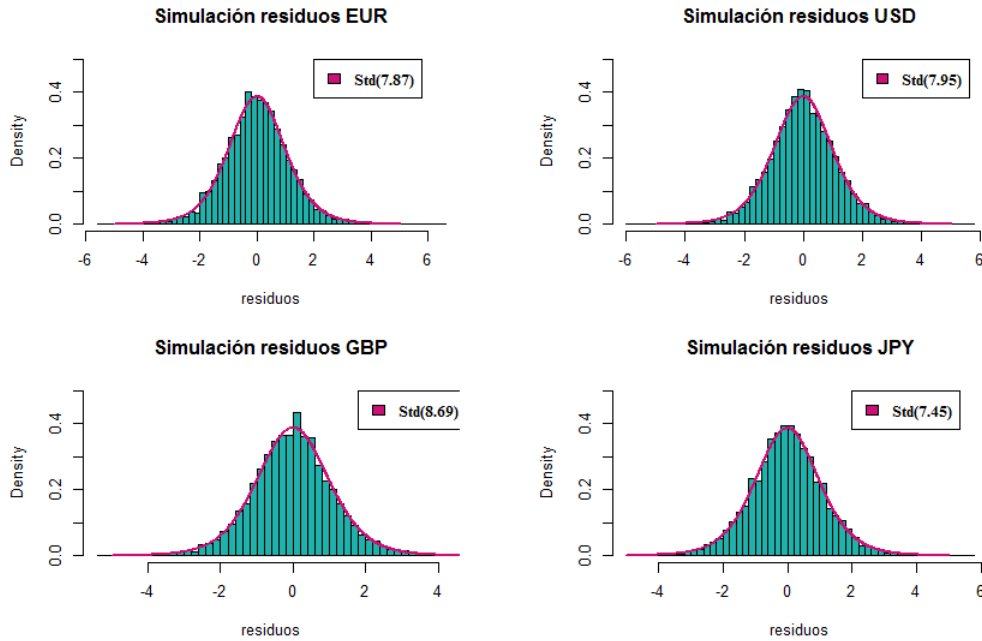


Figura 3.17: Simulación de una muestra de tamaño 5000 con base en la función cópula estimada

De acuerdo con la simulación anterior de la distribución conjunta de los residuos estandarizados y considerando los coeficientes estimados del modelo de la serie de tiempo se tiene una simulación de los rendimientos de los activos que conforman el portafolio. Luego, considerando la estrategia de inversión  $\omega = (0.25, 0.25, 0.25, 0.25)$  y con base en la ecuación del capital  $X_1$  al tiempo 1 de la ecuación (2.1) calculamos el  $VaR(0.1)$  cuantil de nivel  $\alpha = 0.1$  de la simulación del capital al tiempo  $t + 1$ . En este caso es,  $VaR(0.1) = 0.0191$  tomando en cuenta que el capital inicial es  $X_0 = 1$ .

El valor en la muestra simulada que dio lugar a este cuantil es la simulación de la distribución conjunta,

$$(0.02623906, 0.01282036, 0.01862230, 0.01873631).$$

En la evaluación del VaR de la sección 3.4 se grafican estos valores únicamente como herramienta visual a modo de comparación con el VaR individual de cada activo. En éste trabajo se calcula el VaR del capital al tiempo  $t$  definido en la ecuación (2.1) con base en la simulación de la distribución conjunta de los rendimientos del portafolio.

Para el cálculo del VaR para un periodo siguiente es necesario repetir todo el procedimiento descrito en esta sección, desde el ajuste del modelo de series de tiempo,

### 3.3. ESTIMACIÓN DE FUNCIONES CÓPULA

---

la estimación de parámetros para las distribuciones marginales hasta la estimación de los parámetros de la cópula. En la Sección 3.4, se muestra una simulación de éste procedimiento par el primer trimestre de 2008.

### 3.3.2. Aplicando marginales con valores extremos y el método PCC

En la sección anterior se presentó una alternativa más simple del enfoque que hemos estudiado, pues las marginales se estimaron con distribuciones  $t$ -Student y una función cópula 4-variada. Ahora en esta sección presentamos los resultados de la estimación de cópulas bivariadas del método *Pair-Copula-Construction* de la Sección 2.4.1 y las distribuciones marginales aplicando valores extremos por medio del ajuste de distribuciones Pareto Generalizadas para las cuales, en la Sección 3.2 se mostraron los resultados de la estimación.

Partiendo de que se han estimado las distribuciones marginales, para cada cópula bivariada que aparece en la descomposición de la distribución multivariada en la ecuación (2.18) se elige la función cópula bivariada que mejor se ajusta a cada par de acuerdo con el valor de la  $\tau$  de Kendall que se introdujo en la Sección 1.3.3. Para la elección de las funciones cópula consideramos los tipos de cópula introducidos en la Sección 1.3, las cuales son Gaussiana,  $t$ -Student, Clayton, Frank y Gumbel.

Es importante mencionar que hay representaciones gráficas de la construcción para la función de distribución conjunta de un vector aleatorio, a través de cópulas por pares las cuales se conocen como *Copulas Vine*. La forma canónica de esta construcción es la que se ha mostrado en la ecuación (2.18). Sin embargo, esta construcción no es única pues hay diferentes posibilidades de condicionar las variables aleatorias del vector aleatorio en cuestión, cada una de estas maneras se pueden representar por medio de árboles anidados. En el trabajo de *Aas et.al.* (2009) se describen dos construcciones principales, una es la forma canónica de las cópulas *Vine* y otra conocida como *D-Vine Copula*, las cuales se describen brevemente en la siguiente sección.

### 3.3.3. Vine Copulas

Note que para vectores aleatorios de dimensión grande, hay un número significativo de posibles construcciones de la distribución conjunta a través de cópulas por pares. Para la organización de estas posibilidades, en el trabajo de *Bedford & Cooke* (2001) se introdujo un modelo gráfico conocido como *Regular Vine* ó *R-Vine*, este tipo de estructuras incluyen los dos tipos de cópulas *Vine* que mencionaremos a continuación.

### 3.3. ESTIMACIÓN DE FUNCIONES CÓPULA

En una estructura del tipo *D-Vine* en todo árbol  $T_j$  ningún nodo está conectado con más de dos ejes, para el caso de 4 variables mostramos la construcción de este tipo en la siguiente figura.

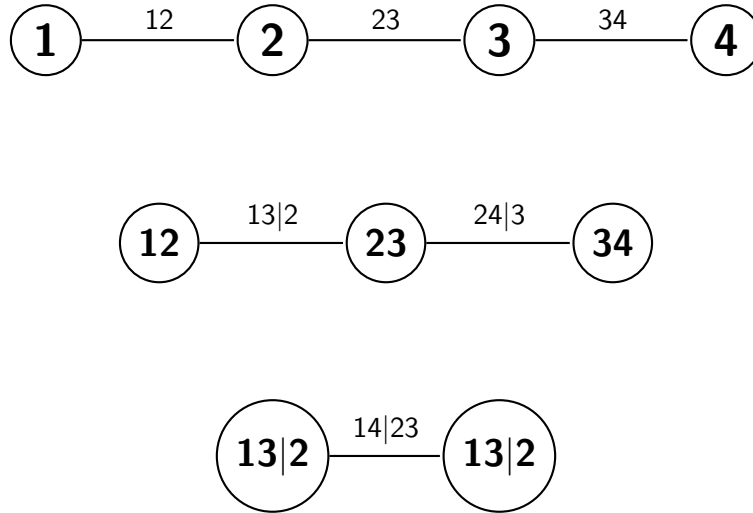
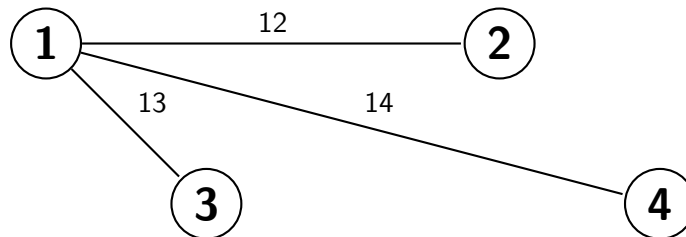


Figura 3.18: Estructura *D-Vine* para la construcción de la densidad conjunta de 4 variables.

