

# TRABAJO FINAL DE MÁSTER

---

**Título:** Modelos de Predicción de Índices de Mercados de Valores mediante el uso de la Lógica Difusa

**Autoría:** Alejandro Bustelo De la Casa

**Tutoría:** Ana María Gil Lafuente

**Curso académico:** 2017-2018

Facultad de Economía y Empresa

Universidad de Barcelona

Trabajo Final de Máster

Máster en Ciencias Actuariales y Financieras

**Modelos de Predicción de  
Índices de Mercados de  
Valores mediante el uso de  
la Lógica Difusa**

Autoría: Alejandro Bustelo De la Casa

Tutoría: Ana María Gil Lafuente

“El contenido de este documento es de exclusiva responsabilidad del autor, quien declara que no ha incurrido en plagio y que la totalidad de referencias a otros autores han sido expresadas en el texto”

## **1. Resumen**

Mediante el siguiente estudio, se va a ahondar en la evaluación de la predicción de una serie histórica del NASDAQ-100, aplicando la metodología ANFIS, concretamente mediante el uso de la *neuro-fuzzy*. Por lo tanto, se puede hablar de la aplicación de modelos de lógica difusa a una serie de datos históricos, a fin de establecer una teoría predictiva del comportamiento de los valores y su proyección, para la toma de decisiones por parte de los *traders*. Con todo, compararemos la predicción con dicha metodología con la encontrada a través de la metodología *Box-Jenkins* sobre los mismos datos históricos.

## **1. Abstract**

Through the following study, we will delve into the evaluation of the prediction of the NASDAQ-100, using and applying the ANFIS methodology. Therefore, we can speak about the application of fuzzy logic models to a series of historical data of the NASDAQ-100, in order to establish relations of belonging to the variables, with the purpose of establishing a predictive theory of the behavior of the values and its projections for the decision, which are made by the traders. To complete the thesis, the reader will be able to see if the prediction has become better or worse than the Box-Jenkins Methodology prediction.

---

**Palabras clave:** [*Fuzzy logic*, predicción, *Box-Jenkins*, *Neuro-Fuzzy*, decisión]

## Índice

1. Resumen.....	1
2. Introducción.....	3
3. Breve historia de la Lógica Difusa.....	4
4. Contextualización.....	6
5. Descripción Teórica de los elementos.....	9
5.1 Propiedades.....	9
5.2 Subconjuntos Borrosos.....	13
5.2.1 Subconjuntos Borrosos Discretos.....	14
5.2.1 Subconjuntos Borrosos Continuos.....	14
5.3 Función de Pertenencia.....	14
5.4 Números Borrosos.....	14
5.4.1 Números Borrosos Triangulares.....	15
6. Exposición de los Datos y del paper <i>Modeling Chaotic Behavior Of Stock Indices Using Intelligent Paradigms</i> [6].....	15
7. Análisis de la Serie mediante Metodología <i>Box-Jenkins</i> .....	22
7.1 Descripción Teórica.....	22
7.2 Análisis de la Serie Escogida.....	25
7.2.1 Análisis de la Estacionariedad de la Serie.....	25
7.2.2 Análisis de los Estadísticos Básicos de la Serie.....	26
7.3 Identificación, estimación y diagnóstico de Modelo para Media y Varianza.....	28
7.3.1 Modelo para la Media.....	31
7.3.2 Predicción y Evaluación del Modelo.....	33
7.3.3 Modelo para la Varianza y estudio de la Volatilidad de la serie.....	35
8. Conclusiones.....	38
Anexo 1: Teoría de Conjuntos Borrosos .....	39
Anexo 2: Código R.....	44
Anexo 3: Leyenda Gráficos .....	51
9. Referencias .....	52

## **2. Introducción.**

Finalizar el Máster de Ciencias Actariales y Financieras redactando las siguientes líneas hace que sienta un ápice de nostalgia a la vez que orgullo. Atrás quedan dos años de estudio y desarrollo personal, dos años en que la forma de ver las cosas ha virado por completo. Considero el definirse, y hallar el perfil dónde poder desarrollar y potenciar las habilidades personales, como uno de los pasos más importantes que se pueden dar.

Precisamente, de este proceso de determinación nace la idea de este TFM. Aun amando la autenticidad que el empirismo otorga y la concreción que la estadística procura, siento la necesidad de expresar, mediante herramientas de dichas disciplinas, la proposición de que no todo se puede explicar mediante los números. El lenguaje [1]. Por suerte, la lógica difusa lo barruntó antes que yo.

