

**VIII.**

Universidad de Chile  
**Facultad de Economía y Negocios**  
Escuela de Economía y Administración

“Proyección del Signo de las Variaciones del Precio de las Acciones: Modelos  
Multivariados Dinámicos”

Seminario de Título Ingeniero Comercial  
Mención Administración

Alumno: María Fernanda Norambuena  
Pape.

Profesor Guía: Franco Parisi Fernández,  
Ph.D.

Profesor Instructor: David Díaz Solís,  
M.F.

Santiago, Chile 2007

## **Tabla de Contenidos**

I. Introducción .....	1
II. Revisión Bibliográfica .....	5
III. Dow Jones Industrial y ALCOA: Algunas Características .....	11
Dow Jones Industrial .....	11
Alcoa Inc.....	12
IV. Metodología y Datos .....	17
Modelos ARIMAX .....	18
a) Modelos de Medias Móviles (MA).....	19
b) Modelos Autorregresivos (AR).....	21
c) Modelo Autorregresivo de Medias Móviles (ARMA) .....	22
Algoritmos Genéticos (AG) .....	24
Rolling Regression .....	26
Evaluación de la Capacidad Predictiva y Significancia Estadística .....	27
V. Análisis de los Resultados.....	29
VI. Conclusiones .....	34
VII. Bibliografía y Referencias .....	36
VIII. Anexos .....	39

## I. Introducción

En la actualidad existen diversos modelos que se utilizan para predecir los precios accionarios. Lo anterior constituye un tema de gran relevancia, ya que los *traders* y analistas financieros desean obtener el máximo de rentabilidad en la conformación de su portafolio. Es así como, en la búsqueda de este objetivo, algunos modelan el comportamiento de las acciones del mercado según el precio exacto de éstas y otros según los signos de las variaciones de dichos precios.

Leung, Daouk y Chen (2000) muestran que la proyección de los rendimientos de las acciones, dado el conjunto de variables de información para realizar las estimaciones, es una realidad aceptada por el mundo financiero. De esta manera, el trabajo de los *traders* y analistas financieros termina centrándose en cómo usar de manera óptima la información para predecir y transar en los mercados. Cabe destacar que estas prácticas, en su mayoría, se basan en la predicción precisa de los precios de los instrumentos financieros. Sin embargo, estudios recientes han dejado en evidencia que aquellas transacciones basadas en la proyección de la dirección de los movimientos de los precios de las acciones son más eficientes y efectivas y, más aún, pueden generar mayores ganancias que aquéllas basadas en la predicción específica del precio de la acción.

Siguiendo en esta misma línea, Hodgson y Nicholls (1991) muestran en su estudio que es de suma importancia considerar la significancia de la proyección dentro de los modelos de predicción de los movimientos de los precios.

El presente seminario tiene como objetivo principal cuestionar si predecir los precios accionarios en base al precio del día anterior, es la manera más eficiente y efectiva para obtener un mayor beneficio. Cabe señalar que este estudio no analiza la significancia económica de la rentabilidad de las estrategias de compra y venta surgidas por las recomendaciones de los modelos predictivos, cuestión que se plantea como un tema para una futura investigación. Se plantean distintos supuestos y modelos econométricos de manera de estudiar la significancia estadística de predecir los precios con diferentes sets de variables explicativas, estando estos últimos compuestos por precios rezagados en términos diarios y/o semanales.

La investigación a efectuar se centra en el empleo de modelos multivariados ARIMAX dinámicos, contruidos a través de Algoritmos Genéticos. Para realizar este estudio se consideraron variables de series de tiempo con hasta un máximo de 4 rezagos para la rentabilidad de la acción en análisis, compuestos en forma semanal; 4 posibles rezagos para la rentabilidad del índice Dow Jones Industrial considerados en forma semanal; 4 rezagos para la rentabilidades de los precios de las acciones de los días anteriores; y lo mismo para el índice Dow Jones Industrial; y, por último, 4 rezagos de error del modelo aplicado.

El estudio presenta evidencia y valida el hecho de que los Algoritmos Genéticos en los modelos ARIMA pueden ser usados como una alternativa estratégica para el mejoramiento en la construcción de los modelos de proyección de series de tiempo, en función de la proyección del signo de las variaciones de los precios de las acciones.

En lo que se refiere al marco teórico, éste centra su discusión en torno a la Hipótesis de Mercados Eficientes, planteada por Fama (1970). En ella se afirma que los precios reflejarán perfectamente la información presente en el mercado o, por lo menos, se espera de que así suceda. Se entiende por eficiencia de mercado que, los precios reflejan completamente la información disponible en el mercado.

La acción que se utilizará para la predicción del signo de la variación del precio será la de Alcoa Inc. (AA), debido a la importancia del aluminio como materia prima a nivel mundial y las futuras proyecciones que se tienen sobre ésta y otras características que se discutirán posteriormente. Los datos utilizados provienen de la NYSE, entre el período del 22 de febrero de 2002 hasta el 28 de agosto de 2007. Los precios utilizados se extrajeron de la página [www.finance.yahoo.com](http://www.finance.yahoo.com).

La presente investigación se divide en 6 secciones: en la sección dos se presentará una breve revisión bibliográfica acerca de la predicción de modelos bursátiles; en la sección tres se presentarán la historia y las características sobre Alcoa Inc. y el índice Dow Jones Industrial; en la sección cuatro se presentará la metodología y los datos del sistema de predicción planteado; en la sección cinco se revisarán los resultados obtenidos; y, por último, en la sección seis se presentarán las conclusiones finales del estudio.

## II. Revisión Bibliográfica

En 1970 Fama plantea la hipótesis de eficiencia de mercado, en el sentido de que los precios de las acciones reflejaban completamente la información disponible en el mercado. De esta manera, la nueva información de mercado, cuyo carácter es aleatorio, haría que el precio de las acciones también lo fuesen; por lo que muchos estudiosos han llegado a creer y defender la posición de que los precios accionarios siguen una caminata aleatoria, mejor conocida como “*random walk*”. Sin embargo, muchos académicos y analistas financieros han comprobado que el precio no sigue un proceso estocástico de estas características, existiendo un margen para la proyección de los retornos, siendo éstos efectivamente previsible en algún grado. Un ejemplo de esto es el trabajo realizado por Parisi, Parisi y Guerrero (2003), en el cual estudian la capacidad de los modelos de redes neuronales para predecir el signo de las variaciones semanales de los índices bursátiles. De acuerdo a Hellström y Hollström (1998), en el contexto de la hipótesis de mercado eficiente no se predicen los cambios futuros que pueden tener los precios de las acciones; esto quiere decir, por ejemplo, que la entrada de una nueva información al sistema es inmediatamente descubierta y eliminada, produciéndose un cambio en los precios de mercado.

