

# UNA PROPUESTA PARA EVALUAR PRONÓSTICOS DE RENDIMIENTOS DE ACCIONES CUANDO LAS DISTRIBUCIONES EMPÍRICAS NO SON NORMALES ESTACIONARIAS\*

**José Carlos Ramírez**

*Instituto Tecnológico y de Estudios Superiores de Monterrey*

**Rogelio Sandoval Saavedra**

*Secretaría de Hacienda y Crédito Público*

*Resumen:* Se analizan los principales problemas estadísticos relacionados con el pronóstico de acciones cuando la distribución de sus series de rendimientos no son normales estacionarias. La exposición comprende una aplicación pormenorizada de las pruebas de normalidad para las acciones que forman el índice *Banamex-30*, así como el empleo de algunas mezclas de distribuciones y de procesos estocásticos para las series que no son normales estacionarias. Como método alternativo de estudio, se propone un modelo de Markov de segundo orden dirigido a obtener pronósticos cuya elaboración no descansa sobre el supuesto de normalidad.

*Abstract:* This paper deals with the main problems related to predictability of asset returns when data series are not normally stationary distributed. The statistical analysis includes several normality tests on returns series of *Banamex-30* stocks first, and then an application of mixture of probability distributions and stochastic processes to series, which are not normal stationary. As a means to avoid the normality assumption when forecasting asset returns, we introduce a second-order Markov model.

*Clasificación JEL: C52, G11*

*Fecha de recepción: 7 V 2002*

*Fecha de aceptación: 20 IX 2002*

---

\* Agradecemos la colaboración de Carlos Vallebuena, Alicia Rodríguez y Héctor Hernández de la Dirección de Riesgos de Banamex-Santa Fe [jcsanch@campus.ccm.itesm.mx](mailto:jcsanch@campus.ccm.itesm.mx).

## 1. Introducción

Aunque es común suponer que los inversionistas que disponen de una sólida base estadística para predecir el comportamiento del mercado financiero reciben, en promedio, mayores rendimientos futuros por sus activos que cualquier otro, la tarea de pronóstico no siempre resulta simple ni suficiente para garantizar una mejor estrategia de inversión en el futuro. Las razones obedecen a que: 1) el desconocimiento de los costos de transacción o del grado de aversión al riesgo del inversionista puede conducir a juicios erróneos sobre las ganancias previstas por un pronóstico estadístico, sobre todo porque éste no dice mucho acerca de las posibilidades reales del inversionista que busca materializar esa ventaja; y 2) los movimientos de los rendimientos de los activos están regidos, también, por elementos de idiosincrasia y *ambientales* que hacen que, en no pocas ocasiones, la mejor estrategia de inversión obedezca a la suerte, Lo y Mackinlay (1990), o al conocimiento empírico del inversionista sobre el mercado, Zenios (1998). Entre estos elementos cabe destacar la diferente estacionalidad de los rendimientos de los activos, la influencia de los anuncios de las ganancias o pérdidas de las empresas sobre los precios de sus acciones o, en un plano más general, los cambios experimentados por las instituciones y el ambiente de los negocios (ver Kon, 1984 y Lo y Mackinlay, 1988).

Por tales razones resulta poco recomendable formular una estrategia de inversión futura sin un estudio profundo de esos elementos y de los factores de riesgo de los activos. Pero aquí hay que aclarar que ese estudio no tendría mucho sentido sin el tratamiento estadístico del pronóstico. Ambos son complementarios. El saber, por ejemplo, que una serie de rendimientos no se comporta como una martingala puede crear condiciones de arbitraje para el inversionista, simplemente porque éste cuenta con mayor información para determinar *por qué* cierta estrategia de inversión puede ser exitosa. La mejor calidad de la información crea, mas no asegura, posibilidades de arbitraje aun en mercados eficientes debido a la ventaja competitiva que le da al inversionista tener un mejor conocimiento que sus competidores sobre el historial de ciertos activos claves.<sup>1</sup>

---

<sup>1</sup> Para evitar que surjan contradicciones entre los conceptos de mercado eficiente y arbitraje, Lo (1997) recomienda utilizar la hipótesis de eficiencia relativa de mercados en lugar de la de eficiencia absoluta, ya que esta última acepta la existencia de arbitraje sólo en caso de que haya *ineficiencias* de mercado, y no por efecto de ventajas competitivas permanentes (que resultarían de tener, por ejemplo, una capacidad superior de innovación financiera) como lo hace, en efecto, la primera versión.

Decidir si los rendimientos de los activos son o no pronosticables es, en buena medida, un asunto relacionado con su distribución de probabilidad. Si la distribución empírica es normal, entonces el pronóstico del rendimiento de un activo o de un portafolio se puede elaborar con cualquier modelo de media-varianza. El análisis del valor en riesgo (VeR) se puede llevar a cabo sin mucha dificultad con el estudio del primer y segundo momentos de la distribución normal.<sup>2</sup> Pero si, en cambio, la distribución empírica no es normal, entonces el pronóstico se complica, porque ahora hay que establecer si la serie de rendimientos de los activos se ajusta a una mezcla de distribuciones discretas, a un valor extremo o a un proceso estocástico en particular. Y en cualquiera de los casos la solución es ardua, pues además de que ya no es posible seguir utilizando fielmente las técnicas tradicionales de VeR, los problemas para encontrar una ecuación de pronóstico para los rendimientos no son nada despreciables.

El objetivo de este trabajo consiste en explicar los problemas estadísticos más comunes en la elaboración de un pronóstico cuando el primer y segundo momentos de una distribución de rendimientos no son normales estacionarios.<sup>3</sup> Para ilustrar los resultados se desarrolla un análisis estadístico detallado del paquete accionario *Banamex-30*.<sup>4</sup> Las conclusiones ahí obtenidas sostienen que en México, al igual que en los países desarrollados, las distribuciones de las series de rendimientos diarios de los activos no son normales estacionarias

---

<sup>2</sup> Las referencias que haremos en el documento en torno al VeR están orientadas, única y exclusivamente, a ejemplificar la manera en que se afectan las técnicas de administración de riesgos cuando las distribuciones empíricas no son normales, pero sin profundizar en su estimación. El supuesto de normalidad estacionaria subyacente en su cómputo, así como su uso generalizado en la estimación del riesgo del capital, hacen que el VeR sea un ejemplo para ilustrar las dificultades asociadas con la infracción de ese supuesto.

<sup>3</sup> Aquí seguimos la recomendación hecha por Campbell, Lo y MacKinlay (1997, p. 9) de concentrarnos en los rendimientos de los activos y no en sus precios, porque: 1) para el inversionista promedio los mercados financieros pueden ser considerados competitivos y, por lo tanto, el tamaño de su inversión no afecta los cambios de precios; y 2) los rendimientos tienen propiedades estadísticas más apropiadas que los precios, tales como estacionariedad y ergodicidad. En aras de la simplicidad aquí usaremos los términos activos financieros y acciones como sinónimos a menos que se especifique lo contrario.

<sup>4</sup> El paquete *Banamex-30* es un portafolio que consta originalmente, como su nombre lo indica, de igual número de acciones. Desafortunadamente, la falta de información nos obligó a reducir el tamaño de la muestra a 21 series accionarias. Aún así, ésta es muy grande, pues abarca series de 1 541 observaciones diarias para cada una de las acciones, entre el 3 de enero de 1994 y el 24 de febrero de

y, por lo tanto, el pronóstico de portafolios óptimos o de activos individuales es poco fiable con base en las técnicas habituales de Ver si no se aplican métodos estadísticos que incluyan procesos estocásticos o funciones de probabilidad transformadas. En concreto se señala que para patrones no lineales de rendimientos, como es el caso de las series aquí estudiadas, las pruebas basadas en cadenas de Markov de alto orden pueden resultar un instrumento de pronóstico más robusto que los tradicionales modelos heterocedásticos lineales.

El trabajo se divide en tres partes. En la primera se enuncian los problemas más comunes en el pronóstico de los rendimientos de los activos, cuando las distribuciones empíricas no son normales estacionarias. En la segunda se ensayan algunas soluciones basadas en la utilización de modelos de heterocedasticidad condicional para aquellos activos del paquete *Banamex-30* cuyas series de rendimientos presentan sesgos y exceso de curtosis en sus distribuciones. Finalmente, se propone una metodología de estudio para pronósticos de series de rendimientos no lineales basadas en cadenas de Markov de segundo orden. Las conclusiones incorporan una breve discusión acerca de los resultados alcanzados.

## 2. Los hechos estilizados sobre el supuesto de normalidad en los rendimientos de los activos financieros

No existe quizás supuesto más fuerte para la teoría de portafolios de media-varianza o los modelos de precios de los activos de capital, que el relacionado con la forma de la distribución de los rendimientos de los activos. En general, se supone que esa distribución es normal multivariada con parámetros estacionarios a través del tiempo (Kon, 1984). La discusión en la literatura se reduce a determinar si las distribuciones empíricas se ajustan a ese supuesto o, más concretamente,

---

2000. Los rendimientos fueron calculados de acuerdo con la práctica habitual:

$$R_t = \ln \left( \frac{X_t}{X_{t-1}} \right),$$

donde  $R_t$  es el rendimiento en el tiempo  $t$  de un activo diario y  $X_t$  es el precio del activo al cierre de cada día. Adicionalmente, y para fines de comparación, incluimos la serie del índice *Banamex-30*, tal como es computado por dicho banco. En el cálculo de cada serie de rendimientos diarios se hicieron los ajustes correspondientes para distintos años calendarios así como para los lunes de cada semana.

a verificar si las hipótesis de normalidad y de valores estacionarios para la media y la varianza son válidas para activos con rendimientos que mantienen cierta correlación entre sí.

Por lo que toca a las pruebas de normalidad, los procedimientos incluyen varios pasos complementarios. El primero consiste en determinar si las series originales de los rendimientos son univariadamente normales. Para tal efecto, se construyen las gráficas  $Q$ - $Q$  como un medio expedito para observar la correlación existente entre el conjunto ordenado de datos observados  $x_{(j)}$  y los cuantiles normales  $q_{(j)}$ . Si la pendiente de la recta compuesta por las coordenadas  $(q_{(j)}, x_{(j)})$  es cercana a los 45 grados y el coeficiente de correlación  $r_Q$  (también llamado coeficiente de *rectitud*) entre esas cantidades es estadísticamente significativo, digamos al 5% del nivel de significancia tal como lo sugiere RiskMetrics<sup>TM</sup> (1995, p. 51), entonces es posible confiar en que el análisis rendimiento-riesgo de los activos, basado en el primer y segundo momentos de una distribución normal, es correcto. Si no es así, entonces, lo recomendable es corregir esos datos mediante transformaciones logarítmicas (método *logit* de proporciones y de correlaciones de Fisher) o de potencias del tipo Box-Cox a fin de darles a esos datos una apariencia más *normal* (véase Johnson y Wichern, 1992 y Kaskey, *et al.*, 1980).

