Herrera, Luis Alberto

Acerca de la Teoría de los valores extremos y su aplicabilidad a la estimación del riesgo financiero

Anuario de la Facultad de Ciencias Económicas del Rosario Vol. IX, 2013

Acerca de la Teoría de los Valores Extremos y su aplicabilidad a la Estimación del Riesgo Financiero

Luis Alberto Herrera[14]
Pontificia Universidad Católica Argentina,
Facultad de Ciencias Económicas

Resumen. La hipótesis clásica de normalidad de las distribuciones de las rentabilidades implica el supuesto que los precios de los activos están formados por agregación de "shocks" aleatorios representativos de los factores que impulsan a los operadores a realizar estimaciones sobre el comportamiento de dichos precios y que estas estimaciones individuales constituyen variables aleatorias con varianza finita, independientes entre sí. Pero, diversos estudios realizados sobre los retornos de los activos financieros en mercados tradicionales y/o emergentes nos indican que los mismos suelen tener "colas" de distribución pesadas, o lo que es lo mismo, suelen presentar mayores probabilidades de ocurrencia de eventos riesgosos. Han surgido desde la teoría moderna de portafolios, distintos intentos de solución para esta problemática. En este sentido podemos señalar que una alternativa muy estudiada últimamente, para representar el comportamiento de las rentabilidades, es a partir de las distribuciones de valores extremos que consideran exclusivamente la distribución de las rentabilidades altas y de pérdidas excepcionales. Este enfoque motiva el presente trabajo de investigación, en el cual estudiaremos la caracterización de las "colas" de distribución pesadas en el contexto de la Teoría de los valores extremos, a partir de este marco teórico, podremos inferir medidas de riesgo adecuadas para caracterizar los retornos de activos en los mercados tradicionales y más precisamente en mercados de

[14] lherrera@uca.edu.ar

las economías emergentes o en vías de desarrollo, caracterizados principalmente por una distribución de sus retornos más leptocúrtica que la distribución de retornos de los mercados más desarrollados.

Palabras clave: Riesgo Financiero, Valor al Riesgo (VaR), Teoría de Valores Extremos (EVT), Administración de Proyectos de inversión.

1. Introducción:

La Teoría moderna de Portafolios, de acuerdo con Dixit y Pindyck (1995), Hull (2007), ha establecido los siguientes supuestos fundamentales sobre los que se ha desarrollado la teoría de los modelos estocásticos aplicada al mercado de capitales:

1) Las rentabilidades se comportan de acuerdo con un proceso estocástico.

2) Las variables aleatorias que integran dicho proceso se generan cuando se produce una modificación en el conjunto de información con que cuenta el mercado.

3) La incorporación de nueva información al mercado se verifica en forma continua y dicha información se incorpora en forma instantánea a los precios, es decir que el proceso estocástico representativo del comportamiento de las rentabilidades, es continuo en el dominio del tiempo.

4) Las rentabilidades de un activo son estocásticamente independientes.

5) Las variables aleatorias que representan el comportamiento de las rentabilidades tienen todas, la misma distribución de probabilidades, con varianza finita.

6) las rentabilidades se distribuyen de acuerdo con una función normal.

Ahora bien, la hipótesis clásica de normalidad de las distribuciones de las rentabilidades (de acuerdo con los postulados del teorema central del límite) implica el supuesto que los precios de los activos están formados por agregación de "shocks" aleatorios representativos de los factores que impulsan a los operadores a realizar estimaciones sobre el comportamiento de dichos precios y que estas estimaciones individuales constituyen variables aleatorias con varianza finita, independientes entre sí. Pero, diversos estudios realizados sobre los retornos de los activos financieros en mercados internacionales (Klein y Lederman (2000), Brigo y Mercurio (2001), Taleb (2004), Costantino y Brebbia (2012)) nos indican que los mismos suelen tener "colas" de distribución pesadas, o lo que es lo mismo, suelen presentar mayores probabilidades de ocurrencia de eventos riesgosos, por estos estudios podemos concluir que:

(1) Las rentabilidades suelen presentar cierto grado de asimetría (habitualmente negativa, lo cual implica una mayor probabilidad de caídas significativas que, de aumentos significativos en los precios del activo).

(2) Las rentabilidades muestran Kurtosis mayor a tres (recordemos que la Kurtosis influye sobre la asimetría).

(3) Las frecuencias contenidas en las "colas" de la distribución son más significativas que las que se deberían esperar en una función normal (efecto "fat tailed", que permite concluir que hay probabilidades marcadamente no nulas de que las rentabilidades, asuman valores alejados de cero).

(4) Estas frecuencias son uniformemente más significativas que las de la Normal (para desvíos de la variable con respecto a su valor medio superiores a 4 σ, se tiene la misma frecuencia que para desvíos superiores a 2 σ).

(5) Esta significatividad de las "colas" se mantiene en el tiempo, es decir que se observa un comportamiento similar para rentabilidades correspondientes a distintos períodos.

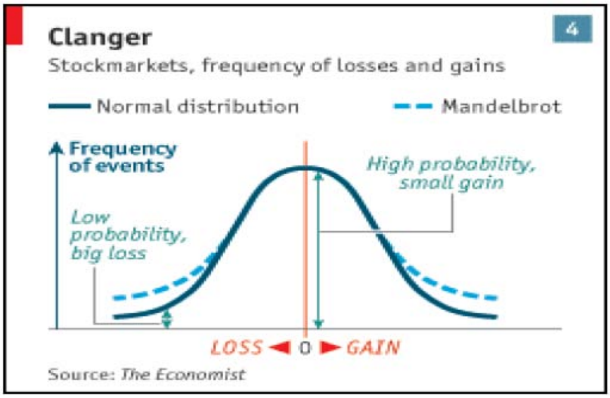

Figura (1): La ilustración siguiente, tomada de The Economist (2009) muestra que el principal problema es ignorar el riesgo de las colas.

Han surgido desde esta misma teoría moderna de portafolios, según lo señalado por Beirlant, Goegebeur, Teuglels, Segers, De Wall y Ferro (2005), distintos intentos de solución para esta problemática. Un intento de solución al comportamiento singular de asimetría y Kurtosis, se basa en la representación de las rentabilidades mediante transformaciones no-lineales de la función normal. Otro intento de solución, particularmente apropiado en activos que presentan saltos en las rentabilidades, a menudo precediendo a períodos de alta volatilidad, consiste en considerar la mezcla de distribuciones normales. También y como alternativas a la función normal para representar el comportamiento de las rentabilidades afectadas por asimetrías, Kurtosis altas y colas significativas, han aparecido las distribuciones no-normales. La distribución hiperbólica, la distribución t con α grados de libertad, la distribución logística y la distribución de Laplace, están entre las distribuciones no-normales más utilizadas. Finalmente podemos señalar que otra alternativa no-normal para representar el comportamiento de las rentabilidades es a partir de las distribuciones de valores extremos que consideran exclusivamente la distribución de las rentabilidades altas y de pérdidas excepcionales. Este enfoque motiva el presente trabajo de investigación, en el cual estudiaremos la caracterización de las "colas" de distribución pesadas en el contexto de la *Teoría de los valores extremos*, a partir de este marco teórico, podremos inferir medidas de riesgo adecuadas para caracterizar los retornos de activos en los mercados tradicionales y más precisamente en mercados emergentes o de países en vías de desarrollo, caracterizados principalmente por una distribución de sus retornos más leptocúrtica que la distribución de retornos de los mercados más desarrollados.