Por otra parte, según Lo y Mackinlay (1988), se puede entender que existe una cierta regularidad o una cierta información dentro de los precios históricos que permiten predecir en un cierto futuro cercano estos mismos. En dicho estudio se registró una correlación serial positiva entre los retornos semanales en una serie de índices y portafolios de tamaño medio. Cabe mencionar que, para sus conclusiones, se utilizaron datos de EE.UU., Europa Occidental y Japón, para el período comprendido entre 1962 y 1985. Además, para el mismo período señalado anteriormente, Conrad y Kaul (1988) comprobaron la existencia de predictibilidad para los retornos en el corto plazo usando datos semanales; la metodología utilizada fue un proceso autorregresivo de primer orden, que permitió obtener las conclusiones ya mencionadas.

De Bondt y Thaler (1985), Fama y French (1988), Poterba y Summers (1988) y Chopra et al. (1992), evidenciaron una correlación serial negativa en los retornos de los activos individuales y en varios portafolios sobre intervalos de tres a diez años. Por otro lado, Jegadeesh (1990), ocupando datos de EE.UU., investigó la predictibilidad de los retornos mensuales sobre activos individuales para los años 1934-1987, y comprobó la existencia de una correlación serial negativa de primer orden altamente significativa para rezagos de dos meses, y una correlación serial positiva para rezagos mayores. Blume, Easley y O'Hara (1994) sugieren la existencia de un modelo teórico que permite una relación significativa entre los rezagos del volumen transado y los rendimientos actuales de los activos individuales.

Por su parte, Ferson y Harvey (1991) estudiaron que la predicción en los retornos de las acciones no se debe solamente a ineficiencias del mercado o al actuar irracional de los inversionistas, sino más bien a algunas variables informativas que están intrínsecamente incluidas dentro del modelo que predice la rentabilidad del precio de las acciones. En esta misma línea, Ferson y Schadt (1996) muestran que la no inclusión de rezagos de los retornos accionarios y de las tasas de interés dentro del modelo, podrían generar resultados ineficientes a la hora de predecir.

No obstante, Hodgson y Nicholls (1991) han señalado que es de suma importancia considerar la significancia económica de predecir los movimientos de los precios de las acciones y, por ende, de la rentabilidad, ya que el grado de predictabilidad del precio no se suele considerar como significativo y relevante. En adición, Leung, Daouk y Chen (2000) compararon la capacidad predictiva de los modelos que predecían el signo o la dirección del retorno de índices bursátiles junto con los de estimación de nivel de precios, con lo que pudieron concluir que estos últimos se desempeñan de manera menos eficiente que los primeros en términos de su tasa de acierto y, por lo tanto, son capaces de generar mayores beneficios.

Posteriormente, los investigadores Wu y Zhang (1997) determinaron que las estrategias de transacción basadas en la estimación del cambio en la dirección de los niveles de precios son más efectivas y eficientes, y permiten obtener

beneficios mayores que aquellas estrategias basadas en la predicción específica de los niveles de precios de los instrumentos financieros. Así mismo, O'Connor (1997) apoya la idea de proyectar la dirección del cambio en el nivel de precios más que el nivel de precios en sí.

Parisi et al. (2004) plantean el conteo de aciertos de la predicción como la forma de evaluar la capacidad predictiva. Para ello utilizaron Algoritmos Genéticos y redes neuronales, siendo el aporte principal la utilización de la técnica de *bootstrapping* para la construcción de intervalos de confianza de las proyecciones de signo y para analizar la robustez de los modelos analizados. Parisi, Parisi y Díaz (2006) analizan la capacidad de los modelos construidos a partir de Algoritmos Genéticos y redes neuronales, para predecir el signo de las variaciones semanales de los índices bursátiles. En su estudio utilizaron modelos multivariados obtenidos a través de Algoritmos Genéticos, los cuales obtuvieron mejores desempeños en términos de rentabilidad corregida por riesgo (medida por los índices de Sharpe y Treynor) que la red neuronal ward, la cual obtuvo una mejor capacidad predictiva, pero no obtuvo una mayor rentabilidad corregida por riesgo. De acuerdo a lo anterior, se apoyan las conclusiones del estudio de Leung, Daouk y Chen (2000), según el cual la predicción de la dirección del movimiento puede arrojar mayores ganancias de capital que la proyección del valor de cierre.

Parisi y Gutierrez (2004) comprueban que usar la metodología *rolling* para la proyección de las redes neuronales artificiales constituye una mejor predicción del retorno de las acciones que el promedio histórico, ya que se puede lograr un mejor desempeño y, por ende, una mayor rentabilidad de la conformación de carteras. Cabe destacar dos investigaciones en que se ha realizado la proyección del cambio de signo del precio de las acciones, utilizando otras metodologías. En una de ellas, Márquez (2006) utilizó la técnica de redes neuronales con rasgos de conciencia para predecir lo anterior; en dicho estudio se obtuvo para la acción en que se centra esta investigación (Alcoa Inc.) un PPS de 58,78%. Por otra parte, en Cortés y Figueroa (2006) se utilizó la metodología de Modelos ARIMAX y óptimos de *rolling* para la predicción de la variación del precio de las acciones, utilizando índices de mercado y sectoriales (S&P 500 y sectoriales, respectivamente). En el estudio se obtuvo para la acción de este seminario un PPS de 55,65% usando el S&P 500, y el mismo resultado utilizando el S&P sectorial.

Luego de analizar los distintos estudios y las diferentes metodologías que existen para predecir los precios de las acciones, se observa que en éstas se consideran como variables a explicar sólo a aquéllas que están en términos de rezagos semanales. Debido a lo anterior, en este seminario se utilizarán variables semanales y diarias para estimar la proyección del signo de las variaciones del precio de las acciones, con el objetivo de mejorar la forma y

procedimientos que hasta hoy se utilizan en la predicción, capturando aquellas tendencias de corto plazo que ocurren en los precios.