Las correcciones no garantizan, sin embargo, el restablecimiento del supuesto de normalidad. La excesiva presencia de rendimientos más bajos o más altos que los previstos por una normal puede anular el efecto suavizador de la transformación operada e, incluso, hacer que el ajuste disminuya el valor original de  $r_Q$ .<sup>5</sup> Esto se comprobó al construir las gráficas  $Q$ - $Q$  del paquete *Banamex-30*, cuyos pares de puntos  $(q_{(j)}, x_{(j)})$  arrojaron valores de  $r_Q$  de .9630 para las series no corregidas y de .9620 para las corregidas con potencias Box-Cox. Ambos valores resultaron, por supuesto, inferiores al valor crítico (.9953) para pruebas de normalidad del coeficiente de correlación con  $\alpha = 0.5$  y  $n > 300$ , lo cual llevó a no aceptar la hipótesis de normalidad para ninguna serie de datos. Cuando el coeficiente de *rectitud* no satisface un valor crítico, aún después de la transformación, lo recomendable es no seguir haciendo inferencias sobre el valor esperado y la desviación estándar de la distribución empírica, debido a que en ese caso los estimadores no convergerían asintóticamente a una  $\chi_v^2$ , con  $v$  grados de libertad. Bajo esas condiciones, los pronósticos sobre la volatilidad

---

<sup>5</sup> En estos casos, el "efecto de arrastre" del operador empleado provoca que valores extremos se conviertan en valores moderados y éstos, a su vez, en valores extremos, alterando aún más el sesgo o la curtosis de la distribución empírica.

de los rendimientos de los activos estarían subestimados en relación con los valores de los parámetros de una distribución normal teórica (Zangari, 1996).

El segundo paso consiste en corroborar esas pruebas univariadas con estimaciones estadísticas sobre el sesgo y exceso de curtosis de los rendimientos de los activos.<sup>6</sup> De acuerdo con el cuadro 1, en el que se condensa la información relativa a los rendimientos diarios de las acciones del paquete *Banamex-30*, tan sólo la serie de GCARSO-A1 mostró un comportamiento parcialmente *normal* al arrojar un *valor p* de 7.99% en su prueba de sesgo (no así en la de curtosis), que significa que la probabilidad de cometer error tipo 1 contra la hipótesis nula de normalidad es muy alta, al menos superior al 5% de nivel de significancia. Sin embargo, este caso no afecta la conclusión general, ya que al realizar las pruebas *ómnibus* de sesgo y curtosis y de Shapiro-Wilk-Francia se confirmó el rechazo a la hipótesis nula de normalidad para GCARSO-A1 y las demás series con *valores p* iguales a cero.<sup>7</sup> Los cálculos basados en simulaciones de Monte Carlo arrojaron los mismos resultados.

Los resultados no hacen más que confirmar que los rendimientos diarios en México siguen un patrón muy similar al descrito por varios autores para el caso de EU y Europa (Affleck-Graves y McDonald, 1989). La ausencia de normalidad en los rendimientos diarios de los activos es un hecho de sobra conocido desde que Mandelbrot (1963) y Fama (1965) mostraron la presencia de colas gordas en las series del índice industrial *Dow Jones*.<sup>8</sup> El sesgo y exceso de curtosis de esas series dio lugar a que se ensayaran dos criterios alternativos para subsanar el problema derivado de la falta de normalidad. Uno de ellos fue someter a prueba la hipótesis de normalidad en escenarios donde

---

<sup>6</sup> Los coeficientes de sesgo y curtosis maestres son definidos, respectivamente, como  $\sqrt{b_1} = m_3/s^3$  y  $\sqrt{b_2} = m_4/s^4$ , donde  $m_3$  y  $m_4$  son los momentos alrededor de la media y  $s$  la desviación de las observaciones respecto a la media muestral.

<sup>7</sup> Las pruebas que aparecen en las últimas dos columnas de los cuadros 1 y 2 son las llamadas pruebas *ómnibus*, porque combinan el sesgo y la curtosis en un solo estadístico. De las dos pruebas, la de Shapiro-Wilk-Francia es la más general pues, a diferencia de la de D'Agostino y Pearson (la otra prueba), contempla los casos en que el sesgo y la curtosis no son independientes entre sí. Para mayores detalles ver Mardia (1980).

<sup>8</sup> La presencia de colas gordas se debe al exceso de curtosis de una distribución empírica en relación con la supuesta por una normal teórica. El exceso aumenta la probabilidad de presenciar valores grandes y pequeños, a costa de disminuir la probabilidad de ocurrencia de los valores moderados.

**Cuadro 1**  
*Pruebas de normalidad univariada para series de rendimientos diarios*

<i>Activo</i>	<i>Sesgo</i>	<i>Valor p</i>	<i>Monte Carlo Valor p</i>	<i>Curtosis</i>	<i>Valor p</i>	<i>Monte Carlo Valor p</i>	<i>Prueba conjunta Valor p</i>	<i>Shapiro -Wilk Valor p</i>
ALFA.A	1.170	0.000	0.000	13.346	0.000	0.000	0.000	0.000
APASCO	0.777	0.000	0.000	6.196	0.000	0.000	0.000	0.000
BANACCI.O	0.295	0.000	0.000	2.979	0.000	0.000	0.000	0.000
BIMBO.A	0.267	0.000	0.000	3.304	0.000	0.000	0.000	0.000
CEMEX.CPO	0.247	0.000	0.000	6.159	0.000	0.000	0.000	0.000
CIFRA.V	0.548	0.000	0.000	5.893	0.000	0.000	0.000	0.000
COMERCI.UBC	0.636	0.000	0.000	11.479	0.000	0.000	0.000	0.000
DESC.B	0.624	0.000	0.000	4.537	0.000	0.000	0.000	0.000
ELEKTRA.CPO	0.667	0.000	0.000	8.225	0.000	0.000	0.000	0.000
FEMSA.UBD	0.501	0.000	0.000	12.466	0.000	0.000	0.000	0.000
GCARSO.A1	-0.016	0.000	0.000	4.851	0.799	0.578	0.000	0.000
GFNORTE.O	0.151	0.000	0.000	4.232	0.015	0.015	0.000	0.000
GMEXICO.B	0.765	0.000	0.000	3.643	0.000	0.000	0.000	0.000
ICA	-0.144	0.000	0.000	9.436	0.000	0.000	0.000	0.000
SAVIA.A	0.533	0.000	0.000	8.404	0.000	0.000	0.000	0.000

Cuadro 1  
(continuación)

<i>Activo</i>	<i>Sesgo</i>	<i>Valor p</i>	<i>Monte Carlo Valor p</i>	<i>Curtosis</i>	<i>Valor p</i>	<i>Monte Carlo Valor p</i>	<i>Prueba conjunta Valor p</i>	<i>Shapiro -Wilk Valor p</i>
SORIANA.B	0.342	0.000	0.000	2.579	0.000	0.000	0.000	0.000
TAMSA	0.518	0.000	0.000	6.220	0.000	0.000	0.000	0.000
TELECOM.A1	0.389	0.000	0.000	4.292	0.000	0.000	0.000	0.000
TELMEX.L	0.626	0.000	0.000	3.221	0.000	0.000	0.000	0.000
TELEVISA.CPO	0.620	0.000	0.000	5.487	0.000	0.000	0.000	0.000
VITRO.A	0.905	0.000	0.000	9.101	0.000	0.000	0.000	0.000
BANAMEX-30	0.222	0.000	0.000	6.456	0.000	0.000	0.000	0.000



los rendimientos cubrían diferentes períodos. La conclusión fue que, en efecto, la hipótesis era altamente sensible a la periodicidad de la serie de los datos y que, por tanto, no eran triviales las consecuencias de tomar tal o cual intervalo (días, semanas o meses). El otro criterio consistió en suponer que la existencia de colas gordas era indicativo de que los cambios en los precios de los activos eran mejor captados por la distribución no normal estable de Pareto, con exponente característico menor que 2, que por una normal idéntica e independientemente distribuida. La evidencia aquí no fue tan favorable como en el otro criterio, debido a que autores como Blattberg y Gonedes (1974) encontraron que los exponentes característicos de la distribución de la suma intertemporal para los rendimientos sobre los activos individuales o portafolios aumentaba con el tamaño de la suma; lo cual era una abierta infracción a la propiedad de estabilidad de la distribución de Pareto.

Los resultados obtenidos con el paquete *Banamex-30* confirmaron que la tarea de pronóstico de los rendimientos de los activos es muy distinta para series diarias que para series mensuales.<sup>9</sup> El valor de  $r_Q$  para las series mensuales corregidas con la transformación de Box-Cox (.9956) resultó normal para el valor crítico correspondiente a  $\alpha = 0.5$ , así como para todas las pruebas omnibus de sesgo y curtosis. El punto a discusión es si esas pruebas son suficientes para hacer el pronóstico del cambio monetario en el valor de un portafolio. La respuesta no es inmediata, pues antes debe verificarse si la correlación cruzada de los rendimientos no afecta el supuesto de normalidad multivariada. Y aquí entramos al tercer paso requerido para probar normalidad en las series de rendimientos.

En efecto, una vez aceptada la hipótesis de normalidad univariada lo que procede es verificar si la serie de rendimientos es también normal multivariada, pues la primera no implica necesariamente a la segunda, para así descartar la presencia de cosesgos y cocurtosis en el análisis de rendimiento-riesgo de un portafolio.<sup>10</sup> Las pruebas incluyen un examen de normalidad bivariada y de momentos conjuntos y otro de orden generalizado con base en el estadístico de Wald.

---

<sup>9</sup> Aquí seguimos la recomendación hecha por RiskMetrics<sup>TM</sup> (1995, p. 68) de extrapolar los rendimientos mensuales a partir de las series no traslapadas de rendimientos diarios.

<sup>10</sup> Es bien conocido que distribuciones que son univariadamente normales pueden no ser normalmente multivariadas debido a la presencia de efectos cruzados entre los rendimientos de los activos, y que distribuciones que no son univariadamente normales nunca pueden ser normales multivariadas (véase Richardson y Smith, 1993).