De acuerdo con la estructura *D-Vine*, la función de densidad conjunta está dada por,

$$\begin{aligned}
 f(x_1, x_2, x_3, x_4) &= f(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\
 &\quad \cdot c_{1,2}(F_1(x_1), F_2(x_2)) c_{2,3}(F_2(x_2), F_3(x_3)) \\
 &\quad \cdot c_{3,4}(F_3(x_3), F_4(x_4)) \\
 &\quad \cdot c_{1,3|2}(F(x_1|x_2), F(x_3|x_2)) \cdot c_{2,4|3}(F(x_2|x_3), F(x_4|x_3)) \\
 &\quad \cdot c_{1,4|2,3}(F(x_1|x_2, x_3), F(x_4|x_2, x_3))
 \end{aligned} \tag{3.1}$$

Por otra parte, en la construcción de la forma canónica en las *Vine-Copula* cada árbol  $T_j$  tiene un único nodo conectado con  $n-j$  ejes. En la siguiente figura se presentan las etapas para esta construcción.



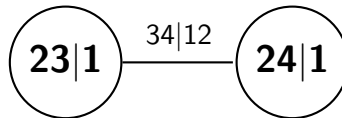
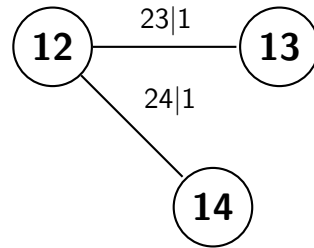


Figura 3.19: Estructura canónica de *R-Vine* para 4 variables.

La expresión para la densidad conjunta aplicando esta estructura es la derivada en la ecuación (2.18). Es relevante notar que para un vector aleatorio de dimensión 4 hay en total 12 estructuras distintas del tipo *D-Vine* y 12 descomposiciones canónicas diferentes, además todas son distintas entre sí; por lo tanto, hay 24 maneras posibles para la construcción de la densidad conjunta para dicho vector aleatorio.

En la práctica, para un tipo de estructura seleccionado se elige cuál de las estructuras es más adecuada con base en calcular la  $\tau$  de Kendall, introducida en la Sección 1.3.3, la cual es una medida de dependencia de acuerdo con las funciones cópulas bivariadas estimadas. En éste trabajo para seleccionar esta estructura, aplicamos el paquete *VineCopula* de R.

De este modo, aplicando esta metodología para las series de residuos estandarizados con el fin de obtener su distribución conjunta, obtenemos que la mejor estructura canónica está representada por los árboles en la siguiente figura.



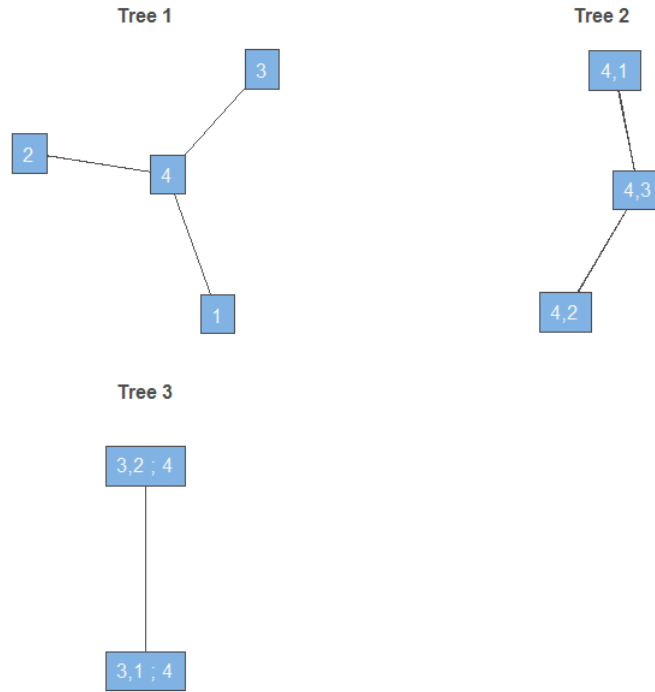


Figura 3.20: Estructura canónica para los residuos estandarizados, la numeración 1,2,3,4 corresponde a las variables EUR, USD, GBP, JPY, respectivamente.

Así, la función de densidad conjunta está dada por,

$$\begin{aligned}
 f(x_1, x_2, x_3, x_4) = & f(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\
 & c_{3,1}(F_3(x_3), F_1(x_1)) \\
 & \cdot c_{3,2}(F_3(x_3), F_2(x_2)) \\
 & \cdot c_{3,4}(F_3(x_3), F_4(x_4)) \\
 & \cdot c_{4,1|3}(F(x_4|x_3), F(x_1|x_3)) \\
 & \cdot c_{4,2|3}(F(x_4|x_3), F(x_2|x_3)) \\
 & \cdot c_{2,1|4,3}(F(x_2|x_4, x_3), F(x_1|x_4, x_3))
 \end{aligned}$$

Las funciones cópulas bivariadas para éstos datos considerando los parámetros estimados se muestran en la siguiente gráfica.

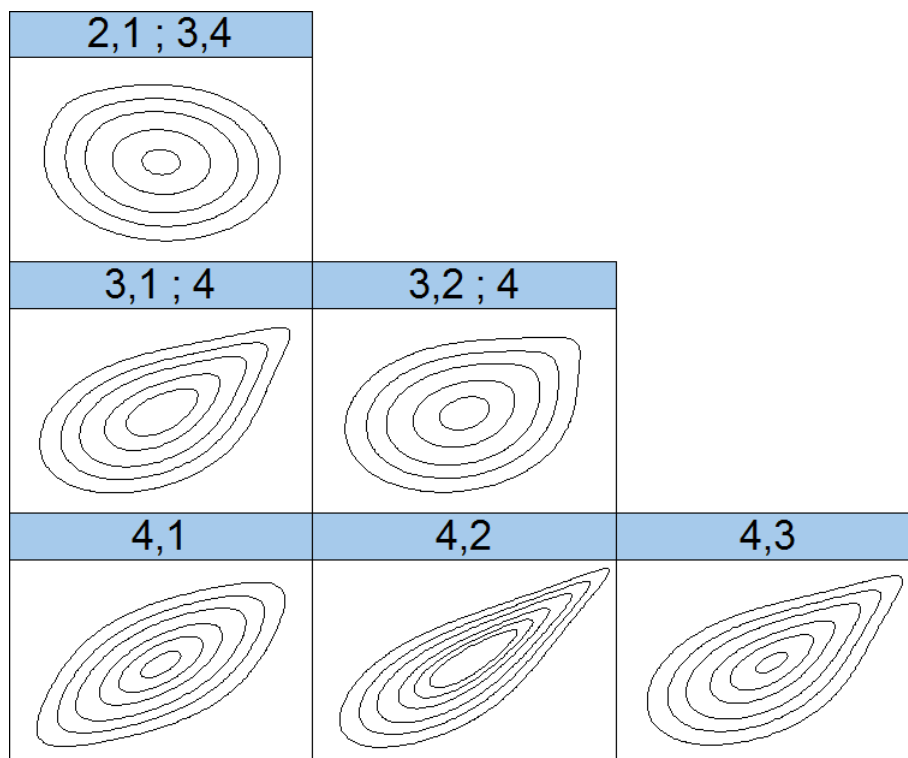


Figura 3.21: Contornos de funciones cópula bivariadas con los parámetros estimados.

Estos resultados fueron obtenidos usando la función *RVineStructureSelect* del paquete *VineCopula* de R. Los parámetros estimados para estas funciones se resumen en la siguiente tabla.

Árbol	Variables	Función Copula	parámetro 1	parámetro 2	$\tau$
1	4,1	t-Student	0.62	7.36	0.43
1	4,2	Gumbel	2.37	0.00	0.58
1	4,3	Gumbel	1.68	0.00	0.41
2	3,1;4	Gumbel	1.51	0.00	0.34
2	3,2;4	Gumbel	1.20	0.00	0.17
3	2,1;3,4	Gumbel Rotada 90°	-1.05	0.00	-0.04

Cuadro 3.6: Funciones cópula bivariadas con sus parámetros estimados por máxima verosimilitud.

### 3.3. ESTIMACIÓN DE FUNCIONES CÓPULA

---

Considerando la construcción para la densidad conjunta anterior y los parámetros estimados para las funciones cópula bivariadas, realizamos una simulación de datos para realizar la estimación Montecarlo del cuantil de la distribución conjunta de los residuos estandarizados. Con 3000 observaciones, los histogramas para cada entrada de los vectores aleatorios simulados, se muestran en las gráficas de abajo.