### III. Dow Jones Industrial y ALCOA: Algunas Características

#### Dow Jones Industrial

El Dow Jones, desde sus inicios hasta la actualidad, se ha utilizado como un indicador de bienestar de los diferentes mercados mundiales. Dicho índice es seguido y considerado por diferentes analistas financieros, estudiosos y *traders*, como un instrumento de gran importancia para sus investigaciones de predicción, cobertura y arbitraje en los mercados bursátiles. Es más, sus movimientos y variaciones son mencionadas por los especialistas siempre que ocurre cualquier evento de importancia mundial, ya que la nueva información inyectada en el mercado afecta a las industrias que listan en este índice, las cuales poseen gran importancia e influencia a nivel mundial. Es por ello que los analistas utilizan este índice como parámetro para visualizar y predecir el impacto generado por esta nueva perturbación.

El Dow Jones se compone de tres índices promedio, los cuales se dividen según la medición que realizan en los sectores empresariales donde se desempeñan las diferentes compañías que lo componen: sector industrial, sector transporte y sector *utilities*.

En el presente trabajo se utilizan los rezagos del índice promedio Dow Jones Industrial Average (DJIA) como variable explicativa dentro del modelo de

predicción de la variación del precio de las acciones. El Dow Jones Industrial Average no sólo mide títulos que pertenezcan al área industrial, sino que también aquéllos que están relacionados con las áreas de servicios, como el área de bienes de consumo, financiero, comercio minorista y sector tecnológico.

El DJIA es el índice promedio más antiguo de Estados Unidos, y abarca dentro de él a las 30 empresas más importantes de las industrias que mueven las bolsas de valores estadounidenses. Las empresas que lo componen constituyen entre el 25% y 30% del valor total de las acciones de aquel país.

#### Alcoa Inc.

El aluminio es un *commodity* que últimamente está siendo fuertemente demandado por países del Asia Oriental, especialmente por China, el cual es el principal consumidor a escala mundial de este insumo.

Hoy en día, China es el líder en el desarrollo tecnológico y en investigación a nivel global, lo que lo ha llevado a ser el determinante de los precios de las materias primas en la economía mundial. En esta misma línea, dicho país consume un cuarto del aluminio producido en el mundo, por lo que el precio de este *commodity* últimamente se ha elevado significativamente, y las proyecciones financieras dicen que esto se mantendrá por mucho tiempo. Es

así como investigadores del Prudential Equit Group han mencionado que la demanda global de aluminio se incrementará en un 6,3% para el presente año, impulsada principalmente por el país mencionado anteriormente.

El precio de las acciones de aquellas empresas cuyo producto principal y mayor fuente de ingresos es el aluminio, entre ellas Alcoa Inc., es de carácter cíclico, ya que el precio de esta materia prima está sujeto a los diferentes cambios que afectan a la economía mundial. Lo anterior se traduce en que, en épocas de expansión de la economía, estas compañías tendrían menor tasa de costo de deuda, junto con una menor tasa de costo de capital, ya que se encontrarían con una fuerte demanda del producto y, por ende, altas utilidades; por el contrario, si nos encontramos en épocas en que la actividad económica es baja o de crisis, el precio de la misma bajaría considerablemente, ya que los demandantes dejarían de utilizar este insumo y buscarían aquél que pueda reemplazarlo a menor costo.

Hoy en día, Alcoa Inc. es el líder mundial de producción de alúmina y el segundo productor de aluminio. Debido a lo anterior, se ha elegido en este seminario a dicha empresa como caso particular para realizar la predicción de la variación del signo del precio de las acciones utilizando modelos multivariados dinámicos, ya que se espera que para los próximos meses sea una de las que posea mejores rentabilidades en el índice Dow Jones Industrial Average. Alcoa Inc. es una empresa que se dedica a la producción y

comercialización de productos de aluminio, en forma de metales primarios y manufacturados, los cuales son vendidos a través de todo el mundo.

Actualmente, la empresa posee cuatro unidades estratégicas de negocios, las que se dividen en: productos realizados de bobina aplastada, como por ejemplo, hojas contenedoras rígidas, papel y planchas de aluminio y de acero; metales primarios; productos y servicios de ingeniería, que incluye extrusiones de aluminio forjado, piezas de fundición, alambres y barras de metal; y, por último, productos de aluminio y químicos, que fabrica bauxita, aluminio y productos químicos derivados del aluminio.

Alcoa Inc. suministra con sus productos a cuatro sectores de la economía. Entre ellos está el área de transporte (automotriz, aeroespacial, ferroviario y marítimo), el área de embalaje, el área de la construcción y, finalmente, el área industrial.

Además, dentro de su proceso productivo, la compañía realiza una serie de actividades relacionadas con el área tecnológica, la minería, la refinación, fundición, fabricación y reciclaje del aluminio.

Alcoa es considerada una compañía de gran influencia y tamaño, ya que posee una considerable participación de mercado, no sólo a nivel local, donde posee un 59% del mercado, sino que también a nivel mundial. En la actualidad, tiene

25 negocios independientes y 228 plantas operativas en 32 países. Un ejemplo de esto es que en el rubro minero ésta se encuentra presente en diversos países, tales como Australia, Brasil, Guinea y Jamaica, entre otros.

