Andrés E. Reyes Polanco
Prof: Asociado U.C.V-ESGN

ESTADÍSTICA Y MODELOS FINANCIEROS

Andrés E Reyes Polanco.[1]

El siguiente artículo trata sobre el uso de la estadística en los modelos financieros utilizados en la selección óptima de carteras. Se efectúa una exposición teórica preliminar y posteriormente se evalúa los aspectos teóricos de la estadística subyacente a estos modelos y, su posibilidad real de aplicación en el mercado de capitales venezolano.

INTRODUCCIÓN.

En el siguiente trabajo se abre una discusión sobre los problemas que suelen presentarse cuando se emplean modelos que parten de estimaciones de parámetros poblacionales mediante observaciones muestrales de series temporales relacionadas con el mercado de capitales de Venezuela. Entre los modelos que se tratan de aplicar están el formulado por Markowitz, la simplificación propuesta por W. F. Sharpe y los modelos de arbitraje. En este contexto se asume que los rendimientos de los títulos son variables aleatorias a igual que los índices del mercado de capitales y otros indicadores macroeconómicos que se han propuesto para formular este tipo de modelos cuantitativos para efectuar una selección óptima de portafolio.

El modelo Markowitz, se formula en forma general como sigue: se minimiza el riesgo de la cartera, medido este en término de varianza de la misma, sujeto a las restricciones presupuestarias y a un valor esperado de rendimiento de la cartera. Esto es:

Si R_i es el rendimiento del título i-ésimo, el cual tiene un comportamiento aleatorio, esto es, una variable con valor esperado E_i y varianza $\sigma_i^2 < \propto$ y además asumimos que existe covariación entre los títulos, es decir, existen covarianzas σ_{ij} y además consideramos la proporción del presupuesto que debe destinarse en la adquisición del título i-ésimo, se representa por x_i. De lo anteriormente expuesto obtenemos que el rendimiento de la cartera es:

$$\mathbf{R_c = x_1R_1 + x_2R_2 + \ldots + x_nR_n}$$

Donde:

$$\mathbf{x_1 + x_2 + x_3 + \ldots + x_n = 1;\ x_1 => 0}$$

[1] Este artículo apareció en **Reyes Polanco, A. E (2000)**
Estadística y Modelos Financieros ,Cuadernos de Postgrado nº 15 FACES-UCV. Modificado Enero 2007

Como puede observarse R_c es combinación lineal de n variables aleatorias, por tanto ella también es una variable aleatoria con media y varianza dada por:

$$\mathbf{E(R_c)=x_1E_1+x_2E_2+\ldots+x_nE_n}$$

El inversionista seleccionará uno o varios valores de $\mathbf{E(R_c)=E^*}$**,** que satisfaga su expectativa.

La varianza del rendimiento de la cartera es $\mathrm{var}(R_c)=\boldsymbol{\sigma_c^2=\Sigma x_i^2\sigma_i^2+2\Sigma\Sigma x_i\, x_i\, \sigma_{ij}}$

El modelo se expresa como:

Min
$$\boldsymbol{\sigma_c^2=\Sigma x_i^2\sigma_i^2+2\Sigma\Sigma x_i\, x_i\, \sigma_{ij}}$$
Sujeto a:

$$\boldsymbol{\Sigma x_iE_i=E^*;\quad \Sigma x_i=1\ ;\ x_i =>0 \text{ para todo } i=1,2 \ldots n}$$

En este modelo habrá que estimarse los parámetros involucrados en la formulación partiendo de datos históricos del comportamiento de los títulos en el mercado financiero para luego resolverlo como un modelo de programación cuadrática en donde las variables de decisión son x_i las obteniéndose un valor óptimo para cada E*. Como la función objetivo del modelo no tiene un término lineal es fácil resolverlo aplicando directamente los multiplicadores de Lagrange lográndose la solución del sistema de ecuaciones lineales obtenido, con la aplicación directa de la matriz inversa de los coeficientes de las ecuaciones. Este modelo está formulado bajo el supuesto que el mercado de capitales sea eficiente.

El modelo simplificado de Sharpe parte de considerar que existe una relación entre cada rendimiento y uno o varios índices cuyos comportamientos afectas estos rendimientos dando origen al modelo de regresión simple cuando se considera un solo índice o un modelo de regresión múltiple cuando se emplean varios índices como variables explicativas.

Cuando se emplean modelos de regresión, debe chequearse si se cumplen los supuestos del modelo de regresión estándar para obtener los estimadores óptimos de los parámetros y poder realizar los contrastes de hipótesis que le interesa al analista y construir los intervalos de confianzas. Cuando se trata de índice único suele emplearse variables como: índice de la bolsa, tasa de cambio, producto nacional bruto, tasa de interés pasiva o activa etc.

En variadas ocasiones los supuestos del modelo estándar no se cumplen lo que obliga al analista usar un método de estimación distinto de los mínimos cuadrados ordinarios, sin embargo no siempre se tiene tal cuidado y lo más frecuente es que se emplean los mínimos cuadrados ordinarios para estimar el modelo de regresión lineal simple o múltiple sin considerar los supuestos subyacentes en la formulación de tales modelos.

Entre los supuesto que menos se chequean están el de rango completo de la matriz de datos cuando el modelo considera varios índices, la homoscedasticidad de las perturbaciones aleatorias, ausencia de autocorrelación de las perturbaciones aleatorias, supuesto de normalidad. La violación de uno o más de los supuestos que se parte al formular el modelo estándar trae varios problemas uno mas grave que otros que puede invalidar las estimaciones obtenidas bien sea con fines explicativos o predictivo o ambos, en otras palabras: De estos modelos así obtenidos no servirán para explicar el comportamiento de los rendimientos de los títulos y tampoco para predecir el comportamiento de los mismos a partir del comportamiento del índice o índices empleados como regresores.

El modelo de regresión múltiple propuesto por Shepar tiene la siguiente especificación:

$$\mathbf{R}_i=\beta_{i0}+\beta_{i1}\mathbf{I}_1+\beta_{i2}\mathbf{I}_2+\ldots+\beta_{ik}\mathbf{I}_k+\varepsilon_i$$

Para cada título i-ésimo obtenemos un vector de observaciones del rendimiento histórico del título correspondiente de T componentes representado por $\mathbf{R}_i$, los vectores $\mathbf{I}_1, \mathbf{I}_2, \ldots \mathbf{I}_k$ de T componentes cada uno de las observaciones históricas observables de los índices seleccionados, el vector de parámetros de k+ 1 componentes $\boldsymbol{\beta}_i$ y finalmente, el vector de perturbaciones aleatorias no observables también de T componentes $\boldsymbol{\varepsilon}_i$.

En notación matricial $\mathbf{R}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i$. Donde $\mathbf{X} = [\mathbf{1} \mid \mathbf{I}^*]$ es una matriz formada por el vector de unos $\mathbf{1}^t=(1,1....1)$, $\mathbf{I}^*$ es la matriz de datos formada por los vectores $\mathbf{I}_j$ columnas j=1,2,3...k.