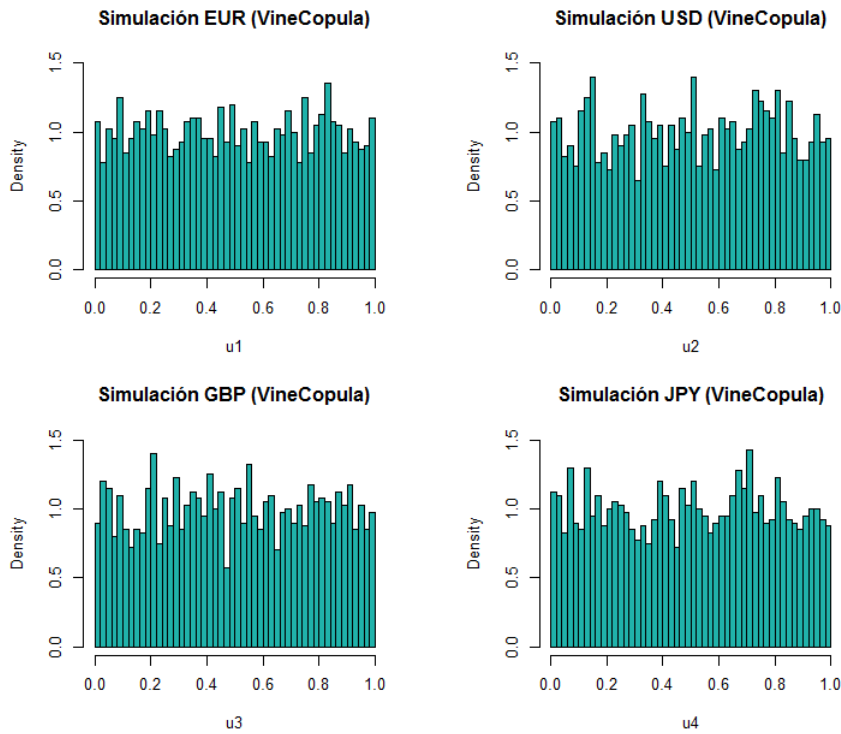


Figura 3.22: Simulación de datos con la función de densidad conjunta estimada.

Luego transformamos estos datos con la inversa de la función de distribución para cada variable, de acuerdo con las distribuciones estimadas en la Sección 3.2.2. Con ello, obtenemos datos que se distribuyen como los residuos estandarizados y que contienen la estructura de dependencia estudiada. Los histogramas de los datos transformados se muestran a continuación.

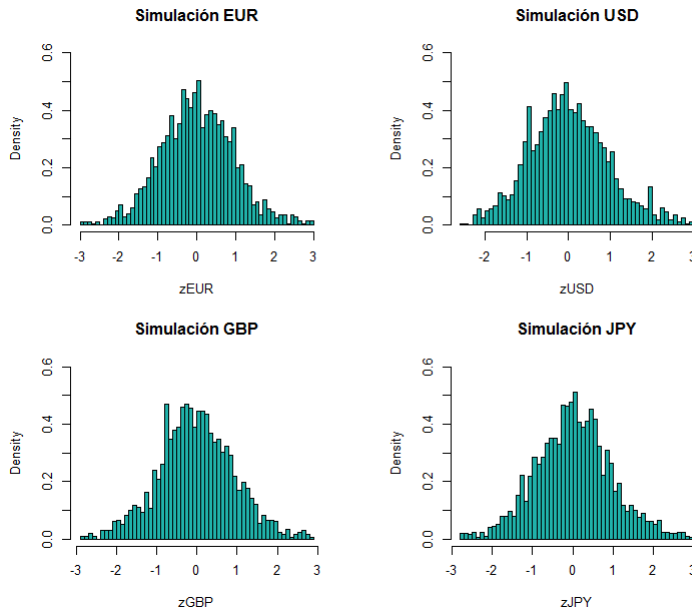


Figura 3.23: Transformación de los datos simulados aplicando la inversa de las distribuciones marginales en cada caso.

Considerando esta simulación y la estrategia de inversión  $\omega = (0.25, 0.25, 0.25, 0.25)$  se va a calcular el cuantil empírico de nivel  $\alpha = 0.1$  para el capital  $X_{t+1}$  de la ecuación (2.1). Para ello, para cada elemento de la simulación de residuos estandarizados a través de la distribución conjunta estimada, se recuperan los rendimientos para cada activo de acuerdo con la estimación de los coeficientes para el modelo de series de tiempo:

$$\begin{aligned}\mu_t &= (0.0004743632, -0.0002785863, 0.0003916893, 0.0003097917) \\ \sigma_t &= (0.01022235, 0.01674258, 0.01400639, 0.01119044)\end{aligned}$$

De este modo, de la ecuación (2.7) respectiva a el modelo de series de tiempo de tenemos que los rendimientos del portafolio están dados por  $\hat{r}_t^q = \mu_t + \hat{z}_t^\alpha * \sigma_t$  y aplicando la ecuación (2.1), tenemos que el cuantil empírico de nivel  $\alpha = 0.1$  es,

$$VaR(0.1) = 0.02333741$$

El vector de rendimientos que dio lugar a este valor en la muestra simulada fue el vector  $(0.01935699, 0.03292662, 0.02033777, 0.02072824)$ , el cual se grafica en la siguiente sección a modo de comparación con el VaR individual de cada activo.

De manera análoga al caso anterior, en la Sección 3.4, se muestra una simulación de éste procedimiento par el primer trimestre de 2008 y otro durante 2009 como un ejemplo del comportamiento del modelo cuando se tienen datos con alta volatilidad. Finalmente se presentan los resultados de backtesting como evaluación de la estimación del VaR para este modelo.

## 3.4. Cálculo del VaR del portafolio

En esta sección presentamos los cálculos del VaR a un nivel de confianza del 90% para el portafolio

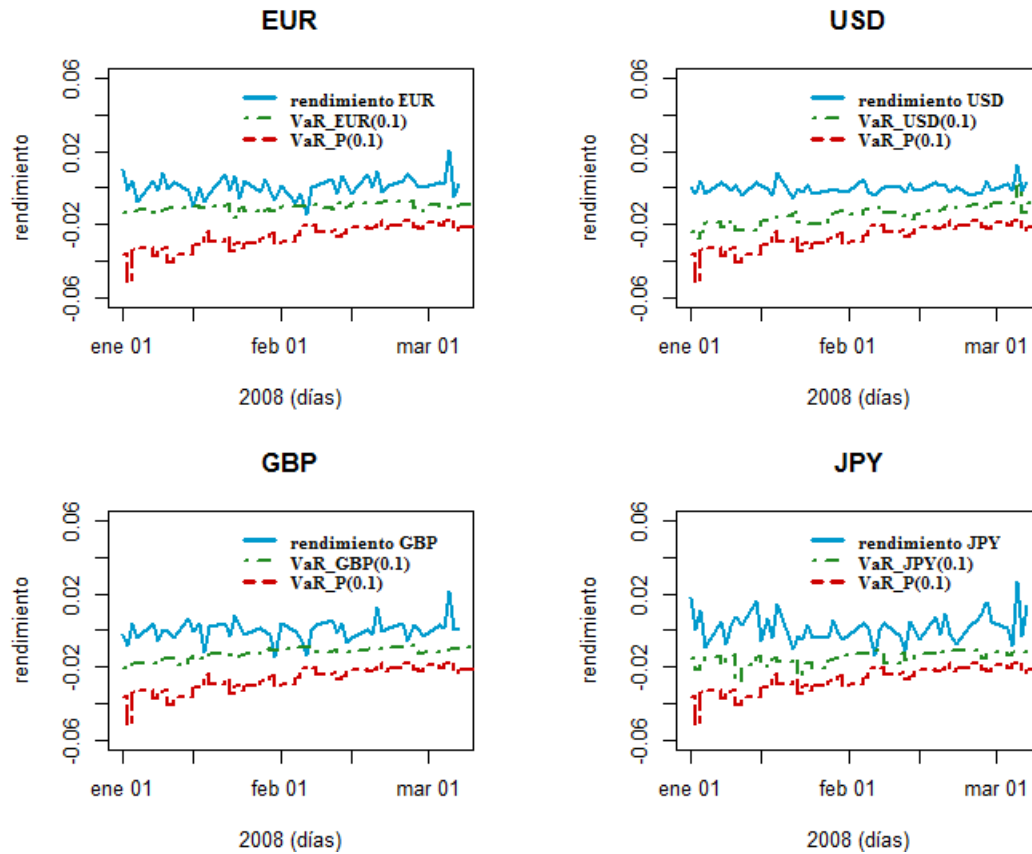
$$\omega = (0.25, 0.25, 0.25, 0.25),$$

en cada periodo del rango en el que evaluamos se sigue esta estrategia de inversión.

Notemos que para obtener estos resultados, en cada periodo de tiempo se requiere hacer la estimación de los parámetros del modelo de series de tiempo, ajustar las distribuciones marginales de las divisas y de las funciones cópula. Esto es, para cada periodo de tiempo se repite el procedimiento descrito de la Sección 3.1 a la 3.3 considerando en cada caso los 1000 rendimientos históricos de las divisas, contando con ésto se calcula el VaR en un horizonte de inversión correspondiente al primer trimestre de 2008.