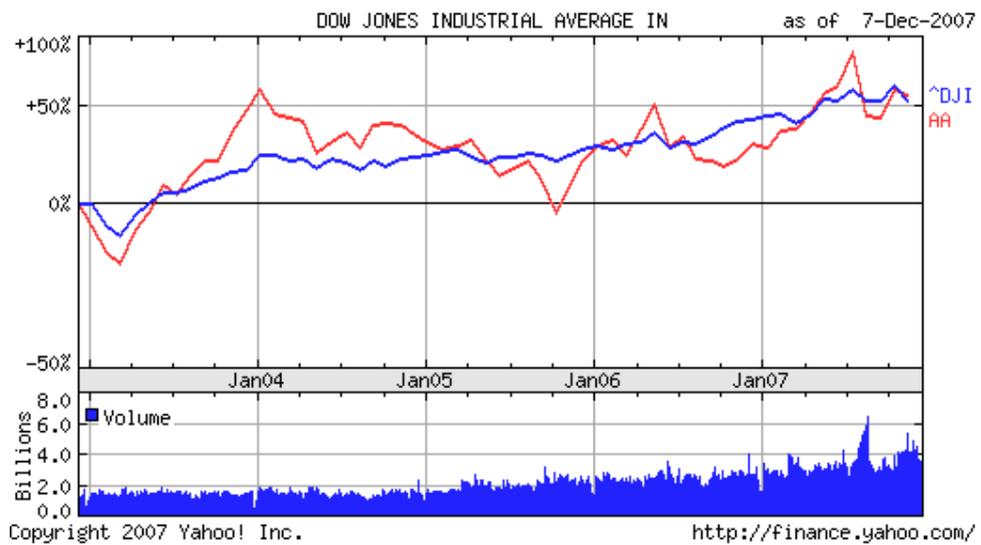
A continuación se observa la Figura 1, en la que se muestra la evolución del retorno accionario de Alcoa Inc. durante los últimos cinco años,<sup>1</sup> junto con el índice Dow Jones Industrial. Se observa tal como se mencionó anteriormente, que últimamente el aluminio y, por ende, el precio de la acción tiene buenas proyecciones, que el precio y la demanda han aumentado para esta materia prima. Además, se puede observar que Alcoa Inc. está directamente relacionada con la economía de EE.UU., por lo que, cuando el índice DJI va al alza, el precio de esta acción también lo hará, y en el caso contrario, cuando el índice va a la baja, el precio también.

---

<sup>1</sup> En el Anexo 1, se puede ver el Gráfico de Alcoa Inc. y el del índice Dow Jones Industrial por separado.

**Figura 1**

Evaluación de la rentabilidad de Alcoa Inc. y Dow Jones Industrial. Período Enero 2003-Diciembre 2007.



La Figura 1. muestra cómo se comporta la rentabilidad de la acción de Alcoa Inc. (AA), en relación al retorno del índice Dow Jones Industrial. Además, en el presente gráfico se puede verificar cómo el comportamiento de la acción AA es fuertemente afectado por factores de la economía, en este caso particular por el índice Dow Jones Industrial (^DJI). Esto ocurre principalmente porque el aluminio está sujeto a factores cíclicos de la economía mundial.

#### IV. Metodología y Datos

El objetivo de este seminario es encontrar un modelo multivariado dinámico que prediga mejor la dirección del movimiento del precio de la acción en particular, es decir, que tenga una tasa y un rendimiento económico significativo, a fin de proveer una herramienta útil para la conformación de sus portafolios.

Para el desarrollo de este seminario se utilizaron 261 rentabilidades semanales, las cuales fueron calculadas usando precios de cierre del índice Dow Jones Industrial y de la acción Alcoa Inc. (AA), correspondientes al período del 22 de febrero de 2002 hasta el 27 de agosto de 2007. Los datos fueron obtenidos desde [www.finance.yahoo.com](http://www.finance.yahoo.com).

La variable dependiente que se utilizó fue la rentabilidad de la acción en un cierto período, y las variables explicativas fueron los rezagos de las rentabilidades del índice Dow Jones Industrial, los rezagos de las rentabilidades del precio de la acción y los residuos rezagados estimados. El hecho de utilizar las variables en rezagos permite que las variables consideren el desfase lógico generado por la diferencia que existe entre el momento en que se realiza la predicción, el período actual o  $t-1$ , y el período futuro para el cual se realiza la proyección o momento  $t$ . Entonces, para realizar esta estimación,

se utilizó aquella información rezagada (desfasada) con respecto a la proyección, razón por la cual, se utilizaron modelos dinámicos con rezagos en las variables independientes.

De esta forma, se logró estimar los mejores modelos ARIMAX, dentro de un grupo de modelos multivariados dinámicos, determinado por las múltiples combinaciones que se obtuvieron gracias al uso de Algoritmos Genéticos. Además, se aplicó el test de Pesaran y Timmermann (1992), el cual permitió identificar cuál de los modelos estimados era el que reportaba una capacidad predictiva estadísticamente significativa. De acuerdo a éste, si los signos de lo observado y lo proyectado coincidían se anotaba un acierto del modelo; y si ocurría lo contrario, disminuía la capacidad predictiva del modelo. Cabe mencionar que, luego de encontrar aquellos modelos que reportaban el mayor porcentaje de predicción estadísticamente significativo, se aplicó la técnica de *Rolling Regression*, otorgando de esta manera mayor importancia a la información recientemente incluida en los modelos.

### Modelos ARIMAX

En el estudio se utilizó una amplia gama de Modelos ARIMAX, los cuales se pudieron estimar gracias a la utilización de Algoritmos Genéticos, para luego encontrar aquel modelo que fuese más eficiente en cuanto al número de aciertos en relación a la variable estimada, es decir, que tuviese mejor

Porcentaje de Predicción de Signo (PPS) y que además tenga mayor rentabilidad. Los modelos ARIMA fueron desarrollados por Box y Jenkins (1970), y tienen como objetivo proyectar una variable en función del comportamiento pasado de series de tiempo o, mejor dicho, variables temporales. Este modelo utiliza información más reciente como valor de partida y luego analiza los errores de proyección más reciente para así ajustarlo apropiadamente. Se examinan varios modelos con el fin de encontrar el modelo correcto.

La metodología de Box y Jenkins supone que no se conocen o que no están disponibles aquellas variables que determinan el valor proyectado, siendo en esta situación donde los Modelos ARIMA tienen una gran ventaja sobre los modelos de regresión típicos.

Existen tres metodologías básicas que permiten encontrar el modelo correcto:

- a) Modelo de Medias Móviles (MA)
- b) Modelo Autorregresivo (AR)
- c) Modelo Autorregresivo de Medias Móviles (ARMA)

- a) Modelos de Medias Móviles (MA)

Este modelo se caracteriza por explicar el valor de una determinada variable en un cierto período  $t$ , en función de una variable independiente y una serie de rezagos de los errores, ponderados según la temporalidad que posean.