Se asume que este modelo cumple con todos los supuestos del modelo estándar que mencionamos anteriormente y sobre los cuales volveremos.

Tanto los rendimientos de los títulos como los índices seleccionados son variables aleatorias observables, mientras que $\boldsymbol{\varepsilon}_i$ es un vector de variables aleatorias no observables que como se indicó en el párrafo anterior recoge los otros múltiples factores que afectan aditivamente de los rendimientos, pero que separadamente el efecto de cada uno de ellos es casi nulo o despreciable.

Como puede observarse tanto el valor esperado y la varianza de cada rendimiento está en relación de los valores esperados y varianzas de los indicadores considerados en el modelo de regresión.

Esto es:

$$\mathbf{E_i = E(\beta_{i\,0}+\beta_{i\,1}I_1+\beta_{i\,2}I_2+\ldots+\beta_{i\,k}I_k+\varepsilon_i)} \qquad \textbf{(1)}$$

$$\sigma_i^2=\mathbf{Var(\beta_{i\,0}+\beta_{i\,1}I_1+\beta_{i\,2}I_2+\ldots+\beta_{i\,k}I_k+\varepsilon_i)} \qquad \textbf{(2)}$$

Bajo el supuesto del modelo de regresión estándar, la esperanza de ε_i es nula, para cada uno de sus componentes, esto es, $\mathbf{E(\varepsilon_i)=0}$. La $\mathbf{Var(\varepsilon_i)}$ es igual para cada periodo t en donde se puede observar el rendimiento de cada título, esto es, $\mathbf{Var(\varepsilon_{t\,i})}$ es constante para todo t, o dicho de otra forma, las perturbaciones aleatorias asociada a cada título cumplen con el supuesto de homoscedasticidad. Por otra parte, se asume que no existe correlación entre dos perturbaciones aleatorias en periodos distintos del mismo rendimiento i-ésimo: $\mathbf{corr(\varepsilon_{t\,i}, \varepsilon_{t+k,,i})}$= : $\mathbf{corr(\varepsilon_{t-k,,i}, \varepsilon_{t-,i})}=0$. Además se parte de que las correlaciones entre las variables $\mathbf{I_j}$ j=1,2,3...k y ε_i , y dos vectores de perturbaciones de títulos diferentes $\boldsymbol{\varepsilon_i, \varepsilon_k}$ son nulas.

Resumiendo tenemos los siguientes supuestos:

1.- Las variables $\mathbf{I_j}$ son linealmente independientes es decir la matriz **X** es de rango completo, pero no necesariamente son estocásticamente independientes.

2.- $\mathbf{E(\varepsilon_{t\,i})=0}$ para todo t del título i-ésimo.

3.- $\mathbf{Var(\varepsilon_{t\,i})=}$es una constante para todo t del título i-ésimo **:** γ_i^2

4.-Los γ_i^2 /= γ_j^2 , i =/j es decir, la que mide el riesgo específico de cada título una vez separado el efecto de los indicadores deben ser distintos entre sí.

5.- $\mathbf{corr(\varepsilon_{t-k,,i}, \varepsilon_{t-,i})}=0$.

6.- $\mathbf{Corr(I_i\, I_j)=0}$

El supuesto 3 y 4 puede expresarse matricialmente como: $\gamma_i^2\ \mathbf{I}$ donde **I** es la matriz identidad.

7.-Se asume que el vector de perturbaciones cumple con el supuesto de normalidad esférica: $\mathbf{N(0}, \gamma_i^2\ \mathbf{I)}$**.** Si se cumple este supuesto entonces podrá usarse las pruebas clásicos del modelo de regresión estándar e igualmente se podrá construir los intervalos de confianzas.

Este último supuesto, en el caso del mercado de valores de Venezuela queda entre dicho, más plausible es considerar distribuciones asimétricas tipo beta o normales contaminadas, es decir, con altas kurtosis Esto último sugiere junto con el no cumplimiento de rango completo de la matriz **X,** emplear métodos de estimación tipo $\mathbf{L_1}$, que consiste en minimizar la suma de los valores absolutos de las desviaciones en vez de la suma de los cuadrados de las desviaciones que da lugar a los mínimos cuadrados ordinarios. El problema del uso del estimador $\mathbf{L_1}$ es que debe contarse con un software que resuelva modelos de programación lineal. Existe en el mercado el SAS que permite el análisis estadístico de los datos con modelos univariantes y multivariantes y además tiene rutinas de programación matemática, incluida programación lineal.

Puede consultarse los elementos básicos del estimador $\mathbf{L_1}$ en Reyes P, Andrés (1984). De no cumplirse los supuestos 3 y 5 deberán emplearse los métodos estadísticos apropiados; para revisar este último punto, puede consultarse en otros textos el de A. Novales (1993).

De todo lo dicho anteriormente, puede desarrollar las expresiones (1) y (2). La primera queda:

$$\mathbf{E_i = \beta_{i0} + \beta_{i1} (E\, I_1) + \beta_{i2} (E\, I_2) + ... + \beta_{ik} (E\, I_k)}$$

Si los índices $\mathbf{I_j}$ no están correlacionados entre sí, situación que podemos decir que en el caso venezolano no se da: entonces:

$$\sigma_i^2 = \beta^2_{i1}\,\mathbf{Var\,(I_1)} + \beta^2_{i2}\,\mathbf{Var\,(I_2)} + ... + \beta^2_{ik}\,\mathbf{Var\,(I_k)} + \gamma_i^2$$

Si los índices están correlacionados:

$$\sigma_i^2 = \beta^2_{i1}\,\mathbf{Var\,(I_1)} + \beta^2_{i2}\,\mathbf{Var\,(I_2)} + ... + \beta^2_{ik}\,\mathbf{Var\,(I_k)} + \beta_{i1}\beta_{i2}\,\mathbf{covar\,(I_1, I_2)} + ... + \beta_{i1}\beta_{ik}\,\mathbf{covar\,(I_1, I_k)} + ... + \beta_{ik-1}\,\beta_{ik}\,\mathbf{Var\,(I_{k-1}, I_k)} + \gamma_i^2$$

Para simplificar la expresión anterior podemos escribirla como sigue:
$\sigma_i^2 = \beta_i^t \mathbf{V} \beta_i + \gamma_i^2$,donde $\beta_i^t = (\beta_{i1}\beta_{i2...}\,\beta_{ik})$ y **V** es la matriz de varianza covarianza de los títulos $\mathbf{I_j}$, γ_i^2 es la varianza de las perturbaciones aleatorias.