El grado de pertenencia de las “cosas” a un cierto conjunto añade a las decisiones la capacidad de pensar en tonalidad de grises y no solo en blanco y negro, o entre 0 o 1, de esto trata a groso modo la metodología empleada en el trabajo.

Las sensaciones, las preferencias, todo lo que la teoría económica aboga, influyen en la toma de decisiones de los inversores, y por ende, en la evolución de los valores en los mercados financieros. Me pregunto si somos capaces de cuantificar esas sensaciones, si realmente las podemos ordenar y agrupar. Ciertamente, considero que hay estudios en los que modelos matemáticos comunes, estadísticos o econométricos, es decir, todos los cuantitativos son incapaces de captar las características intrínsecas de los aspectos cualitativos a los que los inversores se refieren cuando tratan con activos.

Pues bien, en las siguientes páginas se va a tratar de explicar al lector cómo los elementos de la *fuzzy logic* son tratados y usados mediante redes neuronales en paradigmas inteligentes, para medir el error de predicción de dicha metodología y compararla con el error de aplicación y predicción de una metodología contrastada de tratamiento de series, como es la *Box-Jenkins*.

### 3. Breve historia de la Lógica Difusa.

Para hablar de **lógica difusa** hay que remontarse a periodos históricos lejanos a nuestra contemporaneidad. Pues bien, ya *Aristóteles* o *Platón* consideraban las cosas con la cualidad de pertenecer a diversas categorías, agregando un nivel de pertenencia a cada una de ellas, o mejor dicho, explicaban que las cosas no pertenecen a un nivel u otro, sino que existe una escala intermedia entre ambos niveles. Considerando también, niveles de falsedad o veracidad a los sucesos. Tras ellos, *Hume* o *Kant* detectaron contradicciones en la lógica clásica. A principios del siglo XX, el filósofo y matemático británico *Bertrand Russell* divulgó la idea de que la lógica produce contradicciones. Realizó un estudio sobre las vaguedades del lenguaje, concluyendo con precisión que la vaguedad es un grado. También en este tiempo *Ludwing Wittgenstein*, estudió las diferentes acepciones que tiene una misma palabra. Éste llegó a la conclusión de que en el lenguaje una misma palabra expresa modos y maneras diferentes. [4]

En 1920, *Jan Lukasiewicz*, desarrolló la primera lógica de vaguedades. Para él los conjuntos tienen un posible grado de pertenencia con valores que oscilan entre 0 y 1, y en este intervalo existen un número infinito de valores. [4]

El padre del término "borroso" fue *Lofti Asier Zadeh* cuando en 1965 publicó "**Fuzzy Sets**" (**Conjuntos Difusos**). Las tesis que propone surgen del estudio de pensadores de distintas disciplinas que como él, tenían una visión de los problemas diferente de la lógica tradicional. La paradoja del conjunto de *Bertrand Russell*, el principio de incertidumbre de la física cuántica de *Werner Heisenberg*, la teoría de los conjuntos vagos de *Max Black* y la aportación de *Jan Lukasiewicz*, influyeron para que *Zadeh* publicase sus ensayos en diversas revistas y medios del sector. [4]

Al comienzo las ideas publicadas por *Zadeh* no calaron en la comunidad científica del momento, pero con el tiempo comenzó a tener seguidores, esto produjo que sus teorías fuesen ampliadas y se asentaran sus conocimientos.

La intención de *Zadeh* era la creación de un formalismo para manejar de forma más eficiente la imprecisión del razonamiento humano. Es en 1971, cuando realiza la publicación de "*Quantitative Fuzzy Semantics*" en donde aparecen los elementos formales que dan lugar a la metodología de la Lógica Borrosa y de sus aplicaciones tal y como se conocen en la actualidad, es una publicación de severa trascendencia en la evolución y desarrollo de la lógica borrosa.

A partir de 1973, con la teoría básica de los controladores borrosos de *Zadeh*, otros investigadores comenzaron a aplicar la lógica borrosa a diversos procesos. Se establecen varios grupos de investigación en lógica difusa en algunas pequeñas universidades japonesas.

En 1974 *Assilian* y *Mamdani* en el Reino Unido desarrollaron el primer controlador difuso diseñado para la máquina de vapor. La implantación real de un controlador de este tipo no fue realizada hasta 1980 por *F.L. Smidth & Co* en una planta de cemento en Dinamarca.

En 1987 *Hitachi* usa un controlador *fuzzy* para el control del tren de *Sendai*. Es también en este año cuando la empresa *Omron* desarrolla los primeros controladores difusos comerciales y es que