El modelo de medias móviles para  $n$  períodos posee la siguiente forma:

$$Y_t = e_t + w_1 \cdot e_{t-1} + w_2 \cdot e_{t-2} + \dots + w_n \cdot e_{t-n} \quad (1)$$

donde  $e_t$  es ruido blanco<sup>2</sup>.

Ahora bien, los modelos de medias móviles se clasifican según dos características principales, las cuales determinarán si es una serie  $MA(1)$ ,  $MA(2)$  u otra. Estas características son: Autocorrelación y Autocorrelación Parcial.

La Autocorrelación se refiere al grado de asociación entre los valores de una variable en distintos estados de tiempo. Lo anterior está determinado según si una serie de tiempo posee autocorrelación, lo que llevará a una dependencia entre los valores de la serie en diferentes períodos. Estudiar la autocorrelación entre valores seguidos de una serie de tiempo permite saber qué modelo es apropiado para su representación.

---

<sup>2</sup> Ruido blanco es una sucesión de variables con esperanza cero, varianza constante e independiente para distintos  $t$ .

Por otro lado, la Autocorrelación Parcial se caracteriza por medir la relación que existe entre  $Y_t$  y  $Y_{t-k}$ , cuando los rezagos de los demás períodos de la variable dependiente permanecen constantes. La Autocorrelación Parcial se utiliza cuando no se tiene seguridad acerca del orden del proceso autorregresivo para corregir la serie de tiempo. Entonces, esta segunda característica permitiría identificar la relación entre los períodos actuales y pasados de una serie de tiempo.

De esta forma, si existe un coeficiente de correlación de carácter global o parcial, esas variables deberán ser eliminadas del modelo, identificando así a “ $p$ ” de un modelo  $MA(p)$ .

#### b) Modelos Autorregresivos (AR)

Un Modelo Autorregresivo es aquél donde la variable dependiente en un cierto período  $t$  es explicada por las observaciones de la misma variable, pero de períodos anteriores, incluyéndose además un término de error. Luego, un modelo Autorregresivo tiene la siguiente forma:

$$Y_t = A_0 + A_1 \cdot Y_{t-1} + A_2 \cdot Y_{t-2} + \dots + A_n \cdot Y_{t-n} + e_t \quad (2)$$

donde  $A_i$ , con  $i = 1, 2, \dots, n$  son los coeficientes que acompañan a las variables independientes, y  $e_t$  es ruido blanco.

c) Modelo Autorregresivo de Medias Móviles (ARMA)

El Modelo Autorregresivo de Medias Móviles es una combinación de un Modelo AR y un Modelo MA. Entonces, un modelo Modelo Autorregresivo de Medias Móviles de orden  $p$  y  $q$ ; posee la siguiente estructura:

$$Y_t = A_0 + A_1 \cdot Y_{t-1} + \dots + A_p \cdot Y_{t-p} + w_1 \cdot e_{t-1} + \dots + w_q \cdot e_{t-q} \quad (3)$$

donde  $w_j$ , con  $j = 1, 2, \dots, q$  son los coeficientes que acompañan a los rezagos de los errores.

Para poder explicar y conceptualizar qué son los modelos ARIMA, es necesario señalar dos tópicos que son de suma relevancia, es decir, proceso estocástico y estacionariedad.

El Proceso Estocástico es un proceso que sucede de variables aleatorias ordenadas, pudiendo  $Y_t$  tomar cualquier valor de  $t$  entre  $-\infty$  y  $\infty$ . Estas variables poseen diferentes cualidades, como por ejemplo, que tienen su propia distribución; que un par de variables poseen su propia distribución de probabilidad conjunta. Sin embargo, identificar estas funciones de distribuciones conjuntas resulta complejo, por lo que identificar un proceso estocástico, basta sólo con identificar la media y la varianza de cada variable y, por supuesto, la covarianza  $Y_t$  para cada par de variables con distintos valores en  $t$ .

Los procesos estocásticos que nos interesa conocer con más profundidad son dos, el ruido blanco y el proceso estocástico estacionario, donde este último sucede cuando las funciones de distribución conjuntas no cambian con el paso del tiempo. En relación a la estacionariedad, podemos decir que se caracteriza principalmente porque las esperanzas de las variables aleatorias son constantes y no dependen del tiempo, al igual que las varianzas. Sin embargo, las covarianzas de dos variables aleatorias obtenidas de dos períodos distintos dependen del tiempo transcurrido entre ellas.

Hay que mencionar que las series de tiempo en el mundo real son, en su mayoría, no estacionarias, con lo que cambiaría el valor medio de la serie de tiempo, ya que supuestamente existiría una tendencia en el valor medio, aumentándolo o disminuyéndolo. Para solucionar este problema, se proponen dos vías de apoyo:

1. La diferenciación de las series de tiempo.
2. La aplicación del logaritmo en las series, para luego calcular las primeras diferencias.

Cuando se realiza la diferenciación, como medio de corrección de esta no estacionalidad, es común referirse a un modelo ARIMA  $(u,v,w)$ , donde  $u$  es la cantidad de términos autorregresivos (AR);  $v$  es la cantidad de diferencias;  $w$  es la cantidad de términos de medias móviles (MA).

### Algoritmos Genéticos (AG)

Los Algoritmos Genéticos fueron dados a conocer por Holland en 1975, y consisten en una función matemática que simula el proceso evolutivo de ciertos hechos, acciones y especies, con el fin de encontrar la solución óptima a problemas específicos de maximización y minimización.

En el contexto de este seminario, el problema de optimización se define como maximizar la cantidad de aciertos en la proyección del signo de un modelo multivariado, escogiendo para ello una combinación óptima de variables explicativas.

Según Goldberg (1989), los algoritmos genéticos consisten en la búsqueda de la mecánica de la selección natural y de la genética, combinando las estructuras más aptas con estructuras de secuencia con un agregado de información y aleatorizado. Dentro de esta mecánica se tiene como variable de entrada una generación de posibles soluciones para el problema en cuestión, y como variable de salida a la estructura más apta u óptima, de manera de poder generar a futuro las mejores proyecciones posibles.

Los Algoritmos Genéticos trabajan con códigos que generan las posibles soluciones al problema planteado, pero antes de empezar a trabajar con

Algoritmos Genéticos es necesario establecer un rango de las posibles soluciones.