Tomemos ahora el caso en que los rendimientos dependen solo de un índice el cuál se conoce como modelo de índice único MIU. Esto es:

$$\mathbf{R_i = \beta_{i0} + \beta_{i1} I_1 + \varepsilon_i} \qquad \mathbf{(1)}$$

El valor que tome β_{i1} permitirá clasificar el rendimiento del título i-ésimo como: volátil, poco volátil, muy volátil por tanto este parámetro recibe el nombre de coeficiente de volatilidad. $\beta_{i1}=1$ se considera la volatilidad como normal, si $\beta_{i1} <1$ entonces el título es poco volátil y finalmente un $\beta_{i1}>1$ tendrá una alta volatilidad.
Como media y varianza obtenemos:

$$\mathbf{E_i = \beta_{i0} + \beta_{i1} (E\, I_1)} \qquad \mathbf{(2)}$$

$$\sigma_i^2 = \beta^2_{i1}\,\mathbf{Var\,(I_1)} + \gamma_i^2 \qquad \mathbf{(3)}$$

El riesgo del título se ha descompuesto en dos sumando, el primero mide el riesgo sistemático debido al patrón asociado al índice considerado como variable explicativa y el segundo el riesgo específico.

Partiendo que el rendimiento de la cartera es:

$$\mathbf{R_c=x_1R_1+x_2R_2+\ldots+x_nR_n} \qquad \textbf{(4)}$$

Sustituyendo cada rendimiento en (4) por lo dado en (1) obtenemos:

$$\mathbf{R_c= x_1 (\beta_{1\,0}+\beta_{1\,1}I_1+ \varepsilon_1)+ x_2 (\beta_{20}+\beta_{2\,1}I_1+ \varepsilon_2)....+ x_n (\beta_{n\,0}+\beta_{n\,1}I_1+ \varepsilon_n)=}$$
$$\mathbf{\Sigma_i\, x_i\, \beta_{i\,0}+ \Sigma_i\, x_i\, \beta_{i\,1}\, I_1+\Sigma\, x_i\, \varepsilon_1 = \Sigma_i\, x_i\, \beta_{i\,0}+ I_1\Sigma_i\, x_i\, \beta_{i\,1} +\Sigma\, x_i\, \varepsilon_1} \qquad \textbf{(5)}$$

Tomando esperanza matemática y considerando la expresión (2) obtenemos:

$$\mathbf{E(R_c) = x_1 (\beta_{10}+\beta_{i\,1}\, E(I_1))+ x_2 (\beta_{2\,0}+\beta_{2\,1}\, E(I_1))\ldots\ldots + x_n (\beta_{n\,0}+\beta_{n\,1}\, E(I_1))}$$

$$\mathbf{E(R_c) =\Sigma_i\, x_i\, \beta_{i\,0}+ E(I_1)\Sigma_i\, x_i\, \beta_{i\,1}}$$

Tomando varianza y considerando la expresión (3) y (5) obtenemos:

$$=(\mathbf{\Sigma\, x_i\, \beta_{i\,1}})^2\, \mathbf{Var(I_1)+ \Sigma\, x_i^{\,2}\gamma_i^2}$$

Retomemos el problema de múltiples índices:

$$\mathbf{R_i=\beta_{i\,0}+\beta_{i\,1}I_1+\beta_{i\,2}I_2+...+\beta_{i\,k}I_k+\varepsilon_i}$$

$$\mathbf{R_i= x_1 (\beta_{10}+\Sigma\beta_{1\,j}I_j+\varepsilon_1)+x_2 (\beta_{2\,0}+\Sigma\beta_{2\,j}I_j+\varepsilon_2) +x_3 (\beta_{3\,0}+\Sigma\beta_{3\,j}I_j+\varepsilon_3)+...+ x_n (\beta_{n0}+ \Sigma\beta_{ij}I_j+\varepsilon_n)}$$

$$\mathbf{R_i=\Sigma_i\, x_i\, \beta_{i\,0}+\Sigma_i\, x_i\, \Sigma_j\, \beta_{i\,j}I_j+\Sigma_i\, x_i\, \varepsilon_i}$$

Tomando esperanza matemática y varianza obtenemos:

$$\mathbf{E(R_i)=\Sigma_i\, x_i\, \beta_{i\,0}+\Sigma_i\, x_i\, \Sigma_j\, \beta_i\, jE(I_j)}$$

Asumiendo que **Corr(I_i I_j)=0**

$$\mathbf{Var(R_i)=\Sigma_j\, (\Sigma\, i\, x_i\, \beta_i\, j)^2Var(I_j) +\Sigma_i\, x_i^{\,2}\gamma_i^2}$$

Dado que nos interesa el aspecto estadístico subyacente en estos modelos no entraremos en al discusión de cómo se obtienen carteras eficientes partiendo de cada uno de los modelos. Para ello pude consultarse las siguientes obras: W. F. Sharpe 1974, A. S. Suarez 1977 Messuti et al 1994.

Pasaremos ahora a considerar la naturaleza estadística del mercado de valores de Vzla. Y de acuerdo a esta naturaleza el buen uso de estos modelos para la toma de decisiones.

II.- COMENTARIOS GENERALES SOBRE LAS CARACTERÍSTICAS DE LAS VARIABLES IMPLÍCITAS DE LOS MODELOS.

Antes de analizar directamente los supuestos con los que se construyen estos modelos debemos ver cual es el comportamiento estadístico de los datos históricos de las variables utilizadas para alimentar los modelos brevemente expuestos en la introducción.

En primer lugar tenemos la variable rendimiento de cada título. La primera hipótesis que planteamos es que su comportamiento histórico responde a un patrón perfectamente identificable y que es válido en periodos futuros más un elemento aleatorio cuyo efecto es casi despreciable, lo que permitirá predecir los futuros valores de este rendimiento o por el contrario la hipótesis que establecemos es que el comportamiento de este rendimiento es altamente sensible a las condiciones iniciales, es decir, responde a un comportamiento caótico determinístico. De ser cierta la última hipótesis lo más que podrá hacerse es predicciones muy a corto plazo y tratar de descubrir al atractor extraño, es decir, un atractor de dimensión fractal que está asociado a ese comportamiento caótico.

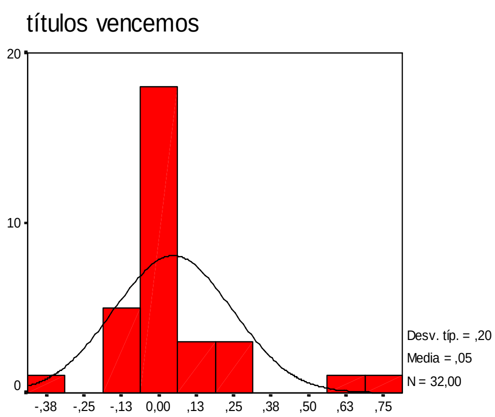
Si asumimos como plausible la primera hipótesis, es decir, que existe un patrón, debemos determinar el modelo que mejor se ajuste a estos datos históricos, si se determina que el comportamiento sigue una relación funcional perfecta sin ningún margen de error, el riego es nulo. Pero en practica el elemento aleatorio está presente en el rendimiento de los títulos y por consiguiente es perfectamente válido hablar de esperanza y varianza del rendimiento. Por otra parte, la variabilidad que se presenta en el rendimiento de los títulos medido como: (valor de cierre – valor apertura)/valor de apertura hace presumir que el riesgo puede resultar alto si se toma como la medida del mismo la varianza. Pero esto trae como consecuencia una mayor incertidumbre del inversionista ya que la varianza está afectada por valores extremos.