2. Riesgo en Mercados Financieros

En finanzas, suele entenderse el riesgo como la probabilidad de enfrentar pérdidas. Sin embargo, en sentido estricto debe entenderse como *la probabilidad de observar rendimientos distintos a los esperados*. Si se observan rendimientos extraordinariamente positivos o negativos, la probabilidad de enfrentar rendimientos distintos a los esperados en el futuro, es decir, el riesgo, crece. Si no se considera como una señal de alerta el observar rendimientos muy superiores a los esperados, se omite el análisis de las causas de tal desempeño extraordinario y, por lo tanto, se construyen las bases para enfrentar en el futuro pérdidas también extraordinarias. Dado que resulta imposible evitar por completo el riesgo, la necesidad de administrarlo es tácita. Por lo tanto, primero deben identificarse, en finanzas, todos los factores que pueden ocasionar la obtención de rendimientos distintos a los esperados, es decir, los factores de riesgo. Cada factor distinto define en sí mismo un tipo particular de riesgo, dentro de los cuales estamos interesados en este trabajo en los Riesgos Financieros.

3. El Valor en Riesgo:

Se ha establecido como el Valor en Riesgo, Value-at-Risk, o simplemente VaR como la medida estándar para cuantificar el riesgo de mercado (Comité de Basilea) y al parecer, la que se establecerá también para otros tipos de riesgos como el asegurador y el establecimiento de requerimientos de capital en el negocio del seguro. Podemos decir que la gestión del riesgo financiero surgió como una disciplina autónoma hacia el final de la década de los años 60 debido, principalmente, a la extraordinaria expansión de los mercados financieros. Como bien señala Bernstein (1996), el contexto devino en uno de alta volatilidad, exponiendo a los participantes del mercado a altos niveles de riesgo financiero e incentivando a las instituciones a buscar métodos precisos para lidiar con el riesgo de mercado. Recordemos que fue a mediados de 2004 cuando el Comité de Supervisión Bancaria de Basilea publicó el documento "Convergencia internacional de medidas y normas de capital. Marco revisado” (Basilea II), disponiendo nuevos criterios para la determinación del capital regulatorio de las entidades financieras. Un tiempo después, para la industria del seguro, el Acuerdo Solvencia II, con raíces en el modelo de Basilea II, buscó mejorar la seguridad en el sistema financiero al enfatizar los controles internos de las instituciones así como los modelos y procesos de administración de riesgos, utilizando principalmente modelos estadísticos elaborados con bases de datos históricas de las empresas, a efectos que cada entidad realice una cobertura de sus posibles pérdidas considerando la calidad histórica de su cartera. Analizando Basilea II podemos inferir que está organizado sobre la base de tres premisas fundamentales, la primera está referida a los requisitos mínimos de capital; la segunda respecto al proceso de revisión del supervisor y la tercera sobre disciplina de mercado. A diferencia de su antecesor, el Acuerdo de 1988 (Basilea I), Basilea II tiene una visión más general del tratamiento de los riesgos que toman las entidades y a la vez brinda mayor flexibilidad al permitir una variedad de enfoques para la medición del capital. A nivel local, en el año 2007 el Banco Central de la República Argentina (BCRA) dispuso la adopción del Enfoque Estandarizado Simplificado para riesgo crediticio, cuya implementación efectiva rige desde enero del año 2010, por lo que resulta de interés el estudio de las mejores prácticas para su exitosa puesta en funcionamiento. En ese marco, surge como valiosa la tarea del desarrollo de modelos que faciliten la interpretación de la dinámica de los riesgos en Argentina comparada con otros países, la proyección de variables y la estimación del impacto de políticas y regulaciones así como la elaboración de documentos de investigación empírica y teórica sobre esos temas vinculando la problemática local con la internacional. Tengamos en cuenta que a pesar de los defectos teóricos de esta metodología (señalados por Sarykalin, Serraino y Uryasev (2008)), el VaR posee una serie de características prácticas útiles, además, aún siendo de uso obligatorio, cada entidad tiene la libre elección del esquema a utilizar para su cálculo, resultando entonces que la estimación adecuada del VaR es materia de mucha importancia por una variedad de razones, que abarcan desde el establecimiento de los requerimientos del capital mínimo hasta su influencia en las decisiones de toma de riesgos. Pero, la mayoría de las técnicas concebidas para la cuantificación del VaR se basan en el supuesto que los retornos financieros siguen una distribución normal.

4. La Teoría de Valores Extremos:

La amplia literatura referida a la Estadística de los valores extremos, impide determinar con precisión los orígenes de esta teoría, podemos encontrar sus primeros indicios en Nicolás Bernoulli (1709) quién planteó el problema de la distancia media máxima desde el origen de "n" puntos distribuidos de forma aleatoria en una línea recta de distancia fija (t). También encontramos antecedentes importantes en los trabajos de Chappin, que estudia el problema de los efectos del tamaño en la resistencia de materiales, desde el punto de vista de valores extremos, los trabajos de Dodd, que estudia el problema de las distribuciones del máximo y del mínimo de una muestra y de Tippett, quién analiza el mismo problema pero para poblaciones normales. No obstante, como resultados centrales en la teoría de los valores extremos encontramos la demostración de Féchet, quien identificó una distribución limite posible para valores máximos y luego Fisher y Trippett, quienes demostraron que sólo es posible tres familias paramétricas de distribuciones límites para máximos y sus equivalentes para mínimos, y es a partir de las demostraciones formales de Féchet, Fisher, Trippett, que empiezan a proliferar los trabajos en este área y que no deja de crecer hasta nuestros días. De gran interés en el inicio de esta especialidad fueron también los trabajos de Finetti, Gumbel, Mises y Rice que abordaron el problema de la distribución de los extremos y que culminaron con la prueba en forma general, por Gnedenko, del teorema de los tipos de extremos. Podemos decir que una de las razones fundamentales del éxito y desarrollo logrado por la estadística de valores extremos es su relación con el diseño y los proyectos en ingeniería. En esta especialidad, el diseño viene siempre condicionado por los extremos, por ello, algunas especialidades de ingeniería como la meteorología, la ingeniería estructural, la hidráulica, la resistencia de materiales, la ingeniería eléctrica, la ingeniería del tráfico, etc. dependen inevitablemente de la estadística de los valores extremos. Esta teoría es aplicada desde hace varios años en hidrología como también por actuarios de la industria del seguro. Independientemente de que se esté tratando con movimientos de precios de mercado adversos, riesgo operativo, riesgo crediticio o riesgo de aseguramiento, uno de los mayores desafíos que intenta resolver esta teoría es implementar modelos que contemplen estos eventos y permitan la medición de sus consecuencias. En contraste, la teoría de los valores extremos (EVT) es una herramienta que trata de brindar una estimación de las colas de la distribución original haciendo uso solamente de los valores extremos de la serie de datos. Tomando como contexto teórico las ideas desarrolladas por McNeil, Frey y Embrechts (2005), en este trabajo utilizamos la técnica conocida como “picos sobre un umbral” (POT). Ahora bien, si queremos representar de una forma más genérica a las distribuciones de los valores extremos, podemos partir entonces de una determinada serie cronológica formada por rentabilidades observadas de un activo:

$$\{v_t\} \longleftarrow (t = 1,2,\ldots n) \qquad (1)$$

Con esta serie construimos (k) muestras no superpuestas de tamaño (nk) de la siguiente forma

$$n_k \left(n = k n_k\right), \quad v_1, v_2, v_3, \ldots, v_{n_k} \quad (2)$$

Suponemos además que (v_t) son realizaciones de las correspondientes nk variables aleatorias, además se cumple que:

$$v_{t,n_k}^{(\min)*} = \min\left(v_1, v_2, \ldots, v_n\right) \qquad (3)$$
$$v_{t,n_k}^{(\max)*} = \max\left(v_1, v_2, \ldots, v_n\right) \qquad (4)$$

Y tenemos las correspondientes variables estandarizadas que las obtenemos de restar las variables originales un coeficiente (λ) que representa su posición y de dividir por un coeficiente (δ) que representa su factor de escala:

$$v_{t,n_k}^{(\min)*} = \frac{v_{t,n_k}^{(\min)} - \lambda}{\delta} \qquad (5)$$
$$v_{t,n_k}^{(\max)*} = \frac{v_{t,n_k}^{(\max)} - \lambda}{\delta} \qquad (6)$$

En el límite cuando (nk) tiende a infinito, estas variables representan las rentabilidades extremas estandarizadas de la forma:

$$v_t^{(\min)*} = \lim_{n_k \longrightarrow \infty} v_{t,n_k}^{(\min)*} \qquad (7)$$

$$v_t^{(\max)*} = \lim_{n_k \longrightarrow \infty} v_{t,n_k}^{(\max)*} \qquad (8)$$

a. Admiten tres tipos de distribuciones definidas de la siguiente forma:

$$\begin{aligned} &Tipo1: \quad F_{v_t*}(v) = \exp\left(-e^{-v}\right) \\ &Tipo2: \quad F_{v_t*}(v) = e^{-v^{-\alpha}}; \quad (v \geq \alpha) \qquad (9) \\ &Tipo3: \quad F_{v_t*}(v) = e^{-v^{\alpha}}; \quad (v \leq \alpha) \end{aligned}$$

Estos tipos de funciones que aparecen en la expresión (9), dan origen a distintas familias de distribuciones tal como lo son las distribuciones de Gumbel, Weibull, Fréchet y de Pareto-Lévy. Específicamente en nuestro caso prestaremos fundamental atención a la distribución generalizada de Pareto (DGP).

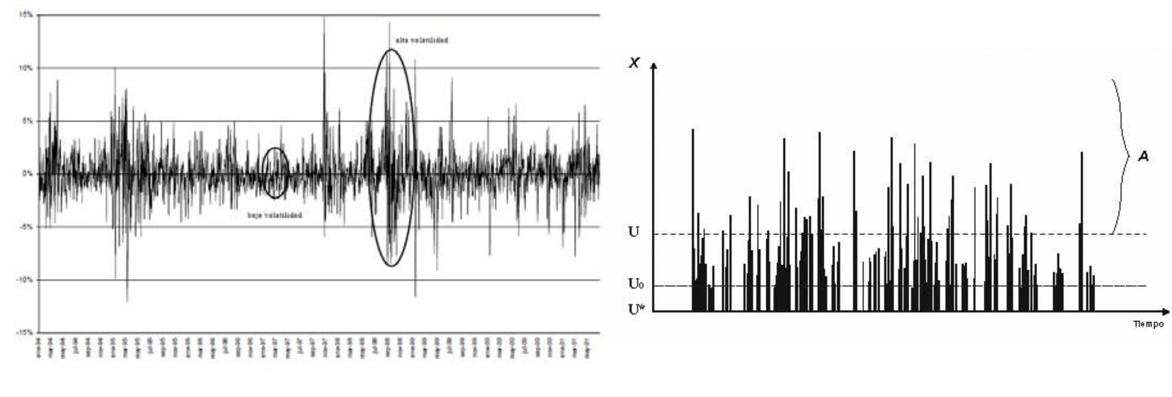

Figura (2) Distribución de Valores Extremos de Retornos financieros según el criterio de “picos sobre un umbral” (POT).

5. Definición de Valor en Riesgo (VaR):

En el contexto de este trabajo el término “riesgo” puede interpretarse como la pérdida potencial que puede sufrir el valor de un activo, mientras que la medida del riesgo está vinculada a la probabilidad de sufrir esa pérdida. La “administración del riesgo” hace referencia a un conjunto de métodos y procedimientos destinados a (i) identificar (y clasificar cualitativamente) los riesgos, (ii) medirlos (cuantificar) y (iii) controlarlos (eliminándolos o disminuyéndolos). El denominado *Value at Risk* o Valor en Riesgo, $VaR(\alpha;\Delta_t)$, resulta ser una medida probabilística del riesgo de mercado que permite estimar la pérdida máxima que puede sufrir el valor de un activo, en un intervalo de tiempo, $\mathbf{\Delta_t}$, especificado y con un nivel de confianza (probabilidad), 1-$\boldsymbol{\alpha}$, dado. En general el VaR de los retornos, pronosticado en el momento “t” para el momento”t+1”, se suele calcular mediante la fórmula siguiente:

$$VaR_{t+1}(\alpha\Delta t) = \sigma_{t+1} * F^{-1}(\alpha)$$

σ_{t+1}: la predicción de volatilidad obtenida de algún modelo, ajustada al período Δt.