Existen diferentes maneras de realizar la codificación de los Algoritmos Genéticos. Entre ellas encontramos la codificación usando alfabetos con diferentes cardinalidades y la codificación que presenta las soluciones por medio de cadenas binarias, es decir, tomando valores de ceros y unos. Esta última forma es la codificación más utilizada por los diferentes analistas y, según Bauer (1994), permite ser utilizada fácilmente en aplicaciones financieras. En esta misma línea, Davis (1994) demuestra que la utilización de Algoritmos Genéticos (AG) en la calificación de créditos bancarios posee un resultado mejor que otros métodos, como el de redes neuronales, ya que los resultados en los AG se presentan de manera más transparente.

Kingdon y Feldman (1995) usaron Algoritmos Genéticos con el fin de predecir y encontrar las reglas que podrían generar una posible bancarrota de alguna compañía. Por otro lado, Pereira (1996) utilizó este método para encontrar aquellos valores óptimos que son utilizados en las distintas reglas de transacción para el tipo de cambio dólar americano en base al dólar australiano. Los resultados obtenidos mostraron resultados intramuestrales positivos, los que al ser aplicados fuera de la muestra disminuyeron, aunque siguieron siendo rentables.

Allen y Karjalainen (1999) utilizaron Algoritmos Genéticos para entender las reglas de transacción del índice S&P 500, para luego emplearlas como criterio en su estudio de análisis técnico que, cubriendo los costos de transacción, encontraron que existía un exceso de retorno calculado, la que fue descubierta gracias a la estrategia *Buy and Hold*.

Por otro lado, Feldman y Treleaven (1994) muestran que el gran problema que poseen los Algoritmos Genéticos es la dificultad que presentan para encontrar una técnica de codificación manejable, que permita determinar el tipo de selección y probabilidades de los operadores genéticos, ya que no hay reglas fijas en esta materia.

Sin embargo, Kim y Han (2000) demostraron que los AG pueden reducir la complejidad de factores irrelevantes dentro del modelo eliminándolos, obteniendo modelos mejores para predecir un índice de precios que los resultantes por métodos convencionales.

En este seminario, al igual que Díaz (2005), se ocupó una combinación binaria para representar las posibles soluciones al problema que correspondan a la especificación del modelo.

### Rolling Regression

La metodología *Rolling regression*, es una técnica que permite agregar nueva información al modelo que se encuentra en el proceso de iteración, de manera

que permita otorgarle mayor importancia a aquella información recién agregada y menos a aquella que es más antigua y, en algunos casos, la eliminación de esta última. En esta técnica se asume que los agentes son inteligentes y adaptan sus expectativas dando mayor ponderación a la información más reciente.

En el proceso *Rolling regresión*, se mantiene constante el tamaño del conjunto intramuestral " $m$ ", pero va cambiando su composición, dado que en cada iteración este conjunto se va desplazando a través del conjunto total de datos, de manera que incluye la siguiente observación (o información,  $Y_{t+m+1}$ ) y elimina aquella que es más antigua.

### Evaluación de la Capacidad Predictiva y Significancia Estadística

Luego de encontrar los mejores modelos por medio de la técnica de Algoritmos Genéticos, los cuales se obtuvieron gracias a la medición del desempeño relativo de cada modelo en términos del Porcentaje de Predicción de Signo (PPS), se aplicó el *test de acierto direccional* (DAT) de Pesaran y Timmermann (1992). Cabe señalar que esta medición de desempeño se realiza contando el número de predicciones correctas del signo "*hits*" de la variación del índice.

El test de Pesaran y Timmermann (1992) permite medir la significancia estadística de la capacidad de predicción de los modelos obtenidos. Sin

embargo, la sola aplicación del test no permite conocer la característica de los momentos de la distribución que sigue dicho test, por lo que podría generar conclusiones sesgadas, en las cualidades de la muestra en cuestión.

## V. Análisis de los Resultados

Los diferentes modelos estudiados y analizados en esta investigación fueron contruidos con variables de series de tiempo con hasta un máximo de 4 rezagos para la rentabilidad de la acción en análisis, considerados en forma semanal; 4 posibles rezagos para la rentabilidad del índice Dow Jones Industrial, considerados en forma semanal; 4 rezagos para la rentabilidades de las acciones de los días anteriores; y lo mismo para el índice Dow Jones Industrial; y, por último, 4 rezagos del error del modelo aplicado.

Primero se usó la metodología de Algoritmos Genéticos para conducir el proceso de búsqueda de los mejores modelos, a partir de todas las combinaciones posibles de las variables independientes ya señaladas. Luego se aplicó la metodología *rolling regression* a los mejores modelos obtenidos, y se evaluó el porcentaje de predicción de signo que tenía cada modelo predictivo (el PPS) en particular, con el fin de encontrar aquel modelo que proyectase mejor los movimientos del signo de los precios de las acciones. El procedimiento anterior se realizó comparando el signo de la proyección de la variación observada en cada *t-ésimo* período, donde  $t = 1, 2, \dots, n$ , por lo que si los signos entre la proyección y la observación coincidían, entonces la efectividad del modelo analizado aumentaba y, en caso contrario, la efectividad o capacidad predictiva del modelo disminuía.

La investigación se realizó utilizando una base de 261 datos semanales y realizando un proceso de *rolling regresión* con un tamaño de 30 datos para el conjunto intramuestral. Como resultado de lo anterior, se encontró que los mejores modelos para proyectar el signo de las variaciones del precio de las acciones eran:

$$\Delta\%Y_t = c + \Delta\%aa_{t-1} + \Delta\%dji_{t-1} + \Delta\%dji_{t-2} + \Delta\%dji_{t-3} + \Delta\%aa_{d-1} + \Delta\%aa_{d-2} + \Delta\%aa_{d-4} \quad (\text{i})$$

$$\Delta\%Y_t = c + \Delta\%aa_{t-1} + \Delta\%dji_{t-1} + \Delta\%dji_{t-2} + \Delta\%dji_{t-3} + \Delta\%dji_{t-4} + \Delta\%aa_{d-1} + \Delta\%aa_{d-2} + \Delta\%aa_{d-4} \quad (\text{ii})$$

$$\Delta\%Y_t = c + \Delta\%dji_{t-1} + \Delta\%dji_{t-4} + \Delta\%aa_{d-1} + \Delta\%aa_{d-2} + \Delta\%aa_{d-4} + e_{t-2} \quad (\text{iii})$$