Otro problema que se presenta en el caso de Venezuela es el cambio de metodología que se emplean para obtener los índices económicos originando datos no homogéneos cuando se consideran series de tiempo de más de 20 observaciones anuales.

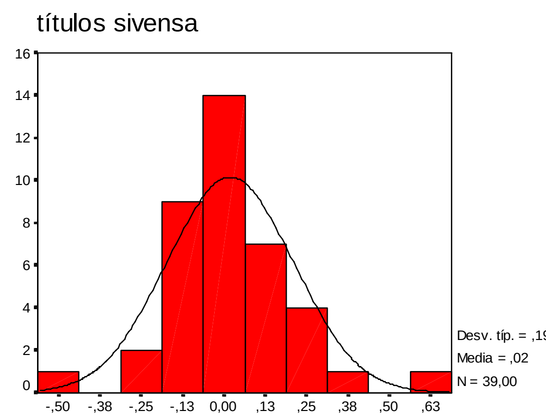
Un caso de estos el el índice de precio que suele emplearse una vez mensualizado para emplearlo como variable explicativa del comportamiento de los rendimientos tanto en modelos de regresión simple (modelo unifactorial o de índice único) o en modelos de regresión múltiple (modelo multi-factoral). Ocurre igualmente con la periodicidad en que se miden los índices: unos se pueden obtener mensualmente, mientras que otros son trimestrales, semestrales o anuales y además están sujeto a revisiones haciendolo de poca utilidad práctica cuando se trata de construir modelos asociados al mercado de capitales.

Recientemente la economía venezolana en su conjunto muestra muy poca estabilidad tanto a nivel macro como microeconómico lo que dificulta la utilización de los modelos clásicos construidos con fines de pronósticos o explicativos tales como modelos de regresión lineal o modelos del tipo ARIMA(p,d,q), y en el caso del mercado de capitales los modelos de selección propuestos por Markowitz y la modificación sugerida por Sharpe parten del supuesto de un mercado de capitales eficiente y resulta en primer lugar que ese mercado en el caso venezolano es pequeño tanto por la cantidad de oferentes y demandantes como por el volumen negociado y en segundo lugar el uso de la varianza para medir el riesgo que como indicamos puede no ser **Corr(I $_i$ I $_j$)=0** lo mas apropiado.

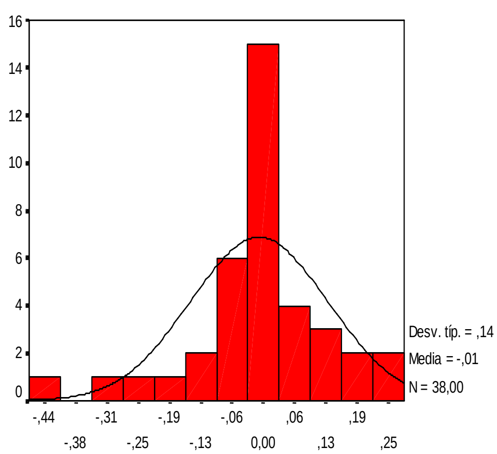
Algunos datos que a continuación presentamos ilustran esta situación. En primer lugar, tenemos un conjunto de histograma que ilustran el comportamiento de algunos títulos:



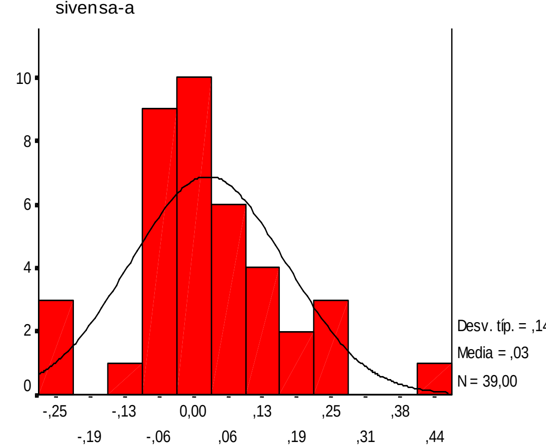
vencemos 2



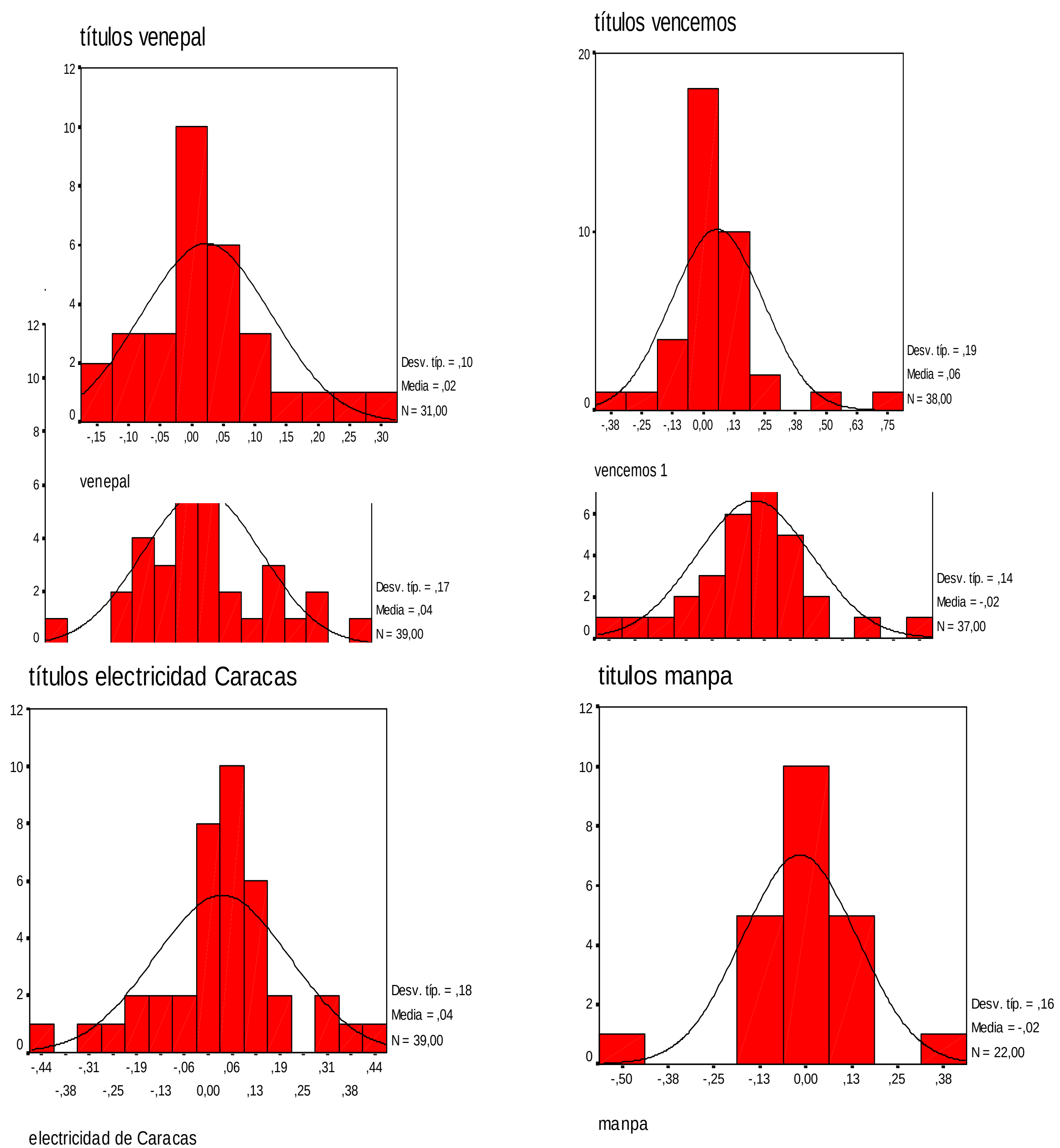
sivensa-a



sudantex-b



banco provincial



Los histogramas anteriores corresponden a los rendimientos de diez empresas muy importantes que cotizan en la bolsa electrónica de Caracas. Puede observarse que en general presentan asimetrías. Los datos se corresponden a los rendimientos de los títulos referidos durante los meses de julio de 1995 a diciembre del 98, medido como: (cierre-apertura)/apertura. Las medidas descriptivas se muestran a continuación:

CUADRO Nº 1
MEDIDAS DESCRIPTIVAS DE RENDIMIENTOS
BOLSA ELECTRÓNICA DE CARACAS
JULIO 1995 – DICIEMBRE 1998

Empresa	Media	Desv.típica	C.V	Asimetría	curtosis
Ele-Caracas	0,03644	0,1767	484,906696	-0,175	1,424
Manpa	-0,016332	0,1558	-953,955425	-1,5	6,586
Mantex	-0,024108	0,1386	-574,912892	0,068	2,587
Mavesa	0,04429	0,1687	380,898623	0,492	0,735
B.Provincial	0,02613	0,1411	539,992346	0,337	1,154
Sivensa-a	0,01884	0,1912	1014,862	0,502	3,178
Sudantex-b	-0,0088448	0,1376	-1555,71635	-0,624	1.713
Vencemos 2	0,04782	0,1972	412,379757	1,714	5,379
Vencemos 1	0,05706	0,1865	326,848931	1,306	4,956
Venepal	0,02351	0,1019	433,432582	0,736	0,881

Fuente: cálculos propios

CUADRO Nº 2
MEDIDAS DESCRIPTIVAS DE RENDIMIENTOS
BOLSA ELECTRÓNICA DE CARACAS
JULIO 1995 – DICIEMBRE 1998

Empresa	Mediana	Moda
Ele-Caracas	0,04486	-0,45
Manpa	-0,0084746	-0,08
Mantex	-0,15152	0,00
Mavesa	0,01596	0,00
B.Provincial	0,000	0,00
Sivensa-a	0,004823	0,00
Sudantex-b	0,000	0,00
Vencemos 2	0,000	0,00
Vencemos 1	0,0177	0,00

Venepal	**0,003086**	**0,00**

Al considerar como rendimiento la diferencia entre el cierre y la apertura obtenemos resultados con características similares que en el caso anterior en donde se hace evidente la asimetría y alta curtosis pero por razones de espacio no podemos mostrar en este artículo. En todo caso, esto nos lleva a pensar que puede ser no conveniente el uso de la varianza en el caso del mercado de capitales tal cual como está formulado y lo mismo ocurre con los métodos de estimación empleado en el modelo multifactorial. Con el fin de ilustrar lo que estamos indicando correremos estos modelos propuestos dentro del ámbito clásico de métodos de selección de cartera óptimas, para proponer posterormente algunos cambios sin que con ello se quiera decir que es la solución definitiva..

III.-EL PROBLEMA DEL PRONÓSTICO.

Un elemento fundamental en la toma de decisiones en el mercado de valores es poder hacer un buen pronostico del comportamiento de los títulos y de los diferentes índices que pueden relacionarse con los precios o rendimientos de los mismos. Para ello se ha propuesto el uso de modelos ARIMA o los del tipo llamado de suavizado exponencial o cualquier otro método clásico desarrollado en el dominio del tiempo asumiendo un mercado eficiente. En la obra John E. Hanke (1995) puede consultarse los diferentes métodos de pronósticos mas populares aplicados a los negocios, igualmente se ha propuesto el uso de modelos no lineales como el ARCH desarrollados por R.F.Engle(1982), Kuldeep Kumar(1986), L.D. Byers (1995) y otras contribuciones y el modelo GARCH que es una generalización del anterior.

Un modelo del tipo ARIMA(p,d,q) tiene la forma general:

$$\mathbf{z_t=\mu+\Sigma_k\varphi_k z_{t-k} - \Sigma_j\phi_{t-j}\, a_{t-j} + a_t} \quad k=1,2,3\ ...p,\ \ j=1,2,3...q$$

Donde $\mathbf{x_t}$ es la variable aleatoria original asociada a la serie de tiempo dada,$\mathbf{z_t=(1-B)^d x_t}$ y $\mathbf{B\ x_t= x_{t-1}}$, φ_k y ϕ_{t-j} son los parámetros autoregresivos y de promedios moviles respectivamente que deben ser estimados atravez de las observación de $\mathbf{x_t}$ una vez diferenciado d veces, $\mathbf{a_t}$ es una variable aleatoria con ley N(0,1). Sobre estos parámetros, una vez estimados, se construyen contrastes de hipótesis de forma similar al modelo de regresión múltiple. Uno de los procedimientos para ver si el modelo seleccionado es un buen ajuste es estudiando los residuos mediante las funciones de autocorrelaciónes, se espera bajo la hipótesis de una buena adecuación del modelo que las correlaciones deben ser pequeñas. Para realizar el contraste de que las correlaciones son nulas, se emplea un test cuya distribución sigue la ley chi cuadrado. Este punto que acabamos de señalar, se encuentra ampliamente desarrollado en la bibliografía especializada