### 3.4.1. Cálculo del VaR aplicando marginales $t$ -Student y cópula 4-variada

En esta sección presentamos primero una versión simplificada del modelo que hemos estudiado, la cual considera el ajuste del modelo con una cópula de dimensión 4 y marginales  $t$ -Student, en la siguiente sección se presentarán los resultados usando teoría de valores extremos y la construcción con copulas *Vine*. En la siguiente gráfica se muestra el cálculo del VaR para el portafolio así como los cálculos de VaR para cada una de las divisas para esta versión simplificada.

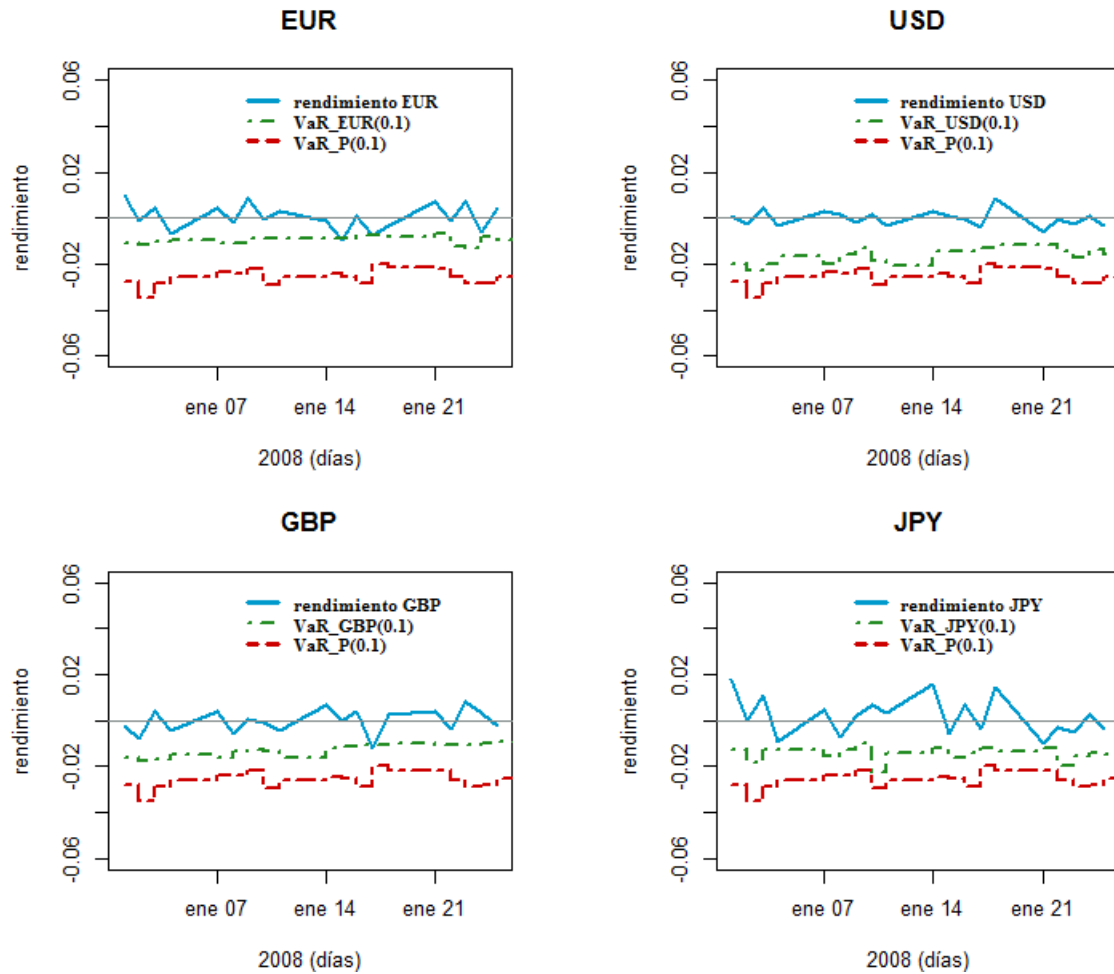


Observamos que para este periodo de tiempo en que hemos evaluado el VaR del portafolio, parece adecuado como una cota para la pérdida máxima esperada de los rendimientos del portafolio. Dicha simulación ha sido obtenida considerando el ajuste de las distribuciones marginales  $t$ -Student para los residuos estandarizados de cada divisa y estimando los parámetros de una cópula 4-variada, sin embargo, esta medida de riesgo no dio buenos resultados para periodos con mucha volatilidad, como en la crisis de 2008.

### 3.4.2. Cálculo del VaR aplicando la estimación de cópulas por pares

En esta sección se presenta la evaluación del cálculo del VaR al nivel  $\alpha = 0.1$  para los rendimientos del portafolio  $\omega = (0.25, 0.25, 0.25, 0.25)$ , de manera análoga a la sección anterior aquí evaluamos el cálculo en el mes de enero de 2008, pues el procedimiento computacional es más elaborado que en la versión que se mostró en la sección anterior.

### 3.4. CÁLCULO DEL VAR DEL PORTAFOLIO



Observamos que durante este periodo el cálculo de VaR para los rendimientos del portafolio parece razonable considerarlo como la pérdida máxima esperada para cada tiempo dentro del periodo que estamos evaluando.

Realizamos también este cálculo en otro periodo de mayor volatilidad, como la primera quincena de 2009, periodo dentro de la crisis que inició en 2008. Se observó que en el caso de aplicar la versión simplificada con la cópula 4–variada y marginales  $t$ –Student el cálculo del VaR no representa la pérdida máxima observada en los rendimientos ni en el capital del portafolio. En cambio, al aplicar la metodología usando marginales aproximadas por distribuciones Pareto Generalizadas y la construcción de cópulas por pares de la Sección 3.3.3 con cópulas *Vine* se obtienen mejores resultados, los cuales se presentan en la siguiente sección.

Presentamos el VaR calculado como una cota del capital invertido en el portafolio, considerando un capital inicial  $X_0 = 1$  y al considerar la estrategia de inversión  $\omega = (0.25, 0.25, 0.25, 0.25)$  en cada tiempo, entonces la evolución del capital y la estimación

del VaR al nivel  $\alpha = 0.1$  se muestra en la siguiente gráfica.

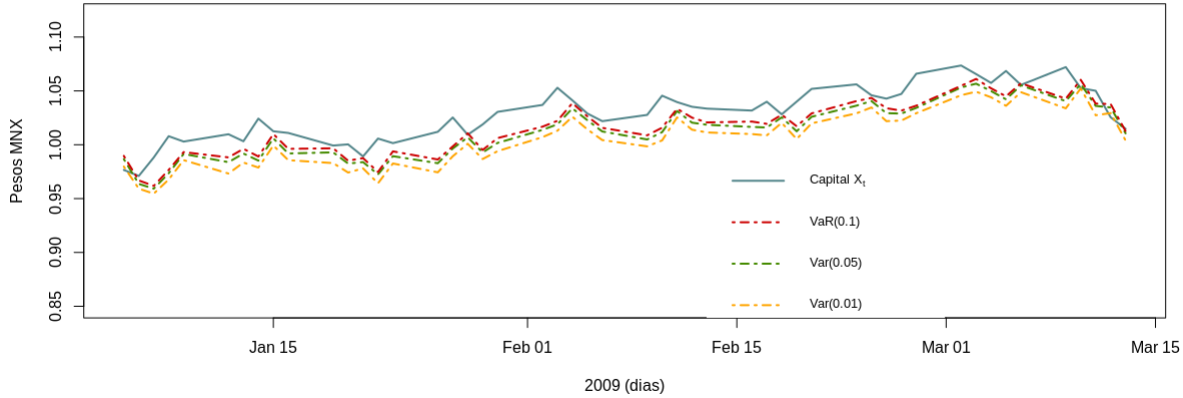


Figura 3.24: Evolución del capital  $X_t$  y cálculo diario del VaR a diferentes niveles de confianza  $\alpha = 0.01, 0.05, 0.10$ . La evaluación se realiza para el primer trimestre de 2009 iniciado con un capital inicial  $X_0 = 1$  y manteniendo la estrategia de inversión  $\omega = (0.25, 0.25, 0.25, 0.25)$ .