donde,

- i)  $\Delta\%Y_t$  es la rentabilidad del precio de la acción Alcoa Inc., en el período  $t$ .
- ii)  $\Delta\%aa_{t-1}$  es el rezago de la rentabilidad del precio de la acción en el período  $t-1$ .
- iii)  $\Delta\%dji_{t-1}$ ,  $\Delta\%dji_{t-2}$ ,  $\Delta\%dji_{t-3}$  y  $\Delta\%dji_{t-4}$  son los rezagos de la rentabilidad del índice Dow Jones Industrial para los períodos  $t-1$ ,  $t-2$ ,  $t-3$  y  $t-4$  respectivamente.
- iv)  $\Delta\%aa_{d-1}$ ,  $\Delta\%aa_{d-2}$  y  $\Delta\%aa_{d-4}$  son los rezagos de la rentabilidad del precio de las acciones, para los períodos  $d-1$ ,  $d-2$  y  $d-4$ , respectivamente.
- v)  $e_{t-2}$  es el rezago del error de medición en el período  $t-2$ .
- vi)  $c$  es la constante de cada modelo.

Cabe mencionar que  $t$  hace referencia a los rezagos en términos de semanas y  $d$  hace referencia a rezagos en términos de días y, además, que ambas terminologías parten en el mismo tiempo en que se realizó la proyección.

El primer modelo alcanzó un porcentaje de predicción del signo de un 70,13%, logrando ser el mejor modelo obtenido con la técnica de Algoritmos Genéticos. El segundo modelo obtuvo un 69,7%, y el tercer modelo un 68,3%. Los resultados de los modelos predictivos se muestran en el cuadro 1.

Cuadro 1:  
Resultados de los mejores modelos predictivos

<b>Modelos</b>	<b>PPS</b>	<b>DA</b>	<b>Akaike</b>	<b>Schwarz</b>
<b>Modelo ( i )</b>	71,83%	6,20	-5,64	-5,39
<b>Modelo ( ii )</b>	71,83%	6,21	-5,65	-5,40
<b>Modelo ( iii )</b>	73,16%	6,59	-5,96	-5,71

Sin embargo, después de realizar el proceso *rolling regression*, el primer modelo alcanzó un porcentaje de predicción del signo de un 71,86%, equivalente al del segundo modelo. El tercer modelo alcanzó un PPS de 73,16%. Se aplicó además, a los distintos modelos, el test de acierto direccional de Pesaran y Timmermann (1992). Si el test alcanza un valor mayor que 2, se puede concluir que el modelo posee capacidad predictiva, y si alcanza un valor menor a 2, se concluye que el modelo no posee capacidad predictiva. Para el primer modelo se obtuvo un valor de 6,2; para el segundo modelo se obtuvo un valor de 6,21; y, por último, para el tercer modelo se obtuvo un valor de 6,59. Se concluye de esta forma, que los tres mejores modelos obtenidos a través de la técnica de Algoritmos Genéticos poseen una

capacidad predictiva significativa de la variación del signo del precio de la acción de Alcoa Inc.

De acuerdo al test DAT se sigue apoyando los resultados obtenidos utilizando la metodología *rolling regression*, en donde el tercer modelo encontrado a través de Algoritmos Genéticos posee el mayor porcentaje de predicción del signo frente a los otros dos modelos, teniendo así un 71,91% el tercer modelo, y el primer y segundo modelo un 70,64%.

Por último, es importante referirse a los criterios de información de Akaike y Schwarz, donde para el primer modelo encontramos un valor de  $-5,64$  para Akaike y  $-5,39$  para el índice de Schwarz; para el segundo modelo encontramos un valor de  $-5,65$  para Akaike y un  $-5,4$  para el criterio de Schwarz; y, finalmente, para el tercer modelo los resultados fueron de  $-5,96$  para Akaike y  $-5,71$  para Schwarz. Lo anterior, permite identificar el modelo más eficiente, que será aquél que presente el menor número en su indicador de información en relación a los otros, y que será utilizado para proyectar el signo de las variaciones del precio de la acción de Alcoa Inc.

Es necesario mencionar además que, como forma de comprobación, se estudiaron los resultados obtenidos en el modelo uno, pero sin las variables diarias, es decir, sin las variables  $\Delta\%aa_{d-1}$ ,  $\Delta\%aa_{d-2}$  y  $\Delta\%aa_{d-4}$ . Esto con el fin de observar el impacto que producía el hecho de no tener dentro del mejor

modelo estimado a través de Algoritmos Genéticos variables de rezagos diarios, sino, sólo de carácter semanal. Para este nuevo modelo, se realizó la metodología *rolling regression* y se aplicó el test DA de Pesaran y Timmermann (1992). Los resultados obtenidos fueron un PPS de 56,7% utilizando la técnica *rolling regression*, DA es de 1,13 y, por último, el criterio de Akaike y Schwarz es de  $-7,12$  y  $-6,87$ , respectivamente. En función de lo anterior, se puede concluir que el modelo no posee capacidad predictiva, y que incluir los rezagos diarios como es sugerido en este seminario puede mejorar la capacidad predictiva.

## VI. Conclusiones

La principal conclusión de este seminario es que la capacidad predictiva de los movimientos de los precios de la acción estudiada, Alcoa Inc., mejora considerablemente si se incluyen variables rezagadas diarias, en relación a otros estudios realizados anteriormente, los cuales contenían sólo variables rezagadas semanalmente. Esta mejora implica un alza de la capacidad predictiva del modelo (4) desde el 58,78% a 70,12% en el período de la investigación.

El mejor modelo obtenido a través de Algoritmos Genéticos logró tener una capacidad predictiva de un 70,12%. Además, se pudo constatar que la capacidad predictiva de los tres modelos encontrados a través de la metodología de Algoritmos Genéticos resultó significativa de acuerdo al test de Pesaran y Timmermann (1992).

Además, el trabajo valida que los Algoritmos Genéticos en los Modelos Multivariados Dinámicos pueden ser usados como una metodología alternativa en la mejora de la construcción de los modelos de proyección de series de tiempo, en función de su capacidad predictiva del signo, con el fin de apoyar a *traders* y financistas en el proceso de la conformación de diversos portafolios de inversión.