$F^{-1}(\alpha)$: la inversa de la función de probabilidad acumulativa (función de distribución) de la variable aleatoria retorno estandarizada. En unidades monetarias tenemos entonces, la siguiente fórmula general:

$$VaR_{t+1}(\alpha\Delta t) = P_t * \sigma_{t+1} * F^{-1}(\alpha)$$

P_t: el precio del activo en el tiempo (t).

Si además, al supuesto (que se verifica empíricamente con bastante aproximación para datos diarios) siguiente:

$$E\left(r_{t+\Delta t}\right)=0$$

Agregamos el supuesto de la normalidad de los retornos $r_{t+\Delta t}$, nos queda entonces la siguiente fórmula para la estimación del VaR, cuando ***z*** es la correspondiente variable aleatoria estandarizada

$$VaR_{t+1}\left(\alpha\Delta t\right)=z_{\alpha} * P_t * \sigma_{t+1} * \sqrt{\Delta t}$$

Donde el período es Δt y el nivel de confianza es:

$$1-\alpha\left(y_{\alpha}=z_{\alpha}=\Phi^{-1}(\alpha)\right)$$

Debido a que el VaR, en su forma más simple, se reduce sencillamente al producto de tres grandes factores: El valor actual del activo o portafolio, marcado a mercado; el factor de riesgo, ajustado por el horizonte de tiempo deseado; y el factor de confianza definido. Entonces el resultado final es sensible, básicamente, a variaciones del valor del portafolio o del activo, del horizonte de tiempo, del grado de confianza y del factor de riesgo, que depende del activo cuyo VaR se desea estimar. Los modelos clásicos de estimación de riesgo financiero, trabajan sobre los cuantiles de las variables aleatorias utilizando información de todo el conjunto de datos. Sin embargo, debido al hecho que los cuantiles al 1% o 5% son valores extremos de la distribución, resulta natural modelar las colas directamente en lugar de considerar la estructura completa de la distribución. La Teoría de Valores Extremos (EVT) provee una justificación teórica para tales procedimientos, ya que desempeña un rol similar al del Teorema Central del Límite en los modelos de variables aleatorias que se distribuyen en forma normal. En términos generales, podemos decir que, los valores extremos pueden ser modelados siguiendo dos procedimientos básicos: a) Los modelos denominados Block Maxima Models (BMM), que emplean las distribuciones generalizadas de valores extremos (GEV) para ajustar una distribución a partir de los máximos o mínimos de un conjunto de datos muestrales agrupados en bloques de similar tamaño, o b) Los modelos denominados modelos Peaks over Thresholds (POT), que utilizan las distribuciones generalizadas de Pareto (GPD) para ajustar una distribución a los valores muestrales que exceden un umbral especificado. El primero tiene el inconveniente de retener para el análisis solamente el valor máximo (o el mínimo) de un gran bloque de datos. En consecuencia, evitar el proceso de agrupamiento en bloques muy grandes deriva en una mejor utilización de los datos. Además, en mercados con series de datos de extensión relativamente corta, es crucial el no tomar en cuenta algunos datos intermedios contenidos en el bloque, ya que podrían llegar a ser extremos si los bloques fueran de diferente tamaño. Estas son algunas de las razones por la cuales en retornos financieros se suele utilizan más frecuentemente los modelos del tipo POT. Veamos a continuación algunas definiciones teóricas sobre este tipo de modelos: Sea una sucesión de variables aleatorias independientes e idénticamente distribuidas (iid) X_1, X_2, … , X_n que tienen la función de distribución $F(y) = Pr(X<y)$ desconocida; en esas condiciones los eventos extremos serán definidos como aquellos valores de X_t que exceden algún valor u, alto y predeterminado. La variable $X – u$ representa los POT (Excesos Sobre el Umbral u). Solamente haremos referencia a umbrales positivos ($u > 0$), pues los resultados para $u < 0$ emplean un razonamiento similar. La relación entre la función de distribución F_u de los citados POT y F, la de la variable aleatoria subyacente, está determinada por la probabilidad condicional:

$$\Pr\{X - u \le y | X > u\} = \frac{F(y+u) - F(u)}{1 - F(u)}$$

$F_u(y)$ puede interpretarse como la probabilidad de que una pérdida exceda el umbral u por un monto igual o menor a y, supuesto que el umbral u ha sido excedido. Esta fórmula muestra que si fuera conocida la distribución original F, sería también conocida la distribución de los excedentes F_u; sin embargo, en las aplicaciones reales sucede lo contrario, obligando entonces a estimar en base a los datos muestrales, la distribución desconocida F para valores altos por encima del umbral. Veremos a continuación una resumida descripción de los fundamentos teóricos básicos que permiten tal estimación. Si se interpreta “valor extremo” como la máxima pérdida en un cierto intervalo de tiempo, entonces puede se lo puede definir como el máximo (o mínimo) de una muestra: $\{x_1, x_2, \ldots, x_n\}$. Se trata entonces de determinar la función de distribución H de esos valores extremos. Suponiendo que ésos son valores de una secuencia: $X_1, X_2, \ldots, X_n$ de variables aleatorias iid con función de distribución común F, se designa:

$$M_n = \max\{X_1; X_2; \ldots; X_n\}$$

y si F cumple con el teorema de Fisher y Tippet, puede probarse que, para valores grandes de n,

$$\Pr\{M_n \leq z\} \approx H(z)$$

La expresión

$$H_\xi(z) = \begin{cases} \exp\left\{-\left[1+\xi\left(\dfrac{z-u}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} & si\,\xi \neq 0;\quad 1+\xi\left(\dfrac{z-u}{\sigma}\right) > 0 \\ \exp\left[-e^{\left(\frac{z-u}{\sigma}\right)}\right] & si\,\xi = 0 \end{cases}$$

La expresión anterior es la fórmula de la función de distribución límite de M_n, conocida como “Distribución Generalizada de Valores Extremos” (GEV). H_ξ depende de tres parámetros: μ (locación), $\sigma > 0$ (escala) y ξ (índice de cola o de forma). El dominio de z está restringido por la condición

$$1+\xi[(z-u)/\sigma] > 0$$

Dado que a los efectos de este trabajo interesa estimar F_u, la función de distribución de los excesos sobre el umbral u, se enuncia a grandes rasgos un teorema que vincula las distribuciones H_ξ y F_u: Para un amplio conjunto de funciones de distribución subyacentes F, existe un valor σ que depende de u,

$$F_u(y) \approx G_{\xi,\sigma}(y)$$

tal que:

sí y sólo si F pertenece al máximo dominio de atracción de la distribución generalizada de valores extremos(GEV) H_ξ , donde $G_{\xi,\sigma}$ representa la “Distribución Generalizada de Pareto” (GPD) dependiente de los parámetros ξ y σ, cuya expresión es:

$$G_{\xi,\sigma}(y) = \begin{cases} 1-(1+\xi y/\sigma)^{-1/\xi} & ; Si\, \xi \neq 0 \\ 1-\exp(-y/\sigma) & ; Si\, \xi = 0 \end{cases}$$

Al igual que para las distribuciones GEV, mayores valores del índice de cola ξ implican un incremento en el tamaño de las colas. Existen tres tipos de distribuciones pertenecientes a la familia GPD, de acuerdo con el valor del índice de cola ξ:

Si $\xi > 0$ entonces $G_{\xi,\sigma}$ es la distribución de Pareto clásica.