Se muestra la evolución del capital  $X_t$ , dado en la ecuación 2.1. Aquí se presenta la evaluación del VaR para el primer trimestre de 2009, es importante notar que en este periodo de tiempo la volatilidad en los precios de las divisas es alta, sin embargo al calcular el VaR con la metodología presentada en este trabajo aplicando las distribuciones de valores extremos y la construcción de cópulas por pares podemos considerar el VaR como una medida de riesgo razonable para el portafolio en cuestión.

Podemos evaluar la estimación del VaR del modelo propuesto por medio de los procedimientos de *backtesting* los cuales consisten en algunas técnicas estadísticas para comprobar si la estimación del VaR es razonable, las pruebas comprendidas por dicho procedimiento se presentaron en la Sección 1.4.1, a continuación se presentan los resultados de estas pruebas para el modelo anterior.

Var(q)	Proporción excepciones VaR( $\hat{q}$ )	Estadística	p-valor
<b>VaR(0.1)</b>	0.1	0	1
<b>VaR(0.05)</b>	0.06	0.09921106	0.75277845
<b>VaR(0.01)</b>	0.04	2.5910982	0.1074658

Cuadro 3.7: Prueba de Kupiec para diferentes niveles en la estimación del VaR.

Observamos que la proporción de excepciones del VaR, es decir, la proporción de veces en que la pérdida observada en los datos reales fue mayor a la estimación del VaR



está cerca del nivel  $\alpha$ , para los casos cuando  $\alpha = 0.01$  y  $\alpha = 0.05$  y cuando  $\alpha = 0.1$  coincide con la proporción de excepciones. Sin embargo, una de las recomendaciones al aplicar las pruebas de *backtesting* es que se evalúe la estimación del VaR en periodos por lo menos de un año. En este caso de estudio, por la complejidad computacional es complicado evaluar el modelo en periodos grandes de tiempo. En la sección siguiente se presenta la estimación para un periodo de tiempo mayor y además se realiza el ajuste de las series de tiempo únicamente para la volatilidad.

#### 3.4.3. Evaluación de la estimación del VaR

*Atendiendo a las recomendaciones para el ajuste del modelo de series de tiempo, se presenta lo siguiente:*

En esta sección se va a evaluar la estimación del VaR a los niveles de confianza 95 % y 99 % para el modelo estudiado, donde en particular se ajusta un modelo de series de tiempo GJR–GARCH(p,q), es decir, sin la parte ARMA. Esto con motivo que al analizar los datos para ajustar la serie de tiempo, se tiene que para este caso es más adecuado este tipo de modelos. Para el ajuste del modelo de series de tiempo, se aplicaron las funciones *ugarchspec* y *ugarchfit* del paquete *rugarch*, a diferencia del caso anterior, aquí se consideró la distribución de error generalizada asimétrica "sged" para el cálculo de la verosimilitud, a través de la función,

```
ugarchspec(variance.model=list(model="gjrGARCH", garchOrder=c(p,q),  
variance.targeting=TRUE), mean.model=list(armaOrder=c(0,0),  
include.mean=T), distribution.model="sged")
```

Luego se obtienen los residuos estandarizados con base en este ajuste y se procedió a estimar las distribuciones de valores extremos Pareto Generalizadas, la estimación de las cópulas por pares y la simulación Montecarlo para el cálculo del VaR del capital  $X_t$  de la ecuación (2.1), en la siguiente gráfica se muestra la evaluación durante el año 2009.

Recordando que el objetivo de este trabajo es calcular el riesgo del portafolio considerado, no necesariamente la estrategia de inversión que se considera es la mejor. Sin embargo, aún en escenarios de alta volatilidad, como se presenta en el periodo de tiempo que estamos considerando, quisiéramos evaluar si el modelo propuesto brinda estimaciones razonables del riesgo por medio de la validación del cálculo del VaR. Para ello, considerando la prueba de Kupiec y de Christoffersen se obtienen los resultados siguientes.

Prueba	Proporción excepciones VaR( $\hat{q}$ )	Estadística	p-valor
Kupiec	0.068	1.5402866	0.2145751
Christoffersen	0.068	2.257224	0.323482

Cuadro 3.8: Backtesting del VaR(0.05) evaluado en días de 2009.

Por lo tanto, como el p-valor de cada una de las pruebas es relativamente grande, no hay evidencia en contra de la hipótesis que  $\hat{q} = \alpha$ , lo que hace al modelo razonable para el cálculo del VaR. Aunque notamos que la proporción de excepciones del VaR es más alta de lo que se esperaba.

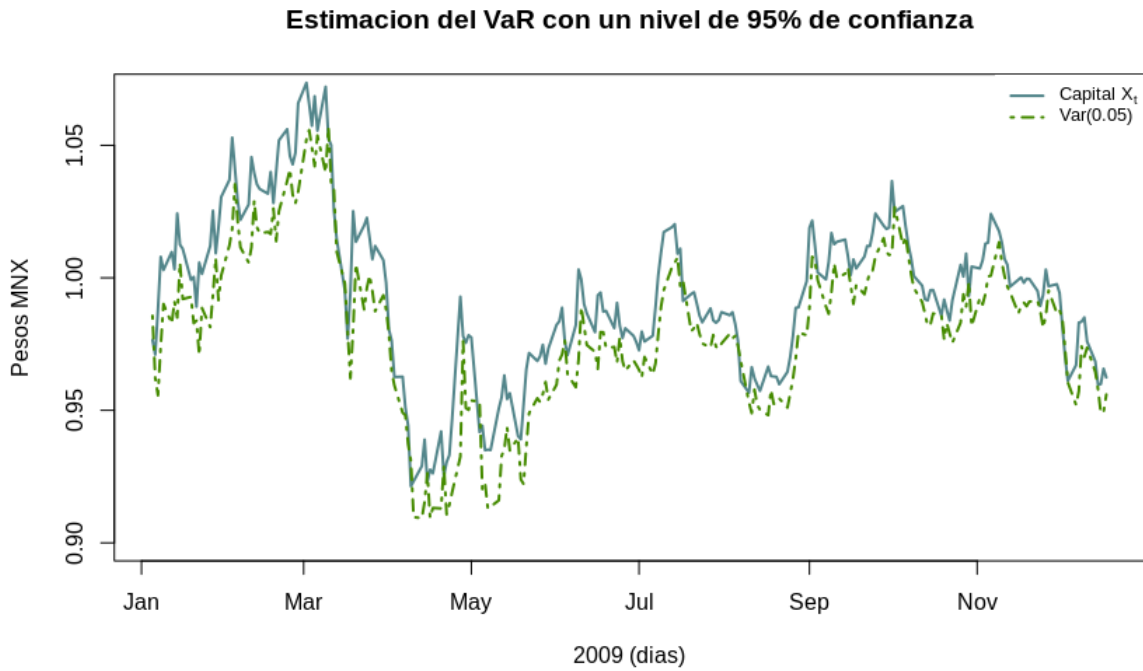


Figura 3.25: Evolución del capital  $X_t$  durante 2009 y cálculo diario del VaR al nivel de confianza  $\alpha = 0.05$ . El capital inicial fue  $X_0 = 1$  y se siguió la estrategia de inversión  $\omega = (0.25, 0.25, 0.25, 0.25)$ .

De manera análoga, se calcula el VaR al nivel  $\alpha = 0.01$  y se realiza la evaluación de dicha estimación para el mismo periodo de tiempo y aplicando el mismo modelo que el presentado anteriormente.