Se plantea, para futuras investigaciones, que el hecho de incluir variables explicativas que son referentes a los días anteriores, no necesariamente permitirán que los beneficios económicos obtenidos sean mayores, siendo necesario estudiar las rentabilidades obtenidas por un modelo con variables diarias versus uno con variables semanales. Como recomendación e ideas para futuros seminarios, se sugiere validar el modelo con otras acciones del mercado tanto estadounidense como local. Además, también se sugiere efectuar una prueba de robustez del estudio, como por ejemplo realizar un análisis *bootstrapping*.

## VII. Bibliografía y Referencias

- Allen, F. y Karjalainen, R. (1999), "Using genetic algorithms to Find technical trading rules", *Journal of Financial Economics* 51, pp.245-271.
- Bauer, R. J. Jr (1994), "Genetic Algorithms and Investment Strategies", *Wiley Finance Editors, John Wiley and Sons*, New York.
- Blume, L., Easley y M O'Hara (1994), "Market Statistics and Technical Analysis: The Role of Volume", *Journal of Finance* 49, pp.153-192.
- Chopra, N., Lakonishok J. y Ritter J. R.(1992), "Measuring abnormal returns: Do stocks overreact?", *Journal of Financial Economics* 31, pp. 235-268.
- Cortés, E., Figueroa, A. (2006), "Modelos Arimax y Óptimos de Rolling para la predicción de variación del precio de las acciones usando índices de Mercado y Sectoriales", Tesis de Magíster en Finanzas, Universidad de Chile.
- De Bondt, W. F. M. y Thaler R. (1985), "Does the stock market overreact?", *Journal of Finance* 40, pp. 793-805.
- Díaz, D (2005), "Prediciendo el cambio de signo del Precio del Oro: Modelos de Redes Neuronales Rolling y Recursivos", Tesis de Magíster en Finanzas, Universidad de Chile.
- Davis, L. (1994), "Genetic Algorithms and Financial Applications", Deboeck (ed. 1994), pp. 133-147.
- Fama, E. (1970), "Efficient Capital Markets: A Review of theory and Empirical Work". *Journal of Finance* 25, pp. 383-417.
- Fama, E. y French, K. R. (1988), "Permanent and temporary Components of Stocks prices". *Journal of Political Economy* 98, pp.247-273.
- Ferson, W., y C. Harvey (1991), "The Variation of Economic Risk Premiums", *Journal of Political Economy* 99, pp.385-415.
- Ferson, W. y Schadt, R. (1996), "Measuring fund strategy and performance in changing economic conditions", *Journal of Finance* 51, pp.425-462.

- Hodgson, A., y D. Nicholls (1991), “The Impact of Index Futures Markets on Australian Share Market Volatility”, *Journal of Business Finance and Accounting* 18, pp.267-280.
- Jegadeesh, N. (1990), “Evidence of predictable behavior of security returns”, *The Journal of Finance* 45, pp. 881-898.
- Kim, K. y Han, I. (2000), “Genetic algorithms approach to feature discretization in artificial neural networks for the predictions of stock price index”, *Expert Systems with Applications* 19, pp. 125-132.
- Kingdon, J. y Feldman, K. (1995), “Genetic algorithms and applications to finance”, *Applied Mathematical Finance* 2, pp.89-116.
- Leung, Mark T., Hazem Daouk y An- Sing Chen (2000), “Forecasting Stock Indices: A Comparison of Classification and Level Estimation Models”, *International Journal of Forecasting* 16, pp.173-190.
- Lo, A. y MacKinley, A. C. (1988), “Stock Market Price do not Follow Random Walk: Evidence from a Simple Specification Test”, *Review of Financial Studies* 1, pp.41-66.
- Márquez, A. (2006), “Redes Neuronales con Rasgos de Conciencia: Aplicación en la predicción de signo de cambios de precios en los componentes del Dow Jones Industrial Average Index”, Tesis Ingeniería Comercial, Universidad de Chile.
- O’ Connor, M., Remus y K.Griggs (1997), “Going Up-Going Down: How good are people at Forecasting Trends and Changes in Trends?”, *Journal of Forecasting* 16, pp. 165-176.
- Parisi, A. (2002), “Evaluación de modelos de Redes neuronales de predicción del signo de la variación del IPSA”, *Estudios de Administración* 9, pp. 67- 103. Chile.
- Parisi, A., Parisi, F. y Cornejo, E. (2004), “Algoritmos genéticos y modelos multivariados recursivos en la predicción de índices bursátiles de América del Norte: IPC, TSE, Nasdaq y DJI”, *El Trimestre Económico* 284, pp. 789-809.
- Parisi, A., y Gutierrez, M. (2004), “Administración de carteras con Redes Neuronales mediante metodología Rolling”, Tesis de Magíster en Finanzas, Universidad de Chile.
- Parisi, A., Parisi F. y Díaz, D. (2006), “Modelos de Algoritmos Genéticos y Redes Neuronales en la Predicción de Índices Bursátiles Asiáticos”, *Cuadernos de Economía* 128, pp. 251-284.

- Pereira, R. (1996), "Selecting parameters for technical trading rules using genetic algorithm", *Journal of Applied Finance and Investment* 1, pp. 27-34.
- Pesaran, M. H., y Timmermann, A. (1992), "A Simple Nonparametric Test of Predictive Performance", *Journal of Business and Economic Statistics* 10, pp. 461-465.
- Poterba, J.M., y Summers L.H (1988), "Mean Reversion in Stock Prices: Evidence and Implications", *Journal of Financial Economics* 22, pp.27-59.
- Wu, Y. y Zhang, H. (1997), "Forward premiums as unbiased predictors of future currency depreciation: A non-parametric analysis", *Journal of International Money and Finance* 16, pp. 609-623.
- Yoda, M. (1994), "Predicting the Tokyo stock Market", G.J. Deboeck (comp.), *Trading on the Edge: Neural, Genetic and the Fuzzy Systems for Chaotic Financial Markets*, New York, Wiley, pp.66-79.

## IX. Anexos

### Anexo 1

#### 1. Gráfico Alcoa, Inc. para un período de 5 años.



#### 2. Gráfico índice Dow Jones Industrial para un período de 5 años.