En el caso venezolano, F. Martínez (1988) propone el uso de un Arima(6,2,0) para hacer pronósticos del valor del precio promedio diario del año 1986 del título del Banco Provincial, demostrando que en aquel entonces el uso de este modelo daba buenos resultados para pronosticar este precio al menos para los siguientes cinco dias. De hecho, es posible encontrar información diaria de los precios promedios de los títulos obteniendo series temporales lo suficientemente grandes como para usar este tipo de modelo.
Como la realidad económica de Venezuela ha cambiado notablemente desde el año 86 hasta el presente, cabe preguntarse si los modelos ARIMA pueden seguir siendo útiles para realizar pronósticos diarios. Podrá pensarse que el comportamiento de los precios o de los rendimientos de los títulos siguen mas bien un comportamiento dinámico caótico, hipótesis que habrá que demostrarse empleando los indicadores adecuados para sustentar esta hipótesis. Otro caso del uso de estos modelos ARIMA para realizar pronósticos, es el caso de la tasa de cambio bolívar versus dólar, tasa de interés activas y pasivas e inflación que fué realizado por el autor de este artículo que lo requirió como insumo para un trabajo realizado para la antigua Corpoven empresa filial de PDVSA, la serie temporal empleada fue el promedio mensual de la cotización del dólar en bolívares, tasas de interés e inflación para el periodo de Enero del 87 a Julio del 91. Los mejores resultados se obtuvieron con la paridad cambiaria. Sin embargo aquí cabe formularse la misma pregunta de la validez actual de estos modelos y la hipótesis formulada de la existencia de patrón. En la actualidad se han presentado un gran número de artículos y libros en donde se postula que el mercado de valores, las tasas de cambio y otros indicadores económicos tasas de interes e inflación a la larga tendrían un comportamiento dinámico caóticos Entre los autores que tocan este tema están A. Fernandez (1994) y T. Vagas (1994) De ser plausible la hipótesis de que estas variables tienen un comportamiento caótico, el problema de pronóstico de la valorización de los títulos tanto en precio promedio diario o rendimiento diario no introduciría gran incertidumbre puesto que, como indica M.S.Bartlet (1990) los sistemas dinámicos caóticos emulan procesos estocásticos del tipo ARIMA al menos en el corto plazo (diario), esto permite entonces realizar los pronósticos de aquellas variables asociadas al mercado de valores mediante la aplicación de modelos ARIMA.

Sin embargo hay de los modelos nombrados al inicio de este artículo que en el caso del mercado venezolano se cuestionan por su poco valor práctico, por las razones que ahora expondremos.

El modelo de Markowitz está construido sobre datos históricos de los precios o de los rendimientos de los títulos y con estos datos debe realizarse el pronóstico de los valores futuros asumiendo que siguen un determinado patrón. Si se tienen suficientes datos, digamos mas de sesenta es posible que los datos correspondan a un proceso ARIMA y se podrá determinar mediante la muy conocida metodología desarrollada por los autores Box y Jenkins (1976) cual ARIMA se ajusta mejor a los datos. Estos datos históricos deben corresponder por lo menos a los valores de apertura y cierre diarios respectivamente para que tenga sentido su aplicación práctica. Hoy con el desarrollo de software no implica ningún problema puesto que cada vez podrá actualizarse los datos, realizando pronostico para los próximos primeros días y, resolver para cada uno de esos días el modelo propuesto por Markowitz para la obtención de la cartera óptima. Este pronóstico realizado a muy corto plazo no debe azorarnos si sospechamos que el comportamiento responde a una situación altamente compleja que llevado por la bibliografía actual, nos haga pensar que lejos de ser un proceso estocástico tipo ARIMA es mas bien un caso de caos determinista puesto que como ya indicamos en los párrafos anteriores el primero emula al segundo y el pronóstico funciona. La situación puede complicarse cuando nos referimos a meses, cosa que al final no tendría tampoco mucho sentido práctico en cuanto a la toma de decisiones. El problema de aplicación del modelo de Markowitz en el caso venezolano puede estar en otro aspecto mas que la necesidad de pronosticar como es el uso de la varianza, utilizada para medir el riesgo. Mas adelante haremos otra propuesta que si bien no cambia la esencia del modelo se presenta como una alternativa que nos permitimos sugerir (no sabemos al escribir este artículo si a otros se le ha ocurrido una idea similar)

Pero que ocurre en el caso venezolano cuando se quiere emplear el modelo MIU o un modelo multifactorial. En el caso del MIU suele emplearse o el índice de la bolsa o el IPC Nosotros conocemos una experiencia en donde se desarrolla un modelo multifactorial que por las variables empleadas corresponde al modelo de arbitraje de Ross, este trabajo lo realizó L.B.Casanova(1999) y en el se demuestra, tomando en cuenta los resultados de las estimaciones de los parámetros y contraste de hipótesis de los mismos, y mas aun, la validación de los supuestos subyacentes de un modelo de regresión múltiple que los resultados son desalentadores.

Por razones de espacio, en este artículo solo estudiaremos el caso del IPC acumulado que se emplea como variable explicativa tanto en el modelo MIU como el multifactorial. La serie parte de Enero 1950 hasta Febrero de 1999. Este hecho nos obliga a emplear los rendimientos o precio de los títulos en periodos mensuales perdiéndose de esta manera información valiosa de la forma como evolucionas estos valores diariamente, por tanto la variación explicada por el modelo es pobre si consideramos tan solo la variación mensual de los rendimientos o precio de los títulos como variable dependiente dentro de un modelo de regresión lineal. Por otra parte los múltiples vaivenes de nuestra economía hace presumir que el IPC no tendrá un comportamiento estadísticamente estable.

La serie histórica del IPC la estudiamos para encontrar el mejor modelo ARIMA que se ajusta a la misma.. El primer ARIMA tentativo con la serie completa es ARIMA(2,3,2) que da mejores resultado que el ARIMA(2,3,1) y otros similares, se probó con los arimas del tipo ARIMA(p,d,q)(P,D,Q) tratando de encontrar patrones estacionales y con ello no se mejoraron los resultados, mas aún se plantearon transformaciones del tipo logarítmico y el modelo originalmente propuesto siguió siendo el mejor, sin embargo si bien las pruebas t para contrastar los parámetros dan resultados muy satisfactorios como muestran sus probabilidades asociadas (valor p), al analizar los residuos de este modelo encontramos que no se puede garantizar que los mismos respondan a un proceso de ruido blanco. Empleando el paquete SAS, el cual prueba automáticamente varios modelos de pronostico en el modulo de ASSIST, encontramos que el mejor ajuste es un doble exponencial (log damped trend esponencial) con un error de estimación mucho menor que los diferentes modelos ARIMA empleados. Esta serie de tiempo se dividió en dos series ambas lo suficientemente largas para verificar el supuesto de estacionaridad. Si la serie original es estacionaria, entonces, los modelos que mejor se ajustan a las dos series obtenidas por la división de la serie original deben ser los mismos para las tres series y solo se observarán cambios en los valores de los estimadores de los parámetros, es decir, las estimaciones las cuales no deben ser significativamente diferentes entre sí. Si los modelos no son similares podría pensarse que las serie original responde a un proceso caótico, que de acuerdo a la clasificación de Allan Wolf (citado por Tonis Vaga) puede ser un sistema caótico del tipo A que se caracterizan por tener un atractor de baja dimensión fractal y es potencialmente predecibles, en contra posición del tipo B cuyo atractor es de alta dimensión fractal e impredesible. Lo que realmente ocurrió cuando se trabajó con las tres series que las mismas renpondía a los mismos modelos ARIMA y al mismo comportamiento logaritmico suavizado exponencial e cuando se empleó el SAS salvo el valor de los estimadores. En una situación como esta hay que contrastar las correspondientes hipótesis de igualdad de parámetros por una parte y, por otra ver el comportamiento del pronóstico se mantiene.