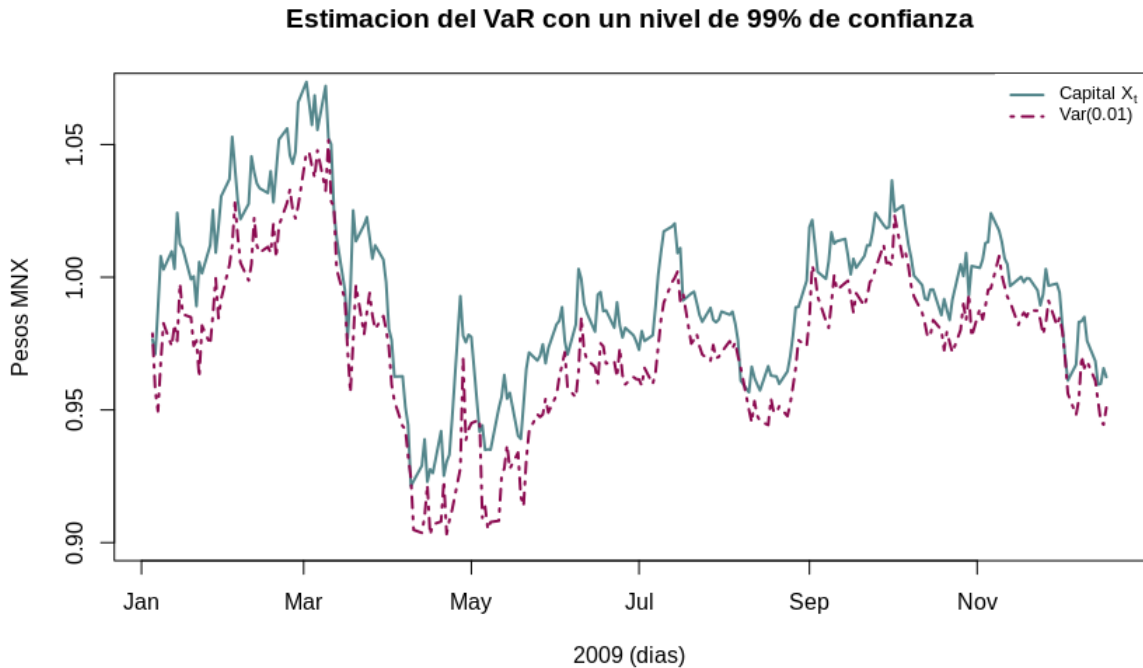


Figura 3.26: Evolución del capital  $X_t$  durante 2009 y cálculo diario del VaR al nivel de confianza  $\alpha = 0.01$ . El capital inicial fue  $X_0 = 1$  y se siguió la estrategia de inversión  $\omega = (0.25, 0.25, 0.25, 0.25)$ .

Prueba	Proporción excepciones $\text{VaR}(\hat{q})$	Estadística	p-valor
Kupiec	0.028	5.49699045	0.01904923
Christoffersen	0.028	5.84342394	0.05384143

Cuadro 3.9: Backtesting del  $\text{VaR}(0.01)$  evaluado en días de 2009.

Notemos que al pedir un nivel de confianza del 99%, la prueba de Kupiec rechaza la hipótesis que  $\hat{q} = \alpha$  esto hace que la estimación del  $\text{VaR}(0.01)$  no sea adecuada pues no se obtiene el nivel deseado. Aquí se está pidiendo un nivel de confianza más alto, por lo que sería adecuado realizar la evaluación en un periodo de tiempo más grande y tomar en cuenta dichos resultados, esto es complicado por la complejidad computacional del modelo completo.

Por otra parte, como trabajo futuro será interesante considerar este modelo en el contexto de optimización de portafolios con el fin de encontrar un portafolio de inversión óptimo pues a lo largo del trabajo se consideró un portafolio particular; sin embargo, el enfoque podría extenderse a resolver dicho problema.



# Capítulo 4

## Conclusiones y trabajo futuro

- En este trabajo se estudió el modelo GJR–GARCH–EVT–Cópula para el cálculo de VaR de un portafolio formado por las divisas euro, dólar, libra esterlina y yen japonés; además se implementó este modelo con datos reales.
- Se realizó un ajuste para las distribuciones marginales de los residuos estandarizados aplicando teoría de valores extremos con distribuciones Pareto Generalizadas, además se consideran criterios para la elección de umbrales razonables.
- Se observó que los residuos estandarizados presentan menos correlación que la serie de rendimientos original, sin embargo, por la naturaleza de los datos aún está presente dicha dependencia, lo que explica algunas de las características observadas en el ajuste del modelo.
- En la implementación de cópulas bivariadas para construir la densidad conjunta se usaron cópulas *Vine*, lo que permite incluir en el modelo las relaciones de dependencia entre las divisas; esto representa una contribución respecto al trabajo de *Omarí et.al.* (2018).
- Se evaluó el modelo en un periodo de tiempo al inicio de 2009, en el cual los precios tuvieron una volatilidad alta. Los resultados de las pruebas de backtesting para la estimación del VaR en los niveles  $\alpha = 0.1$ ,  $\alpha = 0.05$  y  $\alpha = 0.01$  realizadas, no dan evidencia en contra de que la proporción de excepciones del VaR es cercana a la esperada.
- La evaluación del cálculo del VaR se realizó en horizontes de tiempo más cortos a lo esperado debido a que la estimación de todo el modelo para cada periodo de tiempo implican un tiempo computacional muy alto, dado que en cada tiempo se ajusta un modelo de series de tiempo, se estiman las distribuciones marginales y las funciones cópula.

## Trabajo Futuro

El estudio de este trabajo motiva interesantes vertientes a desarrollar, tales como las que se proponen a continuación.

- Realizar un análisis más profundo sobre los parámetros en el modelo de series de tiempo así como de los criterios de selección. Profundizar y aplicar técnicas de selección de modelos.
- Con lo desarrollado hasta el momento, parece factible aplicar este modelo en un enfoque de optimización de portafolios para encontrar estrategias de inversión óptimas dado un nivel de riesgo.
- Analizar el resultado de las pruebas de backtesting para horizontes de tiempo más grandes así como estudiar la potencia de dichas pruebas y obtener intervalos de confianza.

# Capítulo 5

## Apéndice: Código en R

En esta sección se presentan los códigos que permitieron obtener los resultados del Capítulo 3, se muestran las partes esenciales y uso de las librerías de R por limitación del espacio en esta tesis.

```
# Los datos de precios diarios de las divisas euro, dólar, libra esterlina
# y yen japonés fueron descargados del sitio investing.com
rm(list=ls())

library(tidyverse)
library(lubridate)

library(rugarch)
library(rmgarch)

dat <- read.csv("Datos-01-00-01-19.csv")

#Nombramos los datos
EUR<- dat[,2]
USD<- dat[,3]
GBP<- dat[,4]
JPY<- dat[,5]

#Cambiamos el orden de los datos de antes a actual
EUR<-EUR[length(EUR):1]
USD<-USD[length(USD):1]
GBP<-GBP[length(GBP):1]
JPY<-JPY[length(JPY):1]

plot(EUR, main="Precios de divisas en MXN", type="l", col=2, ylim=c(0,30),
      xlab="dias", ylab="Pesos")
```

```

lines(USD, lty=4, col=3)
lines(GBP, lty=5, col=4)
lines(JPY, lty=6, col="purple")
legend(x=4000,y=10,legend=c("EUR","USD", "GBP", "JPY"),
fill=c("red","blue", "green", "purple"),cex=0.5,text.font=7)
n=length(EUR)-1

# Calculamos el logaritmo de la tasa de rendimientos para cada serie
REUR=rep(0, n)
RUSD<-rep(0, n)
RGBP<-rep(0, n)
RJPY<-rep(0, n)
for( i in 1:n){
REUR[i]=log(EUR[i+1]/EUR[i])
RUSD[i]=log(USD[i+1]/USD[i])
RGBP[i]=log(GBP[i+1]/GBP[i])
RJPY[i]=log(JPY[i+1]/JPY[i])
}
#Graficamos rendimientos
par(mfrow=c(2,2), mar=c(4,4,4,4))
plot(REUR, type='l', main="EUR", ylab="rt", xlab="dias", col="deepskyblue4")
plot(RUSD, type='l', main="USD", ylab="rt", xlab="dias", col="deepskyblue4")
plot(RGBP, type='l', main="GBP", ylab="rt", xlab="dias", col="deepskyblue4")
plot(RJPY, type='l', main="JPY", ylab="rt", xlab="dias", col="deepskyblue4")

Ret_divisas<-cbind(EUR, cbind(USD, cbind(GBP, JPY)))
#Matriz de correlación de los rendimientos de las divisas
cor(Ret_divisas)

#Convertimos los datos en tipo series de tiempo
eur.ts=ts(REUR)
usd.ts=ts(RUSD)
gbp.ts=ts(RGBP)
jpy.ts=ts(RJPY)

#Se estiman los parametros del modelo por máxima verosimilitud para cada
elección de órdenes m,n,p y q del modelo.

model.eur <- ugarchspec(variance.model=list(model="gjrGARCH", garchOrder=c(p,q),
variance.targeting=TRUE), mean.model=list(armaOrder=c(m,n), include.mean=T),
distribution.model="std")
gjrgarch.eur = ugarchfit(model.eur,eur.ts)