**Cuadro 2**  
*Pruebas de normalidad univariada para series de rendimientos mensuales*

<i>Activo</i>	<i>Sesgo</i>	<i>Valor p</i>	<i>Monte Carlo Valor p</i>	<i>Curtosis</i>	<i>Valor p</i>	<i>Monte Carlo Valor p</i>	<i>Prueba conjunta Valor p</i>	<i>Shapiro -Wilk Valor p</i>
ALFA.A	-0.508	0.085	0.030	0.578	0.341	0.348	0.105	0.244
APASCO	-0.048	0.872	0.934	0.407	0.502	0.413	0.663	0.708
BANACCIO	0.033	0.912	0.934	1.382	0.023	0.046	0.132	0.013
BIMBO.A	0.564	0.056	0.093	0.410	0.499	0.492	0.096	0.055
CEMEX.CPO	-0.404	0.171	0.183	0.810	0.182	0.146	0.124	0.111
CIFRA.V	0.237	0.422	0.461	0.080	0.895	0.875	0.641	0.721
COMERCI.UBC	0.024	0.934	0.944	0.138	0.820	0.880	0.899	0.758
DESC.B	-0.320	0.277	0.366	1.235	0.042	0.052	0.095	0.225
ELEKTRA.CPO	0.266	0.367	0.320	0.959	0.114	0.065	0.174	0.094
FEMSA.UBD	-0.194	0.510	0.440	0.017	0.977	0.887	0.752	0.455
GCARSO.AI	-0.247	0.403	0.424	0.844	0.164	0.142	0.226	0.376
GFNORTE.O	0.190	0.520	0.486	1.428	0.019	0.044	0.102	0.032
GMEXICO.B	0.000	0.999	0.987	1.885	0.012	0.694	0.050	0.031
ICA	0.039	0.894	0.927	1.403	0.021	0.045	0.127	0.074
SAVIA.A	0.098	0.738	0.761	1.703	0.005	0.035	0.081	0.033

**Cuadro 2**  
(continuación)

Activo	Sesgo	Valor p	Monte Carlo Valor p	Curtosis	Valor p	Monte Carlo Valor p	Prueba conjunta Valor p	Shapiro - Wilk Valor p
SORIANA.B	0.110	0.710	0.748	0.689	0.256	0.240	0.409	0.321
TAMSA	0.032	0.913	0.935	1.988	0.001	0.027	0.058	0.093
TELECOM.A1	-0.107	0.718	0.724	0.677	0.264	0.321	0.418	0.193
TELMEX.L	0.285	0.334	0.440	0.414	0.494	0.410	0.382	0.553
TELEvisa.CPO	0.031	0.917	0.937	1.670	0.006	0.036	0.089	0.102
VITRO.A	0.056	0.849	0.908	0.318	0.600	0.563	0.737	0.404
BANAMEX-30	0.000	1.000	0.969	0.131	0.695	0.641	0.926	0.683
$X^2_{22}$	44.544	0.003		19.521	0.612		0.385	

Nota.- La última fila contiene los valores de la prueba de normalidad multivariada para el estadístico de Wald que converge a una  $x^2_v$ .

El cuadro 3 exhibe el porcentaje de observaciones fuera de los contornos de las elipsoides de las 21 funciones de densidad normal bivariadas construidas al comparar las series de rendimientos de CIFRA-V, que se toma como base de referencia, contra cada una de las series de los activos. Si los contornos del elipsoide están balanceados en torno a los rendimientos medios de la distribución conjunta,  $\bar{R}$ , entonces lo que se espera es que éstos contengan un cierto porcentaje de los valores de los rendimientos bajo la superficie de esa distribución con probabilidad  $1 - \alpha$ , para asegurar normalidad bivariada. El criterio adoptado fue que el porcentaje de observaciones incluido por los contornos de densidad debía ser del 50%, de otra manera se sospecharía de la presencia de efectos de correlación cruzada entre los rendimientos de los activos y los de CIFRA-V.

En virtud de que el porcentaje promedio computado para esas funciones (52.38%) excede ligeramente al 50% de las observaciones que deberían ser contenidas por los contornos de densidad bivariados, se pensaría, inicialmente, en la presencia de efectos cruzados entre las series. La prueba concluyente consiste en decidir si los cosesgos y cocurtosis de estas funciones y de las que resultan de comparar cada una de las series de activos entre sí, son espurios o reflejan, en realidad, la existencia de colas gordas y sesgadas en las distribuciones conjuntas. El cuadro 3 también muestra que los cosesgos y las cocurtosis son aparentemente espurios, debido a que no es posible rechazar la hipótesis de normalidad bivariada para niveles de significancia iguales o menores que 5%. La misma situación se repite al cambiar de base de referencia y comparar cada una de las combinaciones de series de rendimientos. De acuerdo con el último renglón del cuadro 2, los valores críticos del estadístico generalizado de Wald confirman que, en efecto, los momentos cruzados para cualquier combinación de series de rendimientos no afectan el supuesto de *normalidad multivariada*, NMV, de la muestra, con lo que el pronóstico del portafolio con estas series puede hacerse con el procedimiento habitual de VeR.<sup>11</sup>

---

<sup>11</sup> Esta conclusión es válida aun cuando se rechaza la hipótesis de normalidad para el sesgo (no así para la curtosis). El estadístico de Wald ( $W$ ) se calcula como  $W = nb_2'[V(b_2)]^{-1}b_2 \sim \chi_N^2$ ; donde  $n$  es el número de observaciones en la muestra,  $b_2$  es el vector de medidas de curtosis (o  $b_1$  si consideramos los sesgos) para los  $N$  activos, y  $V(b_2)$  es la matriz de varianzas y covarianzas de las medidas de curtosis (o  $V(b_1)$  si consideramos la matriz correspondiente a las medidas de sesgos). La extensión de este indicador a pruebas de cosesgos y cocurtosis y de NMV basadas en una función generalizada de momentos para  $M$ -parámetros puede consultarse en Richardson y Smith (1993).

**Cuadro 3**  
*Momentos conjuntos de sesgo y curtosis y porcentaje de observaciones dentro de los contornos de densidad constante (observaciones mensuales)*

<i>Activo</i>	<i>Correlación</i>	<i>Cosesgo</i>	<i>Valor p</i>	<i>Cocurtosis</i>	<i>Valor p</i>	<i>%</i>
ALFA.A	0.535	-0.044	0.586	0.061	0.807	45.833
APASCO	0.570	-0.095	0.257	0.030	0.911	51.389
BANACCI.O	0.527	-0.080	0.312	0.396	0.104	54.167
BIMBO.A	0.387	-0.071	0.282	-0.198	0.266	54.167
CEMEX.CPO	0.517	-0.292	0.000	0.017	0.944	50.000
COMERCI.UBC	0.599	0.070	0.426	0.008	0.977	50.000
DESC.B	0.530	-0.136	0.089	0.361	0.141	54.167
ELEKTRA.CPO	0.474	-0.293	0.000	0.195	0.368	58.333
FEMSA.UBD	0.722	0.115	0.270	-0.182	0.620	50.000
GCARSO.A1	0.737	0.098	0.354	0.414	0.274	50.000
GFNORTE.O	0.446	-0.232	0.001	0.412	0.042	55.556
GMEXICO.B	0.365	0.130	0.044	0.304	0.074	55.556
ICA	0.639	-0.028	0.759	-0.085	0.784	54.167
SAVIA.A	0.428	-0.104	0.137	0.425	0.029	51.389

**Cuadro 3**  
(continuación)

<i>Activo</i>	<i>Correlación</i>	<i>Cos ego</i>	<i>Valor p</i>	<i>Cocurtosis</i>	<i>Valor p</i>	<i>%</i>
SORIANA.B	0.639	0.013	0.888	-0.211	0.496	48.611
TAMSA	0.420	0.068	0.324	0.594	0.002	54.167
TELECOM.A1	0.610	-0.001	0.990	0.501	0.086	52.778
TELMEX.L	0.640	0.016	0.862	-0.245	0.430	52.778
TELEvisa.CPO	0.593	-0.194	0.025	-0.221	0.432	54.167
VITRO.A	0.555	-0.010	0.904	-0.243	0.346	52.778
BANAMEX-30	0.769	-0.043	0.699	0.158	0.696	50.000
<b>Promedio</b>						52.381

Nota.- CIFRA-V se toma como referencia para la estimación de los cosegos y cocurtosis. En la columna 7 aparecen los porcentajes de observaciones fuera de los contornos de densidad constante.

La verificación del supuesto de NMV para series mensuales no resuelve, sin embargo, un problema crucial en la práctica financiera cotidiana y que consiste en que las técnicas de pronóstico basadas en los estimadores de VeR se desarrollan generalmente sobre bases diarias. Un estimador de VeR corresponde, de hecho, al valor crítico de una distribución de probabilidad de pérdidas y ganancias potenciales diarias de un portafolio (Kupiec, 1995). Por ese motivo, el verdadero problema derivado de la validación del supuesto de NMV se relaciona con el tratamiento de las series de rendimientos diarios que, como vimos, presentan colas gordas y sesgadas en su distribución.

### 2.1. *Las propuestas recientes para corregir series de rendimientos diarios afectadas por sesgo y curtosis*

Las soluciones a este problema han sido muchas y variadas, pero las principales se encuentran resumidas en el documento técnico propuesto recientemente por el grupo de RiskMetrics<sup>TM</sup> (1996), cuyo método para medir exposición al riesgo de un portafolio es el más usado en varios países, entre los que se cuenta México. Para dicho grupo los rendimientos siguen una distribución NMV condicional que no incondicional, lo cual quiere decir que aun cuando las series de rendimientos originales,  $R_t$ , no son NMV, divididas éstas por sus respectivas desviaciones estándar pronosticadas,  $\sigma_t$ , si son NMV con media cero y varianza 1. De esta manera, los tres métodos de pronósticos que propone RiskMetrics<sup>TM</sup> reconocen los cambios de volatilidad que presentan las distribuciones de rendimientos en el tiempo.

El primer método de pronóstico es el llamado método estándar, en el que la varianza de los rendimientos es una función del pronóstico de la varianza de los días previos y del cuadrado del rendimiento.<sup>12</sup> El segundo es el de mezcla de normales, y se introduce para medir eficientemente la falta de simetría y exceso de curtosis que, según evidencia mostrada por el mismo documento en mención, no es capturada por el método estándar cuando pronostica VeR con bandas de intervalo de confianza del 95 y 98% (Zangari, 1996). Finalmente el tercer método, conocido como *distribución generalizada de errores*,

---

<sup>12</sup> RiskMetrics recomienda un factor de decaimiento  $\lambda = 0.94$  y un horizonte histórico de 75 días, en promedio, para llevar a cabo pronósticos de varianza con ventana de uno a diez días. Si vemos con detenimiento la ecuación de la varianza del modelo estándar podemos observar que este grupo atribuye al pronóstico de la volatilidad un peso del 94% a la varianza actualizada del periodo de 74 días, y sólo un 6% a la varianza del día en que se hace el pronóstico.