Los resultados preliminares obtenidos hasta el momento de escribir este artículo son:
a.-Al hacer los histogramas de las tres series mediante el SAS se encontró que los mismos se ajustan mejor a distribuciones asimétrica y con colas gruesa (alta asimetría y kurtosis).
b.- Hay cambio altamente significativos en las estimaciones de varios de los parámetros, además los intervalos de confianza de las predicciones son de longitud muy amplia en todos los casos (tres series).
c.- Esto nos hizo sospechar que el IPC responde a un sistema dinámico caótico del tipo A, el cuál es potencialmente predecible en el muy corto plazo. Para sustentar lo anterior nos abocamos a construir algunos índices.
Generalmente, para determinar si un sistema es caótico se emplean diferentes métodos: el basado en el mayor exponenete de Lyapunov que da una condición necesaria para detectar si un sistema dinámico es caótico o aleatorio, las dimensiones de Lyapunov, la de correlación, la de información, Kaplan-York y finalmente el método basado en análisis de Casdagli (1989).

Estos métodos y otros no mencionados acá han sido aplicados ampliamente en fenómenos físicos. La aplicación a otras ciencias es más limitado por el gran número de observaciones que requiere este tipo de análisis. Por ejemplo, el algoritmo desarrollado por Wolf para obtener el exponente o número de Liapunov requiere al menos 30^d observaciones para un sistema de dimensión **d.** Si el sistema es de baja dimensión, es decir de tipo A, como en el caso de un atractor de Lorenz entonces se requiere por lo menos 900 observaciones. En nuestro caso no tenemos este número de observaciones y si la dimensión es mayor, por ejemplo cercana a 4 se hace casi imposible aplicar este algoritmo a cualquier observación dentro del campo de la economía.
Debemos recordar que la dimensión mide el nivel de complejidad de un sistema dinámico, a mayor dimensión, mayor complejidad. El otro método que tenemos a la mano es el de análisis de rango desarrollado por H.E.Hurst, el cuál es comentado por B.Mandelbrot (1997). Se parte de la relación entre el rango estandarizado $R(X_t)/S(X_t)$ y la raiz de la media cuadrática:

$$\mathbf{R(X_t)/S(X_t) \cong \sqrt{(\sqrt{(\Sigma_i x_i^2 pi)})} = \sqrt{d}. \qquad pi=>0 \quad \Sigma_i pi=1}$$

H.E.Hurst encontró a partir de datos empíricos relacionados con el comportamiento del Nilo que $\mathbf{R(X_t)/S(X_t) \cong d^H}$ en donde H toma valores entre 0,9. B.Mandelbrot demuestra que H toma valores en el intervalo cerrado 0,1. Es importante destacar que el $\mathbf{R(X_t)}$ es

autoajustado es decir, se obtiene no de los datos originales sino a partir de los desvíos con respecto a la media de los mismos. En nuestro caso el valor de H es aproximadamente:0,442610.
Para obtener el valor de H se a propuesto una expresión equivalente dada por:

$$\mathbf{R(X_t)/S(X_t) \cong aN^H.}$$

En donde a es una constante, N es el número de observaciones, H es ekl exponente de Hurst.
La interpretación de H dada por Mandelbrot (1997) es como sigue:
1.-Si **0<=H<=0.5**, entonces el proceso es una serie antipersistente o ergódica, mientras mas cercano esté de cero mayor es la antipercitencia, en contrario, mientras mas cerca este H de 0,5 la antipersistencia es menor. Para Mandelbrot(1997), un valor de H comprendido en este intervalo lo asocia a series con movimientos brownianos fraccionarios antipersistentes." La antipersistencia consiste en la tendencia a regresar continuamente al lugar de procedencia, y por tanto a difundirse mas lentamente que sus homólogos brownianos".
2.-Si **H>0,5** entonces en la serie hay persistencia y ciclos no periódicos. Esta persistencia aumenta en la medida que H se acerque a 1 y consiste en repetir los valores altos o bajos de la serie de los periodos anteriores, es decir, por debajo o por encima de la media En este caso Mandelbrot(1997) señala que los ciclos que pueden detectarse en la muestra no se pueden extrapolar a muestras de mayor duración. Las tendencias y ciclos que muestran este tipo de serie temporales pueden deberse mas a ruidos.

A.Fernandez D (1994) observa que las series persistentes siguen un camino aleatorio sesgado.
Este exponente está relacionado con dos conceptos importantes como los son: la correlación y la dimensión fractal. La correlación en este caso se define como sigue.

$$\mathbf{C(r)= \lim_{n\to\infty}(1/n^2)\{\text{números de pares (i,j) tales que la distancia } |x_i\text{-}x_j| \text{ es menor que r}\}}$$

$$\mathbf{C(r)= \lim_{n\to\infty}(1/n^2)\, \Sigma_i \Sigma_j\, \mathit{f}(r\text{-}|x_i\text{-}x_j|)}$$

.Donde *f(.)* es la función de Heavyside la cual se define como: $\mathbf{(r-|x_i-x_j|)>0}$ entoces $\boldsymbol{f}\mathbf{=1}$; $\mathbf{(r-|x_i-x_j|)<0}$ entonces $\boldsymbol{f}\mathbf{=0}$. H se relaciona con C(r) como:

$$\mathbf{C(\ r) =2^{(2H-1)}-1}$$

Valores de H inferiores a 0,5 se obtiene una correlación negativa, cuando es exactamente 0,5 la correlación es nula y para valore mayores a 0,5 la correlación es positiva, siendo perfecta cuando H es igual a la unidad.
La relación entre H y la dimensión fractal está dada por:

Df=1/H

La relación es correcta si la dimensión D que contiene a la dimensión fractal es D>1/H.
En el caso que nos ocupa C(r)=-0,07647677 y Df=2,25932498: este resultado se interpreta de la siguiente forma: como el coeficiente de correlación es muy bajo y la dimensión fractal es baja también nos hace pensar que la serie puede provenir lo sumo de un sistema dinámico caótico del tipo A, los cuales son predecibles a muy corto plazo.
Quedaría por analizar el comportamiento del precio de los títulos en el mercado de capitales si los mismos responden a un proceso estocástico o a uno caótico del tipo A, en tal caso las predicciones se podrían realizar al menos durantes los próximos dos o tres días. Como señala Wolf no siempre es fácil decidir si el proceso es aleatorio o dinámico a pesar que se han propuestos varios métodos tales como la dimensión de correlación. Uno de los problemas fundamentales en economía es el número de observaciones que requieren las diferentes metodologías para diferenciar el tipo de sistema dinámico y este número es en la practica muy pequeño.

IV.-UNA PROPUESTA RELACIONADA CON EL MODELO DE MARKOWITZ .