```



---

```

#verosimilitud.eur[p,q]=
verosimilitud.eur.opt=likelihood(gjrgarch.eur)
#AIC.eur[p,q]= as.numeric(Infocriteria(gjrgarch.eur)[1])

#Calculamos los criterios AIC y AICC para la elección del modelo.
AIC<-((-2*verosimilitud.eur.opt)/length(eur.ts))
+2*(length(gjrgarch.eur@fit$coef))/length(eur.ts)
AICC<- ((2*(p+q+1)*length(eur.ts))/(length(eur.ts)-p-q-2))
-2*log(verosimilitud.eur.opt)/length(eur.ts)

#Se repite este procedimiento con cada divisa y se seleccionan en cada caso
parámetros m,n, p y q para cada modelo de series de tiempo. Una vez elegidos,
se calculan los residuos estandarizados para cada serie.

zt.eur<- residuals(gjrgarch.eur)/sigma(gjrgarch.eur)
zt.usd<- residuals(gjrgarch.usd)/sigma(gjrgarch.usd)
zt.gbp<- residuals(gjrgarch.gbp)/sigma(gjrgarch.gbp)
zt.jpy<- residuals(gjrgarch.jpy)/sigma(gjrgarch.jpy)

#Estimación de densidades marginales aplicando EVT
library(evmix)
#Estimación de densidad por kernel
library(MASS)

##### Seleccion de umbral
useq<-seq(0.001,3.6,0.001)
x<-sort(as.numeric(zt.eur[1:1000])), decreasing=FALSE)
mrlplot(x, main = "Mean Residual Life Plot EUR upper", xlab = "Threshold u",
  ylab = "Mean Excess")
mrlplot(-x, main = "Mean Residual Life Plot EUR lower", xlab = "Threshold u",
  ylab = "Mean Excess")
#Análisis de la estabilidad del parámetro de forma
par(mfrow=c(1,1))
tshapeplot(x, tlim = c(0, 2.5), ylim = c(-0.7,0.7),xlim=c(1,1.5), main =
  "Estabilidad del parametro de forma (EUR tail upper)", legend.loc = "topleft",
  xlab = "umbral(u)", ylab = expression(paste("Parámetro de forma", (xi))),
  try.thresh=c(1.2,1.3,1.4))
tshapeplot(-x, tlim = c(0, 2.5), ylim = c(-1.2,1.2), main = "Estabilidad del
  parametro de forma (EUR tail lower)")
#El análisis de las graficas sobre la estabilidad del parámetro de forma
  permiten elegir umbrales adecuados, luego de la elección para ambos umbrales
  se estiman los parámetros de las colas con distribuciones GPD.

```

```

# Función de densidad cola derecha
tail_upper_eur<-fgpd(x, u = 1.4)
u_up= as.numeric(tail_upper_eur$u[1])
xx=seq(u_up+0.001,3.5,0.01)
k_up<-x[x>as.numeric(tail_upper_eur$u[1]) ]
k_up=length(k_up)
n_up=length(x)

f_density_upper<- function( z, u, sigmaup, xi, nt, nu){
d= (nu/(nt*sigmaup))*((1+xi*((z-u)/sigmaup))^(-(1+xi)/xi))
return(d)
}

vf_density_upper<-Vectorize(f_density_upper, vectorize.args = "z")
density_up_2=vf_density_upper(xx, u_up ,as.numeric(tail_upper_eur$mle[1]),
  as.numeric(tail_upper_eur$mle[2]), n_up, k_up)

# Función de densidad cola izquierda
tail_low_eur<-fgpd(-x, u = 1.2)
u_lw=as.numeric(tail_low_eur$u[1])
limlw=-u_lw+0.001
yy<-seq(-3.5, limlw, 0.01)
k_lw<-x[x<=-as.numeric(tail_low_eur$u[1]) ]
k_lw=length(k_lw)
f_density_lower<- function( z, u, sigmaup, xi, nt, nu){
d= (-nu/(nt*sigmaup))*((1+xi*((u-z)/sigmaup))^(-(1+xi)/xi))
return(d)
}

vf_density_lower<-Vectorize(f_density_lower, vectorize.args = "z")
density_lower=vf_density_lower(yy, -u_lw, as.numeric(tail_low_eur$mle[1]),
  as.numeric(tail_low_eur$mle[2]), n_up, k_lw)
#Estimación de densidad por kernel gaussiano para el interior
ddist<-density(x, kernel = "gaussian")

FunionDensity_EUR<- function(x){
if(x<=-u_lw ){
# print('lw')
return(-vf_density_lower(x, -u_lw, as.numeric(tail_low_eur$mle[1]),
  as.numeric(tail_low_eur$mle[2]), n_up, k_lw))
}
}

```

---

```

if(x>=u_up){
# print('up')
return(vf_density_upper(x, u_up ,as.numeric(tail_upper_eur$mle[1]),
  as.numeric(tail_upper_eur$mle[2]), n_up, k_up))
}
# else{
# print('m')
i=which(ddist$x>=x)[1]
valf=ddist$y[i]
return(valf)
# }
}

densidadMarginal_EUR<-Vectorize(FunionDensity_EUR, vectorize.args = "x")
DistribucionMarginalEUR<- function(x){
distr<-integrate(densidadMarginal_EUR, lower=-3, upper=x,subdivisions=1000)$value
return(distr)
}
DistribucionMarginalEUR<-Vectorize(DistribucionMarginalEUR, vectorize.args = "x")

#Densidad margunal EURO
Ddens_EUR<-densidadMarginal_EUR(seq(-4,4,0.01))

#Residuos transformados a uniformes (0,1) aplicando la distribución EVT
UResi_EVT_EUR<-DistribucionMarginalEUR(as.numeric(zt.eur[1:1000]))
#hist(UResi_EVT_EUR, breaks = 50, main = 'Residuos Transformados EUR (EVT)',
  xlab = 'residuos', freq = F, col = "lightseagreen", ylab = "Densidaad")
UResi_EVT_USD<-DistribucionMarginalUSD(as.numeric(zt.usd[1:1000]))
UResi_EVT_GBP<-DistribucionMarginalGBP(as.numeric(zt.gbp[1:1000]))
UResi_EVT_JPY<-DistribucionMarginalJPY(as.numeric(zt.jpy[1:1000]))

#La estimación de la distribucion marginal que presentamos para el caso del
euro, se repite para las otras divisas, y para cada una se obtienen los
residuos estandarizados transformados a uniformes (0,1) para proceder a
la estimación de las cópulas.

#####
# Estimacion de Cópula
library(copula)
library(scatterplot3d)
library(ggplot2)

```

```

library(grid)
library(VineCopula)

uni_eur<-UResi_EVT_EUR
uni_usd<-UResi_EVT_USD
uni_gbp<-UResi_EVT_GBP
uni_jpy<-UResi_EVT_JPY

#Construccion de una RVine Copula

udat<-cbind(as.numeric(uni_eur),cbind(as.numeric(uni_usd),
cbind(as.numeric(uni_gbp),as.numeric(uni_jpy))))
udat[which(udat<=0)]=udat[which(udat<=0)]*(-1)
fit <- RVineStructureSelect(udat, type = 1, familyset=seq(1,6,1),
rotations = FALSE) #0-Dvine #1-CVine
# plot(fit)
# fit$Matrix
# fit$family
# fit$par
# fit$par2

##### Simulación Montecarlo
nsim<-5000
RvineData=RVineSim(nsim, fit)
# Transformamos las simulaciones con la inversa de la Distribución estimada
inversa_DistMarginal <- function(y, X, Dist_X){
i=which(Dist_X>=y)[1]
valf=X[i]
if(is.na(valf)){valf=X[length(X)]}
return(valf)
}
vecInversa_DistMarginal<-Vectorize(inversa_DistMarginal, vectorize.args = "y")

X=seq(-4,4,0.01)
Inversos_EUR_Sim<-vecInversa_DistMarginal(RvineData[,1],X, DDist_EUR )
Inversos_USD_Sim<-vecInversa_DistMarginal(RvineData[,2],X, DDist_USD )
Inversos_GBP_Sim<-vecInversa_DistMarginal(RvineData[,3],X, DDist_GBP )
Inversos_JPY_Sim<-vecInversa_DistMarginal(RvineData[,4],X, DDist_JPY )
#Quitamos NAN's para calcular VaR individual
Inversos_EUR_Sim <- Inversos_EUR_Sim[is.na(Inversos_EUR_Sim) == F]
Inversos_USD_Sim <- Inversos_USD_Sim[is.na(Inversos_USD_Sim) == F]
Inversos_GBP_Sim <- Inversos_GBP_Sim[is.na(Inversos_GBP_Sim) == F]