DGE; es el más flexible de todos, porque además de corregir sesgo y exceso de curtosis captura la volatilidad grupal (*volatility clustering*), que es otro de los problemas recurrentes en las series de rendimientos diarios (ver Kim y Kon, 1994).<sup>13</sup>

La sugerencia hecha por el grupo de RiskMetrics<sup>TM</sup> y otros autores, como Duffie y Pan (1997), Hull y White (1998) y Li (1999) es que, una vez calculados los rendimientos y volatilidades para las series corregidas con cualquiera de los tres métodos, es importante comprobar si, en efecto, estos métodos dan cuenta de los problemas de sesgo y curtosis antes de llevar a cabo la estimación definitiva del VeR de un portafolio. Las pruebas comprenden la construcción de intervalos de confianza con base en el enfoque de funciones estimadas (*estimating functions*) o la transformación de distribuciones de probabilidad que son supuestas NMV. Las primeras incluyen la comparación de dos tipos de intervalos de confianza: uno que incorpora explícitamente el sesgo y la curtosis de las series generadas por funciones estimadas y otro construido por métodos, como los vistos arriba, que suponen normalidad (véase Li 1999).<sup>14</sup> La transformación de distribuciones consiste, por su parte, en el mapeo percentil por percentil de una distribución alternativa y una normal estándar (ver Hull y White 1998). La distribución alternativa es una transformación de la distribución empírica que restituye las propiedades de una distribución normal, capaz de reflejar las probabilidades de hasta tres desviaciones estándar en el rendimiento de los activos.

Independiente de la importancia de esta sugerencia, lo realmente interesante de los métodos de pronóstico de RiskMetrics<sup>TM</sup> (1996) es que resumen los esfuerzos realizados por varios autores para obtener una mejor especificación de aquellas distribuciones empíricas de ren-

---

<sup>13</sup> La volatilidad grupal se refiere al comportamiento seguido por los cambios en los precios ante un cambio previo en los precios de los activos. Si la volatilidad es positiva, entonces grandes (pequeños) cambios en los precios tenderán a seguir a grandes (pequeños) cambios en los precios. Si ésta es negativa, entonces se observará una relación inversa entre los cambios de precios. A diferencia del método de mezcla de normales, el DGE tiene la ventaja de actualizar continuamente los cambios en la varianza de los rendimientos de los activos. La gran variedad de formas que puede tomar la DGE permite, también, hacer ajustes más precisos en el pronóstico de valores extremos en el primer y noventa y noveno percentil, tal como lo sugieren los cálculos computados por RiskMetrics<sup>TM</sup> (1996) en un análisis comparativo de los tres métodos (p. 242).

<sup>14</sup> La longitud de los intervalos de confianza basados en las funciones estimadas está relacionada positivamente con la curtosis y negativamente con el valor absoluto del sesgo.



dimientos afectadas por cambios temporales en la media y la varianza (el otro gran tema asociado con las pruebas de normalidad). Dichos esfuerzos los podemos dividir en dos grandes grupos de modelos alternativos: los de heterocedasticidad tempo-independientes (como el de mezcla de normales) y los tempo-dependientes (como el DGE). La discusión acerca de cual de los dos tipos de modelos describe mejor las variaciones de los parámetros en una distribución es en opinión de Kim y Kon (1994), una cuestión empírica, que tiene mucho que ver con la naturaleza del instrumento financiero bajo consideración. Por esa razón, resulta pertinente hacer una comparación de varios tipos de modelos pertenecientes a ambos grupos, a fin de tener una idea más precisa de la naturaleza del problema a resolver cuando se estima  $\text{VeR}$  con distribuciones no normales.

### 3. Comparación de métodos heterocedásticos usados para corregir distribuciones de rendimientos de activos con parámetros no estacionarios

La intuición detrás de la aplicación de los modelos heterocedásticos en finanzas es relativamente simple. Como se sabe, los sesgos y excesos de curtosis que presentan las distribuciones de los rendimientos son debidos a desplazamientos de sus parámetros entre un número finito de valores, ocasionados por cambios en la estructura de capital, eventos exógenos al mercado (desplazamientos de tiempo ordenado) o por la mayor varianza de los rendimientos en algunos días de la semana (desplazamientos cíclicos). La persistencia con que los primeros desplazamientos explican el sesgo observado y los segundos el exceso de curtosis, demanda pues, la utilización de modelos de distribución que capturen las variaciones en la media y la varianza a través del tiempo.<sup>15</sup> Entre los modelos cabe mencionar, por un lado, a la distribución de *Student*, la mezcla discreta generalizada de distribuciones normales y el modelo mixto de difusión con salto (modelos de heterocedasticidad condicional tempo-independientes, HCTI) y, por el otro, a algunos modelos tipo GARCH (modelos de heterocedasticidad condicional tempo-dependientes, HTCD).

Por lo que toca a los modelos HCTI, no hay quizás nada más usual que comparar el poder descriptivo de la distribución  $t$  de *Student* con el de la mezcla de normales. La justificación radica en el alegato de

---

<sup>15</sup> Con estos modelos el sesgo observado se explica por desplazamientos en la media, y las colas gordas por los desplazamientos en la varianza (Kon, 1984).

que la distribución de *Student* captura más valores extremos que una normal estacionaria, pero no la combinación de los desplazamientos cíclicos y estructurales, o de tiempo ordenado, de los parámetros de la distribución (Kim y Kon, 1994). Las pruebas contenidas en el cuadro 4 no parecen, sin embargo, confirmar los resultados encontrados para otros países, pues mientras que la distribución de *Student* describe mejor las series diarias de 15 activos, la mezcla de dos y tres normales sólo lo hace para 6 activos y el índice *Banamex-30*. La conclusión se desprende al observar el signo del estadístico del logaritmo de la tasa de verosimilitud,  $\Lambda_{NS} = \frac{\ell(\theta/r)}{\ell(\delta/r)}$ , entre una y más distribuciones normales ( $N$ ) y una distribución *Student*( $S$ ). Si el logaritmo de  $\Lambda_{NS}$  es positivo, entonces es más probable que la mezcla de normales genere los datos observados de una manera más confiable que la distribución de *Student*, contrario a lo que sucedería si el signo de  $\Lambda_{NS}$  fuera negativo (para mayor información consúltese Kon 1984). Los valores negativos del estadístico en los renglones del cuadro 4 indican que la distribución de *Student* es superior a la descripción hecha por una normal,  $\Lambda_{1S}$ , en todas las series de activos, pero no así al incluir la mezcla de dos normales: en este caso los signos positivos de  $\Lambda_{2S}$  revelan una mayor capacidad descriptiva de la mezcla sobre la distribución *Student* en las series diarias de ALFA, APASCO, CIFRA, COMERCI, GCARSO, ICA, TELECOM y el índice *Banamex-30*. La inclusión de tres normales no mejora la potencia descriptiva del modelo de mezclas sobre la distribución de *Student*.<sup>16</sup>

Para probar la robustez de este resultado comparamos el poder descriptivo de la distribución de *Student* con otros modelos competidores. La primera comparación incluyó al modelo de difusión con saltos, también de HCTI. La popularidad de esta mezcla de procesos estocásticos reside en su capacidad de capturar movimientos irregulares y bruscos que afectan los valores estacionarios de la media y la varianza, en una forma muy parecida a como lo hacen los modelos con parámetros cambiantes en el tiempo (véase Ball y Torous, 1985). La forma funcional de la ecuación aquí empleada supone que cada vez que hay un arribo de información anormal existe un salto instantáneo

---

<sup>16</sup> Las pruebas se hicieron estimando las series de cada activo bajo el supuesto de que las distribuciones se comportaban como una distribución de *Student* y una mezcla de normales, de acuerdo con las siguientes funciones de verosimilitud:  $L_T(\delta) = \sum_{t=1}^T \log g(r_t/\delta)$  (para la distribución  $t$  de *Student*); y  $\ell(\theta/r) = \prod_{i=1}^T \left[ \sum_{i=1}^N \lambda_i p(r/\gamma_i) \right]$  (para la mezcla de funciones) donde:  $\gamma_i = (\alpha_i, \sigma_i^2)$ ; y  $p(r/\gamma_i)$  es la función de densidad normal con media  $\alpha_i$  y varianza  $\sigma_i^2$ .

en el precio de las acciones de tamaño  $J$  que es independiente de  $Z$ ,

$$\log(r_t) = \left(\alpha - \frac{1}{2}\sigma^2\right)t + \sigma Z(t) + \sum_{n=1}^{N(t)} J_n \quad (1)$$

donde:

- $Z(t)$  = proceso estandarizado de Wiener,
- $N(t)$  = proceso de Poisson no homogéneo,
- $J_n$  = una variable aleatoria normal con media cero y varianza que representa el logaritmo de uno, más el porcentaje de cambio en el precio del activo causado por el  $n$ -ésimo salto,
- $\alpha$  = la tasa instantánea condicional esperada de rendimiento por unidad de tiempo del proceso de Wiener,
- $\sigma^2$  = varianza condicional instantánea de los rendimientos por unidad de tiempo del proceso de Wiener.

#### Cuadro 4

*Comparación entre una distribución Student y la mezcla discreta de normales generalizada*

Activo	$\Lambda_{1S}$	$\Lambda_{2S}$	$\Lambda_{3S}$
ALFA.A	-0.659	295.965	ND
APASCO	-0.906	187.764	ND
BANACCI.O	-0.841	ND	ND
BIMBO.A	-1.124	ND	ND
CEMEX.CPO	-0.770	ND	ND
CIFRA.V	-0.858	147.306	184.021
COMERCI.UBC	-0.646	-23.913	ND
DESC.B	-0.848	ND	ND
ELEKTRA.CPO	-0.245	ND	ND
FEMSA.UBD	-0.580	ND	ND
GCARSO.A1	-0.872	123.809	ND
GFNORTE.O	-0.817	ND	ND
GMEXICO.B	-1.006	ND	ND
ICA	-0.648	-15.246	ND
SAVIA.A	-1.100	ND	ND
SORIANA.B	-1.100	ND	ND

**Cuadro 4**  
(continuación)

<i>Activo</i>	$\Lambda_{1S}$	$\Lambda_{2S}$	$\Lambda_{3S}$
TAMSA	-0.767	ND	ND
TELECOM.A1	-0.922	149.136	ND
TELMEX.L	-1.290	ND	ND
TELEVISA.CPO	-0.795	ND	ND
VITRO.A	-0.678	ND	ND
BANAMEX-30	-2.242	2.159	ND

ND: no disponible.