En el siguiente punto proponemos un modelo inspirado en el original de Markowitz pero variando la función objetivo para hacerla robusta frente a valores que siguen una ley de probabilidad asimétrica y de colas gruesas como ocurre con los precios de los títulos en el mercado venezolano. La función es:

Min
$\mathbf{R_c=\Sigma x_i d_i+2\Sigma\Sigma x_i\, x_i\, d_{ij}}$
Sujeto a:

$\mathbf{\Sigma x_i E_i=E^*; \quad \Sigma x_i=1\ ;\ x_i =>0}$ **para todo i=1,2 ...n**

Donde $\mathbf{d_i=\Sigma\,|R_i\text{-}mediana_i|/n}$ es decir la suma de los valores absolutos de la desviaciones de los rendimientos del título i con respecto a la mediana dividido entre el número de observaciones y $\mathbf{d_{ij}=\Sigma\Sigma\,|(R_i\text{-}mediana_i)(R_j\text{-}mediana_j)|/n}$ el cual mide la covariación de los rendimientos de dos títulos diferentes con respecto a sus medianas. Puede proponerse otras medidas como la de Gini pero esto queda solo como una propuesta que habrá que analizar desde el punto de vista teórico y empírico.

En notación matricial el modelo puede expresarse como sigue:

Min
$\boldsymbol{R_c=d^T x+x^T Dx}$
Dujeto a:
Ax=b
Donde $\mathbf{d^T}$ = ($d_{1.}$ $d_{2\ ,}$ $d_{3\ ,}$ --- d_n), $\mathbf{x^T}$ =($x_{1.}$ $x_{2\ ,}$ $x_{3\ ,}$ --- x_n) y D es la matriz de los desvíos con respecto a la mediana del rendimiento de cada título:

$$\begin{vmatrix} d_{11} & d_{12} & d_{13} & & & d_{1n} \\ d_{21} & d_{22} & d_{23} & & & \\ d_{31} & d_{32} & d_{33} & & & \\ & & & & & \\ & & & & & \\ & & & & d_{n-1.n-1} & \\ d_{n1} & d_{n2} & d_{n3} & & d_{n.n-1} & d_{nn} \end{vmatrix}$$

Esta matriz es también simétrica como la matriz de varianza covarianza.
A es la matriz cuyo primer vector fila es: $(Md(R_1), Md(R_2), Md(R_3),\ldots Md(R_n))$ y el segundo vector (1,1,1,…1); b es un vector columna de dos componentes $(Md^*.1)^T$
Este modelo se puede resolver empleando la herramienta Solver de Excel que tiene la.

V.-CONCLUSIÓN.

1.-Los datos disponibles en el mercado de valores en Venezuela nos lleva a concluir, que dada la asimetría y sesgo de los mismos, sería interesante probar el modelo de Markowitz tomando en consideración la modificación propuesta.

2.-Los pronósticos necesarios para aplicar el modelo de Markowitz debe hacerse tomando en cuenta un horizonte de muy pocos días, para que el efecto de ¨caotizidad¨, si existe, no afecte el mismo y pueda aplicarse la herramienta propuesta por Markowitz o su variante dada en este artículo.

3.-El uso de modelos unifactoriales (MIU) o multifactoriales deben replantearse o nuevamente especificarse en el caso venezolano. Este replanteo puede incluir otros métodos de estimación de los parámetros como la norma L1 o, establecer restricciones como la ortogonalidad de los parámetros. La especificación estaría incluir restricciones de tipo económico o buscar variables explicativas no propuestas anteriormente haciendo mas realista el uso de la regresión simple o múltiple para definir las carteras eficientes.

En todo caso, queda abierta la posibilidad de indagar mas en estos modelos que permitan adecuarlos mas a nuestra realidad. O mejor aún, usar la imaginación para proponer otros modelos que no necesariamente sigan el esquema de cada uno de los vistos en este artículo.

BIBLIOGRAFÍA CONSULTADA.

Bartlet, M.S(1990).
Chance or Chaos
Journal Royal Statistica Soc. 153 part3-321-342.

Byers J.D; Peal D.A(1995)
Forescasting Industrial Production using non linear methods.
Jounal of Forescasting. Vol 14. Pag. 325-336

Casanova,L.B (1999).
Análisis de los Factores que Influyen en la Toma de Decisiones de los Inversionista dentro del Mercado de Valores en Venezuela.-Tesis para optar al título de Especialista en Finanzas. U.C.V-FACES

Engle, R.F(1982)
Autoregressive Condicions Heteroscedasticity whit Estimates of the of Variance of United Kindong Inflation,
Econometrica. Vol 50 Nº4 Pag. 987-1007.

Fernandez Diaz A.(1994)
La Economía de la Complejidad-McGraw-Hill.

Mandelbrot B(1997)
Geometría Fractal de la Naturaleza-Tusquets Edicciones-Metatema 49.

Martínez, F.(1988)
Un modelo de Apoyo a la Toma de Decisiones en la Negociación de Acciones. Tesis para optar al título de Magíster en Ciencias Administrativas. U.C.V-FACES

Messuti D.J; Alvarez V.A; Graffi H.R (1994).
Selección de Inversiones-Introducción a la Teoría de la Cartera.-Ediciones Macchi.

Novales. A (1993)
Econometría (Segunda Edición)- McGraw-Hill.

Reyes P, A.E (1985)
Aplicación del Criterio Norma L1 al Modelo Lineal. Tesis para optar al título de Magíster Scienciarum en Estadística. U.C.V-FACES

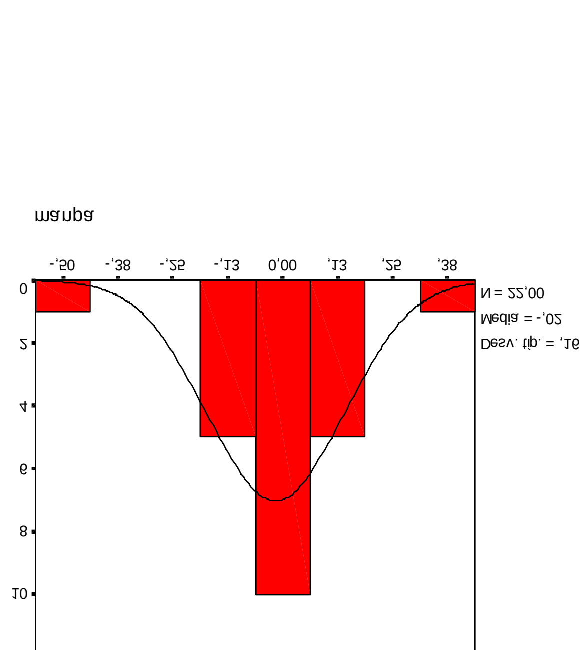
Reyes P, A.E (1996)
Predicción y Caos.
Temas de Fronteras en el Campo de la Gerencia. Cuadernos de Postgrado Nº11Pag. 143-155

Rochen J.(1992)
ARMA covariance structure whit heteroscedasticity for reapet measure experiments. JASA vol 87 Nº 419 pag 777-784.

Sharpe W.F(1974)
Teoría de Cartera y del Mercado de Capitales-Ediciones Deusto.

Suarez.S; A.S (1997)
Decisiones Optimas de Inversión y Financiamiento en la Empresa.-Ediciones Pirámides.

Tonis Vaga(1994)
Profiting from Chaos- McGraw-Hill



manpa
-,50
-,38
-,25
-,13
0,00
,13
,25
,38
0
2
4
6
8
10
N = 22,00
Media = -,02
Desv. típ. = ,16