```

---

```

Inversos_JPY_Sim <- Inversos_JPY_Sim[is.na(Inversos_JPY_Sim) == F]

alpha = 0.01
#VaR individual para cada divisa
VaR_EUR = quantile(Inversos_EUR_Sim, alpha)
VaR_USD = quantile(Inversos_USD_Sim, alpha)
VaR_GBP = quantile(Inversos_GBP_Sim, alpha)
VaR_JPY = quantile(Inversos_JPY_Sim, alpha)

#####
# Se transforman las simulaciones de residuos estandarizados en
# observaciones de rendimientos respecto a los coeficientes estimados
# en cada serie de tiempo.

VaR_EUR_rt[ii]<-mu_t_EUR[tiempo] + as.numeric(VaR_EUR) *
  as.numeric(sigma(gjrgarch.eur)[tiempo])
VaR_USD_rt[ii]<-mu_t_USD[tiempo] + as.numeric(VaR_USD) *
  as.numeric(sigma(gjrgarch.usd)[tiempo])
VaR_GBP_rt[ii]<-mu_t_GBP[tiempo] + as.numeric(VaR_GBP) *
  as.numeric(sigma(gjrgarch.gbp)[tiempo])
VaR_JPY_rt[ii]<-mu_t_JPY[tiempo] + as.numeric(VaR_JPY) *
  as.numeric(sigma(gjrgarch.jpy)[tiempo])

capital[(ii+1)]<-capital[ii] * (1 + c(REUR_C[(t_fin+1)], RUSD_C[(t_fin+1)],
  RGBP_C[(t_fin+1)], RJPY_C[(t_fin+1)]) %*% c(0.25,0.25,0.25,0.25))

SimRtEUR<-mu_t_EUR[tiempo] + as.numeric(Inversos_EUR_Sim) *
  as.numeric(sigma(gjrgarch.eur)[tiempo])
SimRtUSD<-mu_t_USD[tiempo] + as.numeric(Inversos_USD_Sim) *
  as.numeric(sigma(gjrgarch.usd)[tiempo])
SimRtGBP<-mu_t_GBP[tiempo] + as.numeric(Inversos_GBP_Sim) *
  as.numeric(sigma(gjrgarch.gbp)[tiempo])
SimRtJPY<-mu_t_JPY[tiempo] + as.numeric(Inversos_JPY_Sim) *
  as.numeric(sigma(gjrgarch.jpy)[tiempo])

for(j in 1:nsim){
aux_capital[j]= capital[ii]* ( c(SimRtEUR[j], SimRtUSD[j], SimRtGBP[j],
  SimRtJPY[j]) %*% c(0.25,0.25,0.25,0.25))
}
#####
#Calculamos el VaR del capital del portafolio para los niveles
# alpha=0.01,0.05,0.10

```

```

VaR_capital_01[ii]<- quantile(aux_capital, alpha)
VaR_capital_05[ii]<- quantile(aux_capital, 0.05)
VaR_capital_10[ii]<- quantile(aux_capital, 0.10)

dif_capital01=capital[1:nrep]+VaR_capital_01
dif_capital05=capital[1:nrep]+VaR_capital_05
dif_capital10=capital[1:nrep]+VaR_capital_10
##### Calculo CVaR empirico #####
ddistCVaR<-density(aux_capital, kernel = "gaussian")
Fun_ES<- function(s){
  i=which(ddistCVaR$x>=s)[1]
prodES=ddistCVaR$x[i] * ddistCVaR$y[i]
  return(prodES)
}
vec_Fun_ES<-Vectorize(Fun_ES, vectorize.args = "s")
ES_CVaR<- function(nivel){
  k=min(aux_capital)
  VaR=quantile(aux_capital, nivel)
  ES<-integrate(vec_Fun_ES, lower=k, upper=VaR, subdivisions=700)$value
return(ES)
}
CVaR_capital[ii]<- ES_CVaR(alpha)
CVaR_capital_05[ii]<- ES_CVaR(0.05)
CVaR_capital_10[ii]<- ES_CVaR(0.10)
#####
#Backtesting: Evaluacion del VaR del capital del portafolio
# Purebas Kupiec y Christoffesen
library(GAS)
BackTest99 = BacktestVaR(capital[2:(nrep+1)],dif_capital01, 0.01)
BackTest95 = BacktestVaR( capital[2:(nrep+1)],dif_capital05, 0.05)
BackTest90 = BacktestVaR( capital[2:(nrep+1)],dif_capital10, 0.10)

plot(fechasR[(t1+1):(t2+1)], capital[2:(nrep+1)], type="l", ylim=c(0.90,1.07),
  ylab="Pesos MNX", xlab="2009 (dias)", lwd=2,col= "cadetblue4" ,
  main="Estimacion del VaR con un nivel de 95% de confianza")
lines(fechasR[(t1+1):(t2+1)],dif_capital, col="deeppink4", lwd=2, lty=4)
lines(fechasR[(t1+1):(t2+1)],dif_capital05, col="chartreuse4", lwd=2, lty=4)
legend("topright", legend=c(expression(paste("Capital ", X[t])), "Var(0.05)"),
lty=c(1,4), cex=0.8, box.lty=0, text.font=7, lwd=c(2,2),
col=c("cadetblue4", "chartreuse4"))

```

# Bibliografía

1. Omari, C.O., Mwita, P.N. and Gichuhi, A.W. (2018) *Currency Portfolio Risk Measurement with Generalized Autoregressive Conditional Heteroscedastic-Extreme Value Theory-Copula Model*. Journal of Mathematical Finance, 8, 457-477.  
<https://doi.org/10.4236/jmf.2018.82029>
2. Wang, Z.R., Chen, X.H., Jin, Y.B. and Zhou, Y.J. (2010) *Estimating Risk of Foreign Exchange Portfolio: Using VaR and CVaR Based on GARCH-EVT-Copula Model*. Physica A: Statistical Mechanics and Its Applications, 389, 4918-4928.  
<https://doi.org/10.1016/j.physa.2010.07.012>
3. Tsay, Ruey s. (2002) *Analysis of Financial Time Series* Wiley Series in Probability and Statistics.
4. McNeil, Alexander; Lindskog, Filip; Embrechts, Paul; (2001) *Modelling Dependence with Copulas and Applications to Risk Management*
5. McNeil, A.J;Frey, R. (2000) *Estimation of Tail Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach*. Journal of Empirical Finance, 7, 271-300.
6. Tang, J.; Zhou, C.; Yuan, X.; Sriboonchitta, S. (2015) *Estimating Risk of Natural Gas Portfolios by Using GARCH-EVT-Copula Model*. The Scientific World Journal.
7. Aas, Kjersti; Czado, Claudia; Frigessi, Arnoldo; Bakken, Henrik (2007). *Pair-copula constructions of multiple dependence*. Mathematics and Economics 44 (2009) 182-198.
8. Omari, C. O., Mwita, P. N., & Waititu, A. G. (2017). *Using Conditional Extreme Value Theory to Estimate Value-at-Risk for Daily Currency Exchange Rates*. Journal of Mathematical Finance, 7, 846-870.
9. Joe, Harry; Xu, James Jianmeng (1996). *The estimation method of inference functions for margins for multivariate models*. The University of British Columbia, Vancouver.

10. Patton, A.J. (2006). *Estimation of Multivariate Models for Time Series of Possibly Different Lengths*. Journal of Applied Econometrics, 21, 147-173.
11. Embrechts, P., Klüppelberg, C., Mikosch, T.(1997).*Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
12. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). *Coherent measures of risk*. Mathematical Finance 9, 203–228.
13. Balkema, A., de Haan, L., (1974). *Residual life time at great age*. Annals of Probability 2, 792–804.
14. Pickands, J. (1975). *Statistical inference using extreme order statistics*. The Annals of Statistics 3, 119–131.
15. Smith, R. (1987). *Estimating tails of probability distributions*. The Annals of Statistics 15, 1174–1207.
16. Scarrott, C.J. and MacDonald, A. (2012). *A review of extreme value threshold estimation and uncertainty quantification*. REVSTAT - Statistical Journal 10(1), 33-59. Available from <http://www.ine.pt/revstat/pdf/rs120102.pdf>
17. Coles S.G. (2004). *An Introduction to the Statistical Modelling of Extreme Values*. Springer-Verlag: London.
18. Martín González, Ehyter M. (2018) *Notas del curso Teoría de Extremos CIMAT*.
19. P. Richtárik. (2015). *Optimization methods in finance*, lecture notes.
20. Philippe, Artzner; Freddy, Delbaen; Jean-Marc, Eber; David Heath.(1999) *Coherent measures of risk*. Mathematical Finance, 9(3):203–228, 1999.
21. Acerbi, Carlo; Tasche, Dirk. (2002) *On the coherence of expected shortfall*. Journal of Banking & Finance, 26(7):1487-1503.
22. Bedford, T.; Cooke, R.M., (2001) *Probability density decomposition for conditionally dependent random variables modeled by vines*. Annals of Mathematics and Artificial Intelligence 32, 245-268.
23. Kupiec, PH (1995). *Techniques for Verifying the Accuracy of Risk Measurement Models*. The Journal of Derivatives, 3(2), 73-84.
24. Christoffersen, PF (1998). *Evaluating Interval Rorecasts*. International Economic Review, 39(4), 841-862.