La segunda, y última comparación, toma los modelos EGARCH y de *Glosten, Jagannathan y Runkle*, GJR, cuya característica común es suponer que la media y la varianza en el tiempo  $t$  están condicionadas al conjunto de información disponible en  $t - 1$ ,  $F_{t-1}$ , de acuerdo con las siguientes ecuaciones:

$$m_t = E(r_t/F_{t-1}) \quad (2)$$

$$h_t = V(r_t/F_{t-1}) \quad (3)$$

donde los residuales o rendimientos no anticipados, son calculados como  $\epsilon_t = r_t - m_t$ . La utilidad de llevar a cabo un análisis conjunto de ambos modelos reside en las posibilidades de análisis que ofrecen las diferencias de sus supuestos sobre el conjunto de información relevante  $F_{t-1}$  y, en general, sobre la forma de las ecuaciones (2) y (3). Con base en ellas, uno puede modelar, por ejemplo, la media condicional con componentes autorregresivos y de promedios móviles de diferente orden, y agregarle un coeficiente de respuesta a los cambios en la varianza contemporánea para series de rendimientos que acusan desplazamientos ordenados en el tiempo muy violentos. O bien, se les puede combinar con diferentes variantes de modelos ARMA para mejorar el ajuste de series afectadas por desplazamientos cíclicos (motivados por cambios en la varianza condicional) que alteran, a su vez, la dirección de los desplazamientos ordenados (debidos a cambios en la media condicional); o simplemente para capturar alguna característica adicional que se considere relevante de las distribuciones de rendimientos no normales.<sup>17</sup>

<sup>17</sup> La característica que regularmente se asocia a estos modelos es el efecto de

El procedimiento aquí adoptado consistió en suponer una misma ecuación de la media condicional,  $m_1$  del tipo ARMA (1,1) para los dos modelos (ecuación 4), y dos especificaciones diferentes para la varianza condicional,  $h_t^2$ , en las que se incluyen los parámetros  $\gamma$  y  $k$ , utilizados para estimar la volatilidad grupal y apalancamiento, respectivamente (ecuaciones 5 y 6). Con el propósito de incorporar mayor variedad en la especificación de las varianzas se incluyeron los dos parámetros por separado en la ecuación (5) del modelo EGARCH(1,1), y en forma conjunta a través del parámetro  $\eta_1$ , en la ecuación (6) del modelo GJR (1,1).<sup>18</sup> El orden de los modelos, que es el más usado en la literatura, fue seleccionado para hacer hincapié en que el rendimiento y el riesgo de las series dependen crucialmente del valor esperado y de la varianza acumuladas en el último rezago (ver Sánchez Cerón, 2001):

$$m_t = c_0 + br_{t-1} + a\varepsilon_{t-1} \tag{4}$$

$$\log(h_t^2) = \varpi + \beta \log(h_{t-1}^2) + \gamma \left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| + k \frac{\varepsilon_{t-1}}{h_{t-1}} \tag{5}$$

$$h_t^2 = \varpi + \zeta \varepsilon_{t-1}^2 + \eta_1 d_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \tag{6}$$

La comparación de los modelos es un ejercicio para verificar las bondades de las ecuaciones propuestas por RiskMetrics<sup>TM</sup> (1996), ya que, visto con detenimiento, el modelo de difusión con salto no es más que una mezcla infinita de distribuciones normales, en tanto que el

apalancamiento,  $k$ , el cual se introduce para capturar la asimetría en las distribuciones de rendimientos ocasionada por malas noticias en el mercado accionario. En condiciones normales se espera que aumentos en la volatilidad condicional se expliquen por cambios negativos en el residual estandarizado o, dicho de otra manera, que las malas noticias disminuyan los precios de las acciones e incrementen el apalancamiento financiero y la volatilidad.

<sup>18</sup> Los parámetros de la ecuación (5) indican que si  $\gamma > 0$ , entonces la volatilidad condicional tenderá a aumentar o decrecer en razón proporcional al valor absoluto del residual estandarizado, con lo que se puede afirmar que cambios significativos en los precios de las acciones son precedidos por cambios importantes en los precios (volatilidad grupal). De la misma manera, y como se explicó en la nota al pie anterior, un valor negativo de  $k$  indica la existencia del efecto de apalancamiento. En el modelo GJR ambos parámetros son capturados mediante la agregación del término de asimetría  $\eta_1$  a un modelo GARCH en el que  $d_{t-1} = 1$ , si  $\varepsilon_{t-1} < 0$  y  $d_{t-1} = 0$  de otra manera. El impacto del residual cuadrado sobre la volatilidad condicional variará dependiendo si el residual de rezago 1 es positivo o negativo. Por esa razón en este modelo el efecto de apalancamiento se detecta si  $\eta_1 > 0$ .

EGARCH y el GJR son variantes de los modelos DGE. De esta manera, los resultados obtenidos se reducen a confrontar el poder predictivo de las mezclas de normales (que incluye la mezcla de procesos estocásticos) con el de modelos que privilegian la actualización de la varianza de los rendimientos de acuerdo con diferentes especificaciones.

Las conclusiones relativas a la comparación de los modelos las podemos agrupar en dos grandes apartados. El primero, se refiere a las mejoras en la capacidad descriptiva introducida por cada uno de los últimos tres modelos. De acuerdo con el cuadro 5, el modelo de difusión con saltos resultó adecuado para explicar el comportamiento de 20 de los 21 activos y del índice *Banamex-30*. El rechazo a la hipótesis nula de igualdad de medias y varianzas para cualquier nivel de significancia (con valores  $p$  iguales a cero en todos los casos) indica que la presencia de información anormal si afecta el precio de los activos, alterando con saltos la volatilidad de sus rendimientos.

Cuadro 5

*Valores de los parámetros del proceso de difusión con salto*

	$\Lambda$	$\alpha$	$\sum^2$	$\delta^2$
ALFA.A	0.6443	-0.1720	0.3751	0.1879
	0.0000	0.0000	0.0000	0.0000
APASCO	1.4688	-0.2806	0.2322	0.4920
	0.0000	0.0000	0.0000	0.0000
BANACCI.O	1.2568	-0.1995	0.4754	0.3061
	0.0000	0.0000	0.0000	0.0000
BIMBO.A	1.4953	-0.3141	0.4055	0.5243
	0.0000	0.0000	0.0000	0.0000
CEMEX.CPO	0.6994	-0.1018	0.4171	0.2575
	0.0000	0.0000	0.0000	0.0000
CIFRA-V	1.0512	-0.2389	0.3187	0.2057
	0.0000	0.0000	0.0000	0.0000
COMERCI.UBC	0.9904	0.0454	0.3533	0.1882
	0.0000	0.0000	0.0000	0.0000
DESC.B	0.4661	0.0319	0.3034	0.3397
	0.0000	0.2293	0.0000	0.0000
ELEKTRA.CPO	3.4168	-3.1226	-0.4299	0.5208

**Cuadro 5**  
(continuación)

	$\Lambda$	$\alpha$	$\Sigma^2$	$\delta^2$
	0.0000	0.0000	0.0000	0.0000
FEMSA.UBD	0.1393	0.2676	0.2875	0.4303
	0.0000	0.0000	0.0000	0.0000
GCARSO.A1	-1.3239	1.5506	0.2608	0.3012
	0.0000	0.0000	0.0000	0.0000
GFNORTE.0	0.0157	0.8644	0.8033	0.4983
	0.0000	0.0000	0.0000	0.0000
GMEXICO.B	0.6700	-0.1833	0.3188	0.4117
	0.0000	0.0000	0.0000	0.0000
ICA	0.6606	-0.0590	0.3764	0.2344
	0.0000	0.0000	0.0000	0.0000
SAVIA.A	-1.1981	0.9658	0.2522	0.3070
	0.0000	0.0000	0.0000	0.0000
SORIANA.B	NC	NC	NC	NC
TAMSA	0.5592	0.1527	0.3653	0.2499
	0.0000	0.0905	0.0000	0.0000
TELECOM.A1	0.2281	0.6146	0.3416	0.2437
	0.0000	0.0000	0.0000	0.0000
TELMEX.L	0.5972	0.0680	0.3602	0.2330
	0.0000	0.0000	0.0000	0.0000
TELEVISA.CPO	0.6874	0.0844	0.4407	0.3410
	0.0000	0.0000	0.0000	0.0000
VITRO.A	0.0982	0.6886	0.3180	0.2546
	0.0000	0.0000	0.0000	0.0000
BANAMEX-30	3.0245	-3.2211	-0.6177	0.3535
	0.0000	0.0000	0.0000	0.0000

NC: No converge.

La situación para los modelos EGARCH(1,1) y GJR (1,1) es ligeramente distinta, puesto que aun cuando ambos capturan la depen-

dencia lineal en la media y la varianza, a la vez que abaten la correlación serial en residuales y residuales al cuadrado (Q(12) y Q(12) al cuadrado), sus especificaciones no eliminan por completo el sesgo y exceso de curtosis de las series de rendimientos (ver cuadro 6). En efecto, al comparar los cuadros 1 y 6 se puede observar que mientras los dos nuevos modelos disminuyen la curtosis en 12 casos, las estimaciones sobre el sesgo resultaron menores en sólo 9 series, que las registradas por las distribuciones de rendimientos originales. No obstante, los valores de los coeficientes de sus varianzas condicionales rezagadas,  $\beta$ , son significativamente positivos y menores que uno, lo que muestra que las series de rendimientos tienen un alto factor de persistencia (o larga memoria) sobre la varianza condicional del período anterior. En el caso concreto del EGARCH(1,1), las pruebas sobre los dos parámetros arrojaron resultados mixtos, ya que mientras el modelo comprobó la existencia de volatilidad grupal en todas las series con valores  $p$  iguales a cero, no dio cuenta de la presencia del efecto de apalancamiento en 21 activos. Los valores positivos de  $k$  en la mayoría de las series, así como su pobre significancia estadística en 4 activos que presentaron signo negativo, confirman el rechazo a la hipótesis de que las malas noticias del mercado accionario disminuyen el precio de las acciones. El mismo resultado puede observarse en el modelo GJR, donde los valores positivos de  $\eta_1$  para 11 series de activos no resultaron significativos para niveles menores al 5%.<sup>19</sup>

El segundo apartado de conclusiones comprende la comparación de los modelos bajo el criterio de Schwarz,  $CS$ , de probabilidad posterior:  $CS = -2\frac{l}{n} + t\frac{\log n}{n}$ ; donde  $l$  es el logaritmo de la función de máxima verosimilitud. De acuerdo con dicho criterio, la especificación de un modelo es más probable, o tiene la probabilidad posterior más alta, de predecir el comportamiento de una serie si el valor de  $CS$  es más pequeño, o de manera equivalente, si el logaritmo de las probabilidades posteriores del modelo competidor en relación con una normal estacionaria es negativo. El cuadro 7 revela que al compararse los modelos entre, si la  $t$  de *Student* se desempeñó como mejor hipótesis alternativa contra la normal estacionaria, seguido muy de

---

<sup>19</sup> La explicación a esta aparentemente extraña conclusión puede hallarse en el comportamiento desigual de las acciones durante el periodo bajo estudio. En concreto, después de los movimientos irregulares de los primeros meses de 1995 el mercado accionario entró en una fase relativamente estable que, con sus excepciones, se mantuvo hasta el final del periodo (de hecho, las acciones de la muestra siguen cotizando hoy en día); lo cual probablemente inhibió el efecto apalancamiento de las empresas durante 1995. La división del periodo puede, seguramente, ayudar a hacer más notorio ese efecto en mayor número de acciones.



cerca por los modelos EGARCH y GJR (los números en las celdas indican los casos en que cada modelo predice mejor que su competidor el comportamiento de las series accionarias).