Si $\xi = 0$ entonces $G_{\xi,\sigma}$ es la distribución exponencial.

Si $\xi < 0$ entonces $G_{\xi,\sigma}$ es la distribución de Pareto de tipo II de colas cortas.

Ellas se corresponden con las distribuciones de valores extremos GEV de Fréchet, Gumbel y Weibull respectivamente. La precisión de la familia GPD puede ser mejorada si se le agrega un parámetro de localización μ, convirtiéndose entonces en $G_{\xi,\sigma}(y - \mu)$. En sentido amplio, los teoremas anteriores significan que si la distribución de los máximos M_n de los bloques de datos es aproximadamente descripta por H, entonces la distribución aproximada de los excesos sobre el umbral está dada por $G_{\xi,\sigma}$ y en consecuencia pertenece a la familia Generalizada de Pareto (GPD). En forma adicional, los parámetros de la fórmula del modelo GPD están en forma unívoca determinados por los de la correspondiente distribución GEV asociada de los máximos de los bloques muestrales. Una vez que se ha realizado la elección del umbral u y se han estimado los parámetros del modelo GPD, es necesario obtener la expresión para determinar el cuantil relevante que permitirá calcular el VaR. Así, teniendo presente que x = y + u, una estimación de F(x), para x > u podría ser:

$$F(x) = [1 - F(u)]G_{\xi,\sigma}(y) + F(u)$$

Considerando k como el número de observaciones mayores al umbral u, F(u) puede ser fácilmente aproximada de manera no-paramétrica por medio del simple estimador empírico:

$$\hat{F}(u) = \frac{n-k}{n}$$

Introduciendo este estimador en la fórmula de $G_{\xi,\sigma}$ es posible encontrar una estimación de F(x):

$$\hat{F}(x) = 1 - \frac{k}{n}\left[1 + \hat{\xi}\frac{(x-u)}{\hat{\sigma}}\right]^{-\frac{1}{\hat{\xi}}}; \quad Si \longrightarrow \xi \neq 0$$

o bien:

$$\hat{F}(x) = 1 - \frac{k}{n} e^{-\left(\frac{x-u}{\hat{\sigma}}\right)}; \quad Si \longrightarrow \xi = 0$$

Para un nivel de significación $\alpha > F(u)$ o, lo un nivel de confianza $1-\alpha < 1-F(u)$,

La expresión del VaR se puede computar determinando la función inversa de la estimación de F(x)

y resolviendo para x en las siguientes fórmulas:

$$VaR_{t+1}(\alpha) = u - \sigma \ln\left[\frac{n}{k}(1-\alpha)\right]; \quad si \longrightarrow \xi = 0$$

$$VaR_{t+1}(\alpha) = u + \frac{\hat{\sigma}}{\hat{\xi}}\left[\left(\frac{1-\alpha}{k/n}\right)^{-\hat{\xi}} - 1\right]; \quad si \longrightarrow \xi \neq 0$$

6.1. Identificación del umbral y ajuste del modelo GPD:

Como primera premisa a considerar, podemos decir que la selección del umbral es el punto débil de este tipo de técnica (método POT) y aún cuando hay numerosas maneras basadas en fórmulas para escoger el mencionado umbral, no existe un fundamento teórico que sustente firmemente un procedimiento satisfactorio para este objetivo. Es por ello que nos vemos obligados a utilizar una serie de procedimientos ad-hoc para poder determinar u como el punto de partida para la estimación de los parámetros del modelo GPD que ajuste razonablemente a los datos y permita un adecuado pronóstico del VaR. En ese sentido se debe tener en cuenta la necesidad de lograr un equilibrio entre la varianza y el sesgo. En efecto, si el umbral es muy bajo resultará afectada la naturaleza asintótica del modelo, dado que h fluctuará considerablemente para diferentes valores del estadístico de alto orden k, originando un mayor sesgo. Por el contrario, cuanto más elevado sea umbral u, habrá menos valores excedentes, lo que originará alta varianza, consecuentemente, distintos autores han desarrollado sus propias técnicas para identificar el umbral, Sarykalin, Serraino y Uryasev (2008), sugieren una regla práctica consistente en considerar valores extremos sólo aquellas observaciones situadas en el decil superior o en el inferior de la distribución.

Finalmente podemos decir que la elección del umbral "u" es un compromiso entre elegir un valor lo suficientemente elevado como para que el teorema asintótico pueda considerarse exacto, y lo suficientemente bajo como para tener suficiente datos para estimar los parámetros.

6.2. La distribución incondicional:

Cualquier probabilidad condicional puede ser escrita en términos de una probabilidad incondicional de la siguiente forma:

$$F_u(y) = P(X - u \leq y / X \geq u) = \frac{F(y+u) - F(u)}{1 - F(u)}$$

Si ahora reemplazamos (x = u + y) además si combinamos esta igualdad con la expresión obtenida anteriormente:

$$F_u(y) \cong G_{\xi,\beta(u)}(y)$$

se obtiene de esta forma:

$$F(x) = (1 - F(u))G_{\xi,\beta}(x-u) + F(u), \quad para\, x > u$$

Lo único que resta es un estimador empírico de F(x). El candidato natural es: (n-Nu)/n con lo que obtenemos el siguiente estimador de la cola de distribución:

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

A partir de esta expresión es posible deducir una fórmula para el cálculo del VaR.

Dada una probabilidad e invirtiendo la fórmula anterior tenemos:

$$VaR_q^{(EVT)} = u + \frac{\beta}{\xi}\left(\frac{n}{N_u}(1-q)^{-\xi} - 1\right)$$

6.3. La distribución condicional:

El VaR puede estar basado tanto en la distribución incondicional como en la distribución condicional de los retornos. En el primer caso se trata del VaR incondicional o estático, que es útil para evaluar el riesgo de un activo durante períodos de tiempo prolongados y que se puede computar mediante la expresión anterior. El VaR condicional o dinámico, en tanto, es un percentil de la distribución condicional de los retornos y se utiliza para evaluar el riesgo en el corto plazo. Una de las ventajas teóricas por sobre el estático es que tiene en cuenta que lo que es un valor extremo varía según se trate de un período de alta o baja volatilidad. Por ejemplo, en el momento (t) se realiza un pronóstico del VaR para (t+1) teniendo en cuenta la distribución de probabilidad de los retornos condicional al momento (t). Esto contrasta con el VaR estático, en donde la distribución dada por la expresión es incondicional respecto de la historia de la serie hasta t.

6.4. Características del Proceso Estocástico:

Tengamos en cuenta que a la variable aleatoria Xt se le ajusta un modelo AR(1)-GARCH(1,1) de la forma:

$$X_t = \Phi X_{t-1} + \sigma_t Z_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Siendo Z_t la innovación iid del proceso y además cumpliéndose lo siguiente:

$$\begin{aligned} &\Phi < 1; \quad \alpha_0, \alpha_1, \beta > 0 \\ &|\alpha_j| < 1; \quad \alpha_1 + \beta < 1 \\ &\varepsilon_{t-1}^2 = \left(X_{t-1} - \Phi X_{t-2}\right)^2 \end{aligned}$$

En la figura (3), se visualiza, a modo de ejemplo la serie de retornos del índice merval, en donde se observa el efecto de "clustering" en sus retornos, es decir, períodos de alta volatilidad seguidos por períodos de baja volatilidad. Esto es lo que motiva la introducción del componente GARCH, para modelar el proceso de la varianza condicional de los retornos.

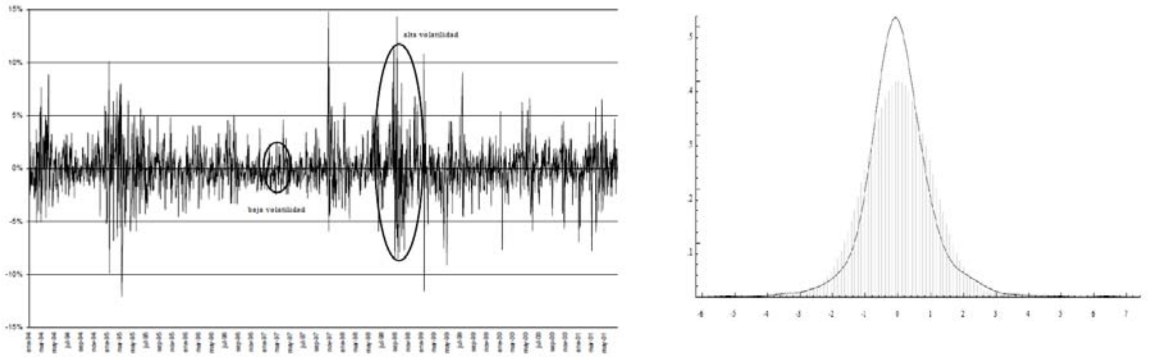

Figura (3): se notan ciertos hechos estilizados de las distribuciones de los retornos de los activos financieros; leptocurtosis mayor que la normal, fat tails (colas gordas) y thin waist (cintura angosta).

La distribución incondicional de los retornos Fx(x) que es en la cual se basa el VaR estático, se puede aproximar a través de un histograma con los retornos de toda la muestra, en el gráfico siguiente, se supone una variable aleatoria con distribución N(0,1) una estimación no paramétrica de la densidad de distribución incondicional de los retornos, estandarizados por la volatilidad incondicional o muestral. Allí se notan ciertos hechos estilizados de las distribuciones de los retornos de de los activos financieros, leptocurtosis mayor que la normal, *fat tails* (colas gordas, mayor probabilidad de ocurrencia de eventos extremos) y *thin waist* (cintura angosta, menor probabilidad de eventos cercanos a la media). El enfoque GARCH convencional, de acuerdo con Landro (2009), para a estimación del VaR dinámico asume que Z_t es una innovación gaussiana independiente de X_t e I_{t-1}, es decir que la distribución $F_z(z)$ es normal. Si bien las innovaciones tienen una distribución iid N(0,1) ellas son multiplicadas por la volatilidad condicional X_t del momento ($\varepsilon_t=\sigma_t Z_t$). En consecuencia, la distribución incondicional de ε_t y luego la de X_t está constituida por muchas distribuciones iid N(0,1) desplazadas alrededor de la media condicional (ΦX_{t-1}) según la volatilidad condicional, obteniéndose distribuciones incondicionales leptocúrticas. Allí se nota claramente la no normalidad del error, a pesar de la normalidad de la innovación. Tengamos en cuenta que, emplear innovaciones gaussianas implica suponer que la distribución condicional de los retornos también es gaussiana, lo que lleva a una subestimación del VaR ya que empíricamente su distribución condicional aparenta tener colas más pesadas que la normal, por lo que hay mayor probabilidad de ocurrencia de pérdidas más grandes. Como se puede inferir de los razonamientos vistos anteriormente, los residuos estandarizados, que en teoría deberían distribuirse de una manera consistente con los supuestos formulados acerca de Z_t, no aparentan surgir de un proceso con innovaciones distribuidas idéntica e independientemente de acuerdo a una normal estándar. A diferencia del enfoque convencional en el que se supone que la innovación tiene una distribución normal, redundando en retornos con una distribución normal, aquí suponemos que Z_t tiene una distribución $F_z(z)$ desconocida con media cero y varianza unitaria, por lo que la estimación del modelo se realiza por "máxima verosimilitud". La modificación del supuesto distributivo de la innovación está motivada por la observación de que aún (controlando) por la volatilidad condicional de X_t, los retornos presentan colas gordas. Esto lleva a pensar que otra fuente de leptocurtosis en los retornos puede venir dada por la distribución de las innovaciones, en el sentido que shocks o pérdidas más grandes tienen una mayor probabilidad de ocurrencia que la que está implícita en una distribución normal, justificando el uso de distribuciones con colas más gordas para no subestimar el riesgo. En síntesis, la distribución incondicional de los retornos puede tener colas gordas (i) con innovaciones gaussianas pero con volatilidad estocástica, y/o (ii) por el hecho de que las innovaciones tengan una distribución con colas gordas, que se refleja también en la distribución condicional de los retornos. En consecuencia, para mejorar el VaR dinámico y siguiendo a McNeil, Frey y Embrechts (2005), aquí no se hacen supuestos distributivos sobre $F_z(z)$, en tanto que se aplica EVT a la cola de la distribución de los residuos estandarizados del modelo, estimando una distribución Generalizada de Pareto. Por definición, la distribución condicional de los retornos es:

$$F_{x_{t+1}|I_t}(x) = P\left(X_{t+1} \leq x | I_t\right)$$

Como se supuso y de acuerdo con Huang y Litzenberger (1995), un proceso AR(1) - GARCH (1,1), se remplaza X_{t+1} por su pronóstico, con lo que se obtiene una estimación de la distribución condicional:

$$\hat{F}_{x_{t+1}|I_t}(x) = P\left(\hat{\mu}_{t+1} + \hat{\sigma}_{t+1} z_{t+1} \leq x | I_t\right)$$

$$\hat{F}_{x_{t+1}|I_t}(x) = P\left(z_{t+1} \leq \frac{x - \hat{\mu}_{t+1}}{\hat{\sigma}_{t+1}} | I_t\right)$$