La escasa diferencia existente entre los valores de  $CS$  arrojado por estas tres últimas especificaciones en el cuadro 7 impide, sin embargo, concluir que la distribución de *Student* sea la única alternativa posible para llevar a cabo el pronóstico del portafolio de los activos aquí incluidos. Cambios en la especificación de los modelos EGARCH o GJR, ya sea en su media o en su varianza, pueden mejorar su potencia descriptiva en relación con la  $t$  de *Student* y, por ende, ofrecer mejor base para el pronóstico del rendimiento-riesgo de los activos. Por esa razón, es más recomendable construir dos tipos de escenarios en los que se incluya, por una parte, un análisis del  $VeR$  del portafolio donde la función transformada sea la distribución de *Student* multivariada y, por la otra, un análisis usando alguna especificación de los modelos tipo GARCH. La adopción de cualquiera de los escenarios dependerá de la verificación del pronóstico arrojado por ambas especificaciones, ya sea mediante la prueba de cociente de fallas o de valores extremos como las sugeridas por Hull y White (1998) o Kupiec (1995).

#### **4. Una propuesta basada en cadenas de Markov para la elaboración de pronósticos de rendimientos de acciones con distribuciones no normales**

El ejercicio anterior deja en claro que la ausencia de normalidad en la distribución de rendimientos diarios supone diversos problemas técnicos que no siempre tienen una solución directa y convincente. Ni siquiera la introducción de procesos estocásticos en el pronóstico de estos rendimientos es una garantía de solución completa al problema de no normalidad estacionaria, debido a que la elección de cualquier proceso conlleva su grado de penalización. Hay, por ejemplo, modelos que son más eficientes que otros en capturar algunas distorsiones evidentes de las distribuciones empíricas, como la leptocurtosis, pero que no dan cuenta de las causas explicativas de las colas gordas, como la volatilidad grupal o modelos que ajustan mejor los desplazamientos cíclicos que los ordenados en el tiempo. Las virtudes de unos pueden aparecer como defectos en otros, y viceversa, por lo que no hay método perfecto ni único que garantice siempre el mejor pronóstico.

**Cuadro 6**  
*Modelos de heterocedasticidad condicional tempo-dependientes*  
**A) GJR**

	$\Omega$	Z	H	B	$\sqrt{b_1}$	$b_2$	Q	$Q^2$
1	0.208	0.146	0.031	0.822	0.679	5.770	8.999	8.291
	0.000	0.000	0.171	0.000			0.532	0.406
2	0.489	0.119	0.089	0.796	0.658	5.596	2.917	18.971
	0.000	0.000	0.000	0.000			0.983	0.041
3	0.046	0.150	-0.143	0.907	0.513	5.091	12.465	6.908
	0.000	0.000	0.000	0.000			0.255	0.734
4	0.172	0.133	-0.074	0.891	0.493	5.122	20.242	11.717
	0.000	0.000	0.000	0.000			0.027	0.304
5	0.133	0.050	0.020	0.916	0.389	5.009	12.120	12.946
	0.000	0.000	0.140	0.000			0.277	0.227
6	0.145	0.108	-0.100	0.919	0.670	5.986	6.521	13.082
	0.000	0.000	0.000	0.000			0.770	0.219
7	0.242	0.095	0.026	0.863	0.412	6.229	7.053	8.312
	0.000	0.000	0.102	0.000			0.720	0.598
8	0.482	0.129	0.003	0.814	0.151	8.026	16.230	28.583
	0.000	0.000	0.858	0.000			0.093	0.001

Cuadro 6  
A) (continuación)

	$\Omega$	Z	H	B	$\sqrt{b_1}$	$b_2$	Q	$Q^2$
9	0.393	0.132	-0.013	0.836	0.684	5.943	15.368	9.186
	0.000	0.000	0.533	0.000			0.119	0.515
10	0.129	0.104	0.014	0.889	0.846	8.523	16.078	5.451
	0.000	0.000	0.464	0.000			0.097	0.859
11	0.240	0.117	-0.072	0.893	0.798	7.297	8.575	13.691
	0.000	0.000	0.000	0.000			0.573	0.188
12	0.164	0.163	-0.121	0.875	0.343	4.316	6.627	14.290
	0.000	0.000	0.000	0.000			0.760	0.160
13	0.067	0.059	-0.056	0.961	-0.066	6.777	5.673	41.498
	0.000	0.000	0.000	0.000			0.842	0.000
14	0.986	0.139	0.109	0.679	0.635	5.173	11.515	16.214
	0.000	0.000	0.000	0.000			0.319	0.094
15	0.147	0.090	0.000	0.905	-0.228	13.587	8.874	3.704
	0.000	0.000	0.996	0.000			0.544	0.960
16	0.032	0.096	0.010	0.901	0.587	8.406	10.617	10.193
	0.000	0.000	0.382	0.000			0.388	0.424
17	0.587	0.130	-0.064	0.815	0.415	4.703	15.804	5.052

## Cuadro 6

## A) (continuación)

	$\Omega$	Z	H	B	$\sqrt{b_1}$	$b_2$	Q	$Q^2$
	0.000	0.000	0.006	0.000			0.105	0.888
18	0.164	0.073	0.007	0.905	0.572	6.626	8.141	7.702
	0.000	0.000	0.561	0.000			0.615	0.658
19	0.172	0.111	-0.080	0.902	0.676	6.289	9.255	9.173
	0.000	0.000	0.000	0.000			0.508	0.516
20	0.095	0.092	-0.065	0.915	0.522	4.568	13.779	12.693
	0.000	0.000	0.000	0.000			0.183	0.241
21	0.148	0.075	-0.103	0.951	0.515	6.268	18.071	14.607
	0.000	0.000	0.000	0.000			0.054	0.147
22	0.550	0.110	0.010	0.834	0.870	7.897	6.965	12.821
	0.000	0.000	0.624	0.000			0.729	0.234

## B) EGARCH

	$\omega$	$\gamma$	k	$\beta$	$\sqrt{b_1}$	$b_2$	Q	$Q^2$
1	-0.126	0.244	0.013	0.969	0.712	5.868	10.78	8.7441
	0.000	0.000	0.284	0.000			0.375	0.364

**Cuadro 6**  
**B) (continuación)**

	$\omega$	$\gamma$	$k$	$\beta$	$\sqrt{b_1}$	$b_2$	$Q$	$Q^2$
2	-0.089	0.324	-0.046	0.928	0.682	5.833	3.4543	18.924
	0.000	0.000	0.001	0.000			0.969	0.041
3	-0.135	0.193	0.091	0.987	0.47	4.969	7.8206	8.0672
	0.000	0.000	0.000	0.000			0.646	0.622
4	-0.100	0.190	0.049	0.981	0.455	4.64	20.018	14.354
	0.000	0.000	0.000	0.000			0.290	0.157
5	-0.039	0.102	0.013	0.980	0.431	5.319	12.689	19.743
	0.000	0.000	0.195	0.000			0.242	0.032
6	-0.093	0.166	0.057	0.983	0.588	5.552	6.237	22.826
	0.000	0.000	0.000	0.000			0.795	0.011
7	-0.095	0.193	-0.013	0.975	0.433	6.295	7.629	14.897
	0.000	0.000	0.208	0.000			0.655	0.136
8	-0.070	0.304	-0.015	0.925	0.111	8.733	17.872	22.19
	0.000	0.000	0.225	0.000			0.057	0.014
9	-0.065	0.263	0.020	0.939	0.669	5.924	14.535	7.272
	0.001	0.000	0.158	0.000			0.150	0.699
10	-0.108	0.245	0.007	0.970	0.939	9.871	17.086	5.204

**Cuadro 6**  
**B) (continuación)**

	$\omega$	$\gamma$	$k$	$\beta$	$\sqrt{b_1}$	$b_2$	$Q$	$Q^2$
	0.000	0.000	0.637	0.000			0.072	0.877
11	-0.101	0.180	0.062	0.984	0.763	7.076	8.814	17.926
	0.000	0.000	0.000	0.000			0.550	0.056
12	-0.109	0.217	0.085	0.968	0.359	4.391	7.496	18.459
	0.000	0.000	0.000	0.000			0.678	0.048
13	0.101	0.162	0.095	0.902	-0.324	9.28	4.598	16.611
	0.000	0.000	0.000	0.000			0.916	0.083
14	-0.003	0.321	-0.022	0.876	0.645	5.201	10.585	18.459
	0.924	0.000	0.153	0.000			0.319	0.048
15	-0.087	0.168	0.015	0.985	-0.087	11.711	9.505	4.573
	0.000	0.000	0.024	0.000			0.485	0.918
16	-0.120	0.202	0.003	0.982	0.563	7.871	10.270	11.409
	0.000	0.000	0.715	0.000			0.417	0.327
17	-0.019	0.190	0.044	0.936	0.416	4.596	16.136	7.430
	0.335	0.000	0.009	0.000			0.096	0.684
18	-0.015	0.145	0.021	0.957	0.607	7.660	9.926	17.882
	0.013	0.000	0.028	0.000			0.447	0.057

**Cuadro 6**  
**B) (continuación)**

	$\omega$	$\gamma$	$k$	$\beta$	$\sqrt{b_1}$	$b_2$	$Q$	$Q^2$
19	-0.075	0.167	0.062	0.972	0.685	6.506	8.462	12.689
	0.000	0.000	0.000	0.000			0.584	0.242
20	-0.083	0.154	0.046	0.975	0.532	4.575	11.944	12.792
	0.000	0.000	0.001	0.000			0.289	0.236
21	-0.074	0.138	0.033	0.986	0.594	5.445	15.135	15.804
	0.000	0.000	0.001	0.000			0.127	0.105
22	-0.025	0.244	-0.006	0.931	0.858	7.939	7.211	17.137
	0.208	0.000	0.660	0.000			0.705	0.071

## Cuadro 7

*Comparación de los modelos con base en el criterio de Schwarz*

<i>Modelo</i>	<i>EGARCH</i> (1, 1)	<i>GJR</i> (1, 1)	<i>Student</i>	<i>Difusión</i> <i>con salto</i>	<i>Mezcla ge-</i> <i>neralizada</i>
<i>EGARCH</i> (1,1)	...	8	16	6	0
<i>GJR</i> (1,1)	15	...	17	6	0
<i>Student</i>	6	5	...	6	0
<i>Difusión</i> <i>con salto</i>	16	16	16	...	0
<i>Mezcla</i>	22	22	22	22	...

En el caso de las acciones hay un elemento adicional que afecta la correcta especificación de los modelos: la presencia de patrones no lineales en el comportamiento de sus series. Con series no lineales, los modelos basados en la actualización de los dos primeros momentos de una distribución, como los ya vistos, pueden no ser muy eficientes en minimizar el efecto de la media y la varianza de períodos anteriores que no se acumulan linealmente en el período más reciente. Por esa razón es conveniente utilizar métodos estocásticos no lineales que aislen el impacto de la media y la varianza de ciertos períodos específicos, no necesariamente el más reciente, con el fin de encontrar un patrón de pronóstico más realista.

El modelo que se presenta a continuación conserva esa cualidad ya que, además de capturar los efectos no lineales sobre los parámetros, tiene la ventaja de evitar los problemas asociados con la falta de normalidad multivariada. La razón estriba en que con la introducción de cadenas de Markov las series de los rendimientos no requieren ajustarse al supuesto de normalidad estacionaria siempre y cuando éstas sean ergódicas o, dicho más técnicamente, que el conjunto de sus estados sean positivos recurrentes y aperiódicos. La idea se basa en los trabajos sobre ciclos económicos desarrollados por Neftci (1984) y Falk (1986), en los que la asimetría de los ciclos económicos es evaluada bajo la hipótesis de que si las series son simétricas a lo largo



del ciclo, entonces las probabilidades de transición entre los estados de una cadena de Markov asociada son supuestas iguales.

McQueen y Thorley (1991) extienden este razonamiento para argumentar que si los rendimientos siguen una caminata aleatoria entonces las probabilidades de transición de la cadena deben ser también iguales. En caso contrario, los rendimientos presentes pueden ser pronosticados con base en la probabilidad de los rendimientos bajos o altos registrados en períodos anteriores, mas no en los rendimientos pasados en sí.

En cualquier caso la estructura sugerida requiere, primero, de ordenar la serie de rendimientos de acuerdo con una secuencia  $I_t$  definida de la siguiente forma:

$$I_t = \begin{cases} 1 & \text{si } r_t > R, \\ 0 & \text{si } r_t < R, \end{cases}$$

donde  $P_t$  es el precio del activo en el tiempo  $t$ ,  $r_t = \ln(P_t/P_{t-m})$  su rendimiento ( $m = 1$  para diarios,  $m = 5$  para semanales, etc.) y  $R$  una medida de la esperanza matemática, que divide a los rendimientos en altos y bajos. Al considerar  $S_T = \{i_1, i_2, \dots, i_T\}$  como una realización de  $\{I_t\}$ , se puede definir a la función de verosimilitud de la cadena de Markov de segundo orden<sup>20</sup> como:

$$L(S_t) = P(I_2 = i_2, I_1 = i_1)$$

$$\prod_{s=0}^{T-3} P(I_{T-s} = i_{T-s} \mid I_{T-s-2} = i_{T-s-2}, I_{T-s-1} = i_{T-s-1})$$

la cual, una vez incorporadas las probabilidades de transición y las del estado inicial,

$$P(I_k = 0 \mid I_{k-2} = i, I_{k-1} = j) = \lambda_{ij}, \quad i, j = 0, 1$$

$$P(I_2 = i_2, I_1 = i_1) = \pi_0$$

la podemos reexpresar de la siguiente manera:

$$L(S_t, \Lambda, \pi_0) = \pi_0 \prod_{ij=00}^{11} \lambda_{ij}^{N_{ij}} (1 - \lambda_{ij}t)^{M_{ij}}, \quad i, j = 0, 1$$

---

<sup>20</sup> En la siguiente sección se justificará la elección del orden del proceso de Markov con fundamento en el criterio de información de Schwarz y en evidencia experimental.

$$\Lambda = [\lambda_{00} \quad \lambda_{01} \quad \lambda_{10} \quad \lambda_{11}]'$$

donde  $N_{ij}$  representa el número de secuencias de *ceros* y *unos*, tales que  $(I_t = 0/I_{t-2} = i, I_{t-1} = j), i, j = 0, 1$  y  $t = 3, 4 \dots T$ , y  $M_{ij}$  el número de secuencias con el patrón  $(I_t = 1/I_{t-2} = i, I_{t-1} = j), i, j = 0, 1$  y  $t = 3, 4 \dots T$ . O, escrita en su forma logarítmica (o función de log-verosimilitud):

$$\ln \{L(S_t, \Lambda, \pi_0)\} = \ln(\pi_0) + \sum_{ij=00}^{11} \{N_{ij} \ln(\lambda_{ij}) + M_{ij} \ln(1 - \lambda_{ij})\}$$

El problema de incluir las probabilidades del estado inicial en la función de log-verosimilitud es que se requiere de métodos numéricos molestos para obtener los estimadores máximo verosímiles. Por tal razón es necesario suponer que el número de observaciones es grande, tal como es el caso de este trabajo, para obviar el cálculo de esas probabilidades en la estimación.<sup>21</sup> De esta manera, al resolver las condiciones de primer orden del problema de maximización de  $\ln(L)$ , ignorando  $\pi_0$ , se obtienen los siguientes estimadores más simplificados:

$$\lambda_{ij} = \frac{N_{ij}}{N_{ij} + M_{ij}}, \quad i, j = 0, 1$$

Las matrices de ocurrencias y de probabilidades de transición del proceso de Markov de segundo orden resultantes de estos estimadores, se resumen en el cuadro 8.

**Cuadro 8**  
*Ocurrencias y probabilidades de transición*

<i>Estados anteriores</i>		<i>Ocurrencias</i>		<i>Probabilidades</i>	
		<i>estado actual</i>		<i>estado actual</i>	
		0	1	0	1
0	0	$N_{00}$	$M_{00}$	$\lambda_{00}$	$1 - \lambda_{00}$
0	1	$N_{01}$	$M_{01}$	$\lambda_{01}$	$1 - \lambda_{01}$
1	0	$N_{10}$	$M_{10}$	$\lambda_{10}$	$1 - \lambda_{10}$
1	1	$N_{11}$	$M_{11}$	$\lambda_{11}$	$1 - \lambda_{11}$

<sup>21</sup> Neftci (1984) demuestra que  $\pi_0$  se puede expresar como una función de los  $\lambda_{ij}$ 's, y que la estimación de las probabilidades de transición necesita de métodos iterativos

La pruebas orientadas a determinar la existencia de un patrón en las series de rendimientos, y que tienen como hipótesis nula el comportamiento de una caminata aleatoria, se reducen en el modelo a hacer un contraste de igualdad entre las probabilidades de transición, o a comparar las probabilidades  $\lambda_{00}$  y  $\lambda_{11}$ , como una opción sugerida por la posible presencia de burbujas racionales.

#### 4.1. Estimación y pruebas

Para estimar el modelo se tomó inicialmente como medida de expectativas, el promedio móvil de los rendimientos diarios del año anterior. Al utilizar esta media móvil, se construyó la serie  $\{I_t\}$  de unos y ceros, y se procedió a clasificar los rendimientos de acuerdo con los valores de las pasadas realizaciones. Con el fin de encontrar el grado óptimo de la cadena de Markov, se utilizó otra vez el criterio de información de Schwarz (*CS*).<sup>22</sup>

Con base en el cuadro 9, la función de log-verosimilitud de la cadena de primer orden es menor que la de segundo orden, además de que la diferencia entre los valores de *CS* de las cadena de primero y de segundo orden es mínima. Esto es, la pérdida de parsimonia debida al mayor número de parámetros a estimar en la cadena de segundo orden no hace menos verosímiles a sus estimadores. Si uno observa los resultados para la cadena de tercer orden, la pérdida de parsimonia es apreciable, por lo que el *CS* es considerablemente mayor a los demás, a pesar de que la función de log-verosimilitud evaluada en sus parámetros es la mayor de las tres. Esto permite pensar en la cadena de Markov de segundo orden como la adecuada para ajustar la serie de rendimientos.

---

<sup>22</sup> El *CS* se define ahora como

$$CS = \frac{-2 \ln(L)}{T} + \frac{k \ln(T)}{T}$$

$$k = 2^p + 2^{p-1}(2^p - 1)$$

donde  $p$  es el orden de la cadena de Markov,  $T$  el número de observaciones y  $\ln(L)$  la función de log-verosimilitud evaluada en los parámetros que la maximizan. El número de parámetros estimados,  $k$ , es la suma de los estimados de las probabilidades de transición (primer término de la suma) y de sus varianzas y covarianzas (segundo término).

**Cuadro 9**

*CS para distintos órdenes de la cadena de Markov*

$p$	$\ln(L)$	$T$	$k$	$CS$
1	-721.10	1291	3	1,1338
2	-699.22	1290	10	1,1396
3	-684.59	1289	36	1,2622

La naturaleza no lineal del patrón de las series de rendimientos es puesta de manifiesto en el cuadro 10, en donde las ocurrencias y las probabilidades de transición revelan una fuerte correlación positiva en los rendimientos de la cartera, lo que significa que la posibilidad de lograr un rendimiento bajo (alto) luego de haber tenido una secuencia de dos rendimientos bajos (altos) es mayor a la de alcanzar un rendimiento bajo (alto) después de dos altos (bajos). La significancia de esta correlación se confirma en el cuadro 11, en el cual queda claro que la hipótesis de la caminata aleatoria es enfáticamente rechazada a cualquier nivel de significancia según los estadísticos de razón de verosimilitud (RV) y de Wald.

**Cuadro 10**

*Ocurrencias y probabilidades de transición estimadas*

<i>Estados anteriores</i>		<i>Ocurrencias estado actual</i>		<i>Probabilidades estado actual</i>	
		0	1	0	1
0	0	160	107	0,5993	0,4007
0	1	67	105	0,3895	0,6105
1	0	106	104	0,6199	0,3801
1	1	104	576	0,1529	0,8471
$\sigma(\lambda_{00})= 0,02999$ $\sigma(\lambda_{01})= 0,03718$ $\sigma(\lambda_{10})= 0,03712$ $\sigma(\lambda_{11})= 0,01380$					

Con el fin de comprobar la robustez de los resultados se tomaron diversas medias móviles de 6, 15 y 18 meses, además de los resultados reportados a 12 meses, y se construyeron las correspondientes series  $I_t$ . En todas las pruebas se rechazaron las hipótesis nulas de igualdad

entre las probabilidades de transición y entre  $\lambda_{00}$  y  $\lambda_{11}$  para niveles de significancia marginal menores que el 1%, con lo que las estimaciones que se muestran en el cuadro 12 resultaron ser robustas ante diversas medidas de expectativas utilizadas.