Luego, como Z_t es una innovación y en consecuencia, independiente de I_t, tenemos:

$$\hat{F}_{x_{t+1}|I_t}(x) = F_z\left(z_{t+1} \leq \frac{x - \hat{\mu}_{t+1}}{\hat{\sigma}_{t+1}}\right)$$

Si q es el nivel de confianza deseado, como por ejemplo 99% tenemos entonces:

$$q = F_z\left(z_{t+1} \leq \frac{x_q - \hat{\mu}_{t+1}}{\hat{\sigma}_{t+1}} \right)$$

$$q = F_z\left(z_{t+1} \leq z_q\right) \xrightarrow{ademas} z_q = \frac{x_q - \hat{\mu}_{t+1}}{\hat{\sigma}_{t+1}}$$

Por último tenemos entonces:

$$z_q\,\hat{\sigma}_{t+1} + \hat{\mu}_{t+1} = x_q \quad ó \longrightarrow \quad VaR_{t+1}^{evt} = VaR(Z)\,\hat{\sigma}_{t+1} + \hat{\mu}_{t+1}$$

Donde Var(Z) es un cuantil superior de la distribución marginal de Zt, por lo que no depende del tiempo.

7. Breve descripción del Método Estático para la determinación de la VaR:

El método estático consiste en determinar para cada serie un umbral "*u*" apropiado (generalmente se adopta un valor que deje por debajo entre el 5% y 8% de las pérdidas extremas) y en estimar los parámetros β y ξ a partir de los N_u datos remanentes, usando máxima verosimilitud u otra técnica apropiada. Una vez que calculamos estos parámetros se puede determinar la VaR a través de la fórmula (7).

$$Var_q^{(EVT)} = u + \frac{\beta}{\xi}\left(\frac{n}{N_u}(1-q)^{-\xi} - 1 \right)$$

Tengamos en cuenta que este cuantil permanecerá estable durante largos períodos, pues en principio, se requeriría de muchos datos adicionales a efectos de agregar suficientes valores extremos nuevos como para cambiar la estimación de los parámetros. Debido a que no necesitamos calcular permanentemente el cuantil es que se conoce a esta metodología como método estático (en general se vuelve a calcular el cuantil pasado el año si que sean observados cambios notables en el VaR).

8. Breve descripción del Método Dinámico para la determinación de la VaR:

El concepto de método dinámico es distinto al concepto de método estático, ya que en el primer caso el VaR se basa en la distribución condicional de los retornos y debe ser nuevamente calculado en forma diaria, con el consiguiente mayor esfuerzo computacional, mientras que el VaR estático se basa en la distribución incondicional y el nuevo cálculo diario no es necesario. En el primer caso, la mayor complejidad tiene como contrapartida un ahorro en el capital económico necesario para cubrir las pérdidas, obteniéndose a pesar de ello un nivel de cobertura similar al del VaR incondicional. Como se explicó anteriormente, para capturar la dinámica de la media y la varianza condicional se estima un modelo AR(1) - GARCH (1,1) sin hacer ningún supuesto sobre las innovaciones más que sean iid de acuerdo a una Fz(z) con media cero y varianza unitaria. Como los residuos estandarizados deberían asemejarse a las innovaciones, para estimar la cola de Fz(z) (que es lo que interesa para conocer el VaR dinámico) se aplicará la técnica de EVT desarrollada en la parte estática a la cola de la distribución empírica de residuos. Finalmente al contar con una estimación paramétrica de la cola de Fz(z) y con las estimaciones de la media y la volatilidad condicional, se puede estimar la distribución condicional de las pérdidas. Siguiendo la metodología desarrollada por McNeil, Frey y Embrechts (2005), vemos a continuación el desempeño del VaR estático y el VaR condicional o dinámico, estimando los modelos auto-regresivos de volatilidad estocástica del tipo AR-GARCH y aplicando EVT para obtener una mejor estimación de la probabilidad de ocurrencia de las pérdidas extremas. En la figura (4) según Definer y Girault (2007), observamos la serie de retornos (acciones de Acindar), donde encontramos períodos de alta volatilidad seguidos de períodos de baja volatilidad, esto es lo que motiva la introducción del componente GARCH para modelar el proceso de la varianza condicional de los retornos.

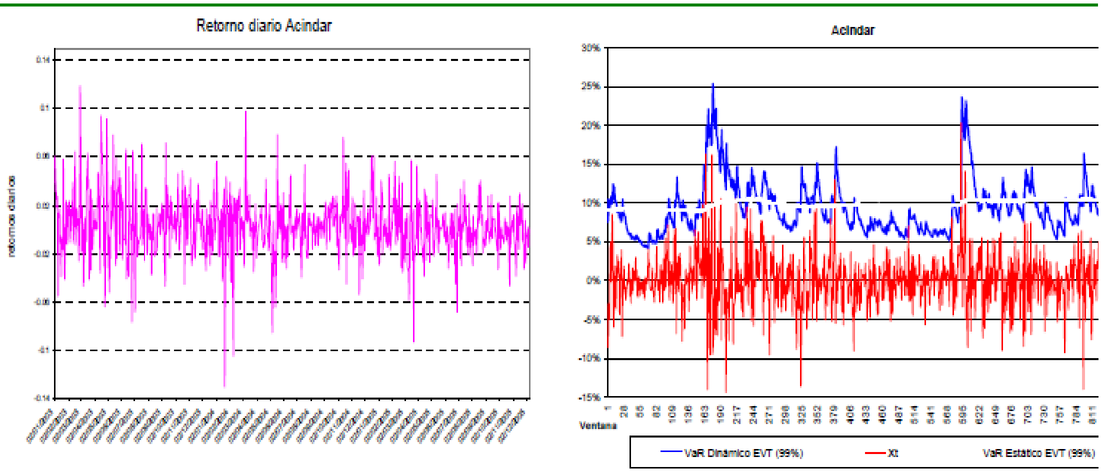

Figura (4): Desempeño del VaR y del VaR condicional para una serie de retornos financieros.

9. Conclusiones y trabajos futuros:

El Value at Risk (VaR), medida de desempeño ampliamente usada responde a la pregunta: ¿cuál es la máxima pérdida que un portafolio puede enfrentar en un horizonte de tiempo definido y con una probabilidad dada? Por ejemplo, el VaR a 95% es una estimación de la máxima pérdida que un portafolio enfrentará en un horizonte definido de tiempo y que es excedida únicamente el 5% de las veces. El VaR puede ser fácilmente calculado cuando los factores de riesgo subyacentes se distribuyen normalmente. Sin embargo, para distribuciones no normales, el VaR puede tener propiedades indeseables tales como falta de subaditividad, es decir, el VaR de un portafolio de dos instrumentos puede ser mayor que la suma de los VaR individuales. Además, el VaR es difícil de optimizar para distribuciones discretas cuando se calcula usando escenarios. En este caso, el VaR es no convexo, no es continuo con respecto a las posiciones y tiene extremos locales múltiples. A pesar de que la administración del riesgo en base al VaR es ampliamente usada y de que existen significativos esfuerzos de investigación en el área, según autores como Galitz (1995), Klein y Lederman (2000), aún no se disponen de algoritmos de optimización eficientes, es por ello que el uso de métodos alternativos, como EVT, proveen medidas de VaR más apropiadas (McNeil, Frey y Embrechts (2005)). La teoría de los valores extremos (ETV) es una herramienta que trata de brindar una estimación de las colas de la distribución original haciendo uso solamente de los valores extremos de la serie de datos. La técnica estática de EVT, muy económica en cuanto al esfuerzo computacional, produce mejores resultados que el cálculo VaR con una distribución normal. Usar el VaR mediante EVT permitiría reducir el sesgo que se origina en asumir una distribución normal y por ende, estar más próximo del verdadero valor al riesgo ante la presencia de eventos extremos. Una medida de pérdidas alternativa, con propiedades más atractivas es el Condicional Value at Risk (CVaR). El CVaR es una medida de riesgo más consistente. El CVaR es consistente con el enfoque media-varianza: para distribuciones de pérdida normales, portafolios óptimos en media-varianza también son CVaR óptimos. Según Sarykalin, Serraino y Uryasev (2008), para distribuciones arbitrarias, experimentos numéricos indican que generalmente la minimización del CVaR también conduce a soluciones cuasi óptimas en términos de VaR ya que el VaR nunca excede el CVaR. Sin embargo, para distribuciones muy sesgadas, los portafolios que minimizan el CVaR pueden diferir bastante de los VaR óptimos. Recordemos que la técnica dinámica, desarrollada por McNeil, Frey y Embrechts (2005), se basa en la distribución condicional de los retornos y combina la estimación de modelos GARCH con la utilización de EVT para estimar de manera diaria un VaR condicional (basado en la distribución condicional de los retornos o las pérdidas). Si bien el desempeño del método dinámico no difiere demasiado respecto al del método estático, el cálculo del VaR es mucho más sensible respecto a los retornos observados y la evolución temporal del VaR dinámico con EVT actúa como una envolvente de los mismos, tal como lo señalaron Definer y Girault (2007). Esto produce claras ventajas debido a una mejor administración del capital inmovilizado en relación al método estático. En síntesis, la elección de la metodología depende de la aplicación que se le termine dando. Por ejemplo, un determinado ente regulador probablemente elegirá el método estático, pues es de simple implementación y fácil supervisión, además de garantizar un razonable nivel de cobertura en todo el período que se desea cubrir. En contraste un "trading desk" probablemente esté dispuesto a un mayor esfuerzo computacional de implementar la metodología dinámica a fin de no inmovilizar tanto capital. La aplicación de la teoría de los Eventos Extremos ha resultado ser de gran utilidad para mejorar la administración del riesgo financiero, en cualquier proyecto de inversión. Como trabajos futuros, nos queda indagar aún más sobre las bondades y robustez de estos métodos (estático y dinámico) en distintas series de retornos financieros provenientes tanto de mercados locales (ROFEX, por ejemplo) como de mercados emergentes. También nos resta indagar si efectivamente esta metodología converge a la caracterización clásica de Markowitz para la selección de portafolios con retornos esperados que se distribuyen normalmente y además resulta presentar mejor desempeño ante la existencia de eventos extremos y/o catastróficos.

10. Agradecimientos: El autor quiere expresar su agradecimiento al profesor Alberto Landro (UCA) por sus valiosos comentarios referidos a las técnicas econométricas aplicadas a los mercados financieros.

11. Referencias Bibliográficas:

- Beirlant J., Goegebeur Y., Teuglels J., Segers J., De Wall D. and Ferro C., (2005). *Statistics of Extremes, Theory and Application.* New York, United States of America: John Wiley.
- Bernstein, P. L. (1996). *Against the Gods: The Remarkable Story of Risk.* New York, United States of America: John Wiley.
- Brigo D. and Mercurio F. (2001). *Interest Rate Models Theory and Practice, with Smile Inflation and Credit.* Toronto, Canada: Springer Finance.
- Costantino M. and Brebbia C. A. (2012). *Computational Finance and it Applications.* Boston, United States of America: WIT Press.
- Definer and Girault (2007). *Aplicación de la teoría de valores extremos al gerenciamiento del riesgo,* Buenos Aires, Argentina: Ediciones Banco Central de la República Argentina.
- Dixit A. K. and Pindyck R. S. (1995). *Investment under Uncertainty.* Princeton, New Jersey, United States of America: Princeton University Press.
- Galitz L. C. (1995). *Financial Engineering. Tools and Techniques to Manage Financial Risk,* I London, UK: Pitman Publishing Company.
- Huang C. F. and Litzenberger R. H. (1995). *Foundations for Financial Economics.* New Jersey, United States of America: Prentice-Hall.

- Hull, J. C (2007). *Options, Futures and Other Derivatives,* New Jersey, United States of America: Fourth Edition, Prentice-Hall.
- Klein R. A. and Lederman J. (2000). *Derivatives Risk and Responsibility, the complete Guide to effective Derivatives Management and Decision Making.* New York, United States of America: Irwin Professional Publishing.
- Landro A. (2009). *Elementos de Econometría de los fenómenos dinámicos*. Buenos Aires, Argentina: Ediciones Cooperativas de la Universidad de Buenos Aires.
- McNeil A. J. Frey R. and Embrechts P. (2005). *Quantitative Risk Management, Concepts, techniques and tools,* Princeton, New Jersey, United States of America: Princeton Series in Finance.
- Sarykalin S., Serraino G. and Uryasev S. (2008). *Value at Risk vs. Conditional Value at Risk in Risk Management and Optimization*. Florida, United States of America: Ed. Informs, Tutorial in Operations Research.
- Taleb N (2004). *Treasury Dynamic Hedging , Managing Vanilla and Exotic Options*. New York, United States of America: Second Edition, John Wiley.