**Cuadro 11**

*Pruebas de hipótesis sobre las probabilidades de transición del proceso de Markov de segundo orden*

$H_0$	$RV$	$p$ -value	Wald	$p$ -value	$\chi^2_1(5\%)$	$\chi^2_1(1\%)$
$\lambda_{00} = \lambda_{11}$	179.53	0.00	182.75	0.00	3.84	6.63
$H_0$	$RV$	$p$ -value		$\chi^2_3(5\%)$	$\chi^2_3(1\%)$	
$\lambda_{00} = \lambda_{01} = \lambda_{10} = \lambda_{11}$	253.31	0.00		7.82	11.34	

4.2. *Probabilidades estacionarias, ergodicidad y pronóstico*

Como ya se mencionó, la prueba aquí propuesta no requiere que la serie original de rendimientos sea NMV, siempre y cuando la cadena de Markov asociada sea érgodica. Para comprobar esta propiedad es necesario calcular primero las probabilidades estacionarias de la matriz, ya que con base en ellas se puede demostrar la ergodicidad de la cadena y, por tanto, determinar si ésta es estacionaria en covarianza. Los valores de estas probabilidades ( $\pi_{00} = 0.2624, \pi_{11} = 0.1975, \pi_{01} = \pi_{10} = 0.2699$ ) se encontraron con arreglo al sistema  $\Pi(t) = A^n \Pi(t - n)$ , en el que el vector de probabilidades estacionarias se definió como  $\lim_{n \rightarrow \infty} \Pi(t) = \Pi^*$ , o en forma compacta como:

$$\Pi^* = A\Pi^*$$

$$1'\Pi^* = 1$$

Ahora bien, estadísticamente se dice que la cadena es érgodica si los unos de los valores propios de la matriz es igual a la unidad y los restantes son menores a uno en valor absoluto (o en módulo, de ser valores complejos). La matriz de transición en este caso es de rango 4, por lo que se espera encontrar 4 valores propios, teniendo en cuenta que éstos son invariantes ante permutaciones de las columnas de la matriz de transición. El cuadro 12 presenta los valores propios y absolutos.

**Cuadro 12**  
*Valores propios de la matriz de transición*

<i>Valor propio</i>	<i>Multiplcicidad</i>	<i>Valor absoluto o módulo</i>
1	1	1
0.3294	1	0.3294
-0.2886 + 0.4628 i	1	0.5454
-0.2886 - 0.4628 i	1	0.5454

La matriz tiene 4 valores propios distintos, dos reales y un conjugado complejo, esto es: uno de los valores propios es uno, y los restantes son menores a la unidad, ya sea en valor absoluto o en módulo; por lo que la cadena es ergódica y, por ello, estacionaria en covarianza. De esta manera se concluye que la prueba se encuentra bien especificada y que es posible hacer pronósticos con base en la cadena de Markov.

El pronóstico o valor esperado de  $\varepsilon_{t+m}$ , condicionado en la información en  $t$  para cualquier realización  $S_t = ij$ , se calcula como:  $E[\varepsilon_{t+m} | \varepsilon_t, \varepsilon_{t-1}, \dots] = L^m \varepsilon_t$ , donde  $L$  es la columna de la matriz de transición del proceso de Markov de primer orden. De esta forma, se pueden pronosticar las probabilidades de transición de la cadena en  $t + m$  elevando a la potencia  $m$  la matriz de transición y multiplicándola por el correspondiente vector de la base canónica de  $R^4$ ,  $e_{ij}$ , usando el ordenamiento de columnas de la matriz del proceso de Markov de primer orden. Así, por ejemplo, el pronóstico  $m$  periodos adelante dado que en  $t$  la cadena se encontraba en el estado  $S_t = 11$  es igual a:

$$\begin{bmatrix} P[S_{t+m} = 11 | S_t = 11] \\ P[S_{t+m} = 00 | S_t = 11] \\ P[S_{t+m} = 10 | S_t = 11] \\ P[S_{t+m} = 01 | S_t = 11] \end{bmatrix} = L^m e_{11}$$

En virtud de que la cadena de Markov asociada a la serie de rendimientos es ergódica las probabilidades pronosticadas tenderán a convergir en sus probabilidades estacionarias. Los valores de convergencia de los rendimientos diarios del paquete *Banamex-30* a sus probabilidades estacionarias se alcanzaron alrededor de  $m=10$ , lo cual indica que el horizonte de pronóstico de este modelo es muy corto y que las predicciones que se pueden hacer son muy inmediatas (de hecho antes de asentarse en sus valores estacionarios). Por esta razón

y recordando que los datos utilizados en este trabajo corresponden a días hábiles, uno puede asegurar con confianza que el horizonte de pronóstico del modelo es de aproximadamente dos semanas: una ventana igual a la recomendada por RiskMetrics<sup>MT</sup> (1995, 1996).

## 5. Conclusiones

El documento pone énfasis en el hecho, muchas veces soslayado por la práctica financiera, de que la evaluación del supuesto de normalidad estacionaria es esencial para formular un análisis adecuado del riesgo-rendimiento de un activo o de un portafolio. Los resultados obtenidos en las dos primeras partes del documento confirman que los rendimientos diarios para las principales acciones mexicanas no son normales estacionarios, no así los mensuales, y que, por tanto, el análisis de VeR del paquete *Banamex-30* requiere de modelos alternativos para el cálculo y pronóstico de su media y varianza. Los modelos más descriptivos de las series de rendimientos no normales estacionarios resultaron, de acuerdo con el criterio de Schwarz, los modelos de *Student*, EGARCH, en buena medida parecido al DGE recomendado por RiskMetrics<sup>MT</sup> y el llamado modelo GJR. La conclusión sugerida en el texto es que la adopción de cualquiera de las tres especificaciones requiere de pruebas de verificación de sus estimaciones sobre el VeR antes de optar por alguno de ellos, ya que aun cuando la distribución de *Student* resultó ser el modelo con más alta probabilidad de explicar las series no normales, los otros dos podrían mejorar su capacidad descriptiva al combinarlos con otros modelos de media y varianza. La selección de determinado modelo debe hacerse sobre bases conceptuales.

La contribución del trabajo se centra en la tercera parte del documento, en la que se propone un modelo que pronostica los patrones no lineales de las series de rendimientos, sin hacer uso de los métodos tradicionales de media y varianza condicional actualizada y sin suponer normalidad multivariada. El modelo busca probar la hipótesis de la caminata aleatoria para los rendimientos con base en cadenas de Markov de alto orden, mediante la expansión de la metodología propuesta originalmente por Neftci (1984), Falk (1986) y McQueen y Thorley (1991).

Después de asociar un proceso de Markov de segundo orden a la serie de rendimientos diarios de la cartera, se encontró que la estructura de las probabilidades de transición del proceso revelaba una fuerte autocorrelación positiva, lo cual implicaba la existencia de

componentes predecibles no lineales en la serie de los rendimientos de la cartera. La significancia estadística de este resultado se probó mediante diversos métodos, encontrándose en todos los casos que la hipótesis nula de la caminata aleatoria era enfáticamente rechazada. Los resultados de estas pruebas se mostraron robustos ante distintas medidas de expectativas. También se demostró que la cadena asociada a la serie de rendimientos era ergódica y por tanto estacionaria en covarianza, por lo que las pruebas fueron bien especificadas. Por último se encontró que es posible obtener pronósticos fiables con un horizonte relativamente corto. El resultado encontrado se ajusta a las recomendaciones hechas por RiskMetrics<sup>MT</sup> (1995,1996) y significa que, de no haber alteraciones en la estructura de las series, la correlación positiva de los rendimientos indicada por la cadena de segundo orden alcanzará sus valores estacionarios en sólo 10 días.

## Bibliografía

- Affleck-Graves J. y B. McDonald (1989). "Nonnormalities and Test of Asset Pricing Theories", *The Journal of Finance*, 44, pp. 889-908.
- Ball C. A. y W. N. Torous (1985). "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing", *The Journal of Finance*, 40, pp. 155-173.
- Blattberg R. y N. Gonedes (1974). "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices", *Journal of Business*, 47, pp. 244-280.
- Campbell, J. Y., Lo, A. W. y A. C. MacKinlay (1997). *The Econometrics of Financial Markets*, Princeton University Press.
- Duffie, D. y J. Pan (1997). "An Overview of Value at Risk", *Journal of Derivatives*, vol. 4, núm. 3, pp. 7-49.
- Falk, B. (1986). "Further Evidence on the Asymmetric Behavior of Economic Time Series over the Business Cycle", *Journal of Political Economy*, 94, pp. 1906-1109.
- Fama, E. (1965). "The Behavior of Stock Market Prices", *Journal of Business*, 38, pp. 34-105.
- Hull, J. y A. White (1998). "Value at Risk When Daily Changes in Market Variables are not Normally Distributed", *The Journal of Derivatives*, vol. 5, núm. 3, pp. 9-19.
- Johnson, R. A. y D. W. Wichern (1992). *Applied Multivariate Statistical Analysis*, Englewood Cliffs, Prentice Hall, New Jersey.
- Kaskey, G. et al. (1980). "Transformations to Normality", *Handbook of Statistics*, vol 1, North-Holland.



- Kim, D. y S. Kon (1994). "Alternative Models for the Conditional Heteroscedasticity of Stock Returns", *Journal of Business*, 67, pp. 563-598.
- Kon, S. J. (1984). "Models of Stock Returns-A comparison", *Journal of Finance*, 39, pp. 147-165.
- Kupiec, P. H. (1995). "Techniques for Verifying the Accuracy of Risk Measurement Models", *The Journal of Derivatives*, 3, Winter, pp. 73-84.
- Li, D. X. (1999). *Value at Risk Based on the Volatility, Skewness and Kurtosis*, Riskmetrics Group, New York.
- Lo, A. W. (1997). "Fat Tails, Long Memory, and the Stock Market Since the 1960's", *Economic Notes by Banca Monte dei Paschi*, vol 26 (2), pp. 213-246.
- y C. Mackinlay (1990). "Data Snooping Biases in Tests of Financial Asset Pricing Models", *Review of Financial Studies*, 3, pp. 431-468.
- (1988). "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test", *Review of Financial Studies* 1, pp. 41-66.
- Mandelbrot, B. (1963). "The Variation of Certain Speculative Prices", *Journal of Business*, 36, pp. 394-419.
- Mardia, K. V. (1980). "Test of Univariate and Multivariate Normality", en P. R. Krishnaiah (comp.), *Handbook of Statistics*, vol. 1, North-Holland, Amsterdam.
- McQueen, G. y S. Thorley (1991). "Are Stock Returns Predictable? A Test Using Markov Chains", *The Journal of Finance*, vol. 46 (1), March, pp. 239-263.
- Neftci, S. N. (1984). "Are Economic Time Series Asymmetric over the Business Cycle?", *Journal of Political Economy*, 92, pp. 307-328.
- Richardson, M. y T. Smith (1993). "A Test for Multivariate Normality in Stock Returns", *Journal of Business*, 66, pp. 295-321.
- RiskMetrics<sup>TM</sup> (1996). *Technical Document*, J. P. Morgan.
- (1995). *Technical Document*, J. P. Morgan.
- Sánchez Cerón (2001). *Valor en riesgo y otras aproximaciones*, Valuación, análisis y riesgo A.C., México.
- Zangari, P. (1996). "Test of Conditional Normality", *RiskMetrics<sup>TM</sup>-Technical Document*, J. P. Morgan/Reuters.
- Zenios S. A. (1998). *Financial Optimization*, Cambridge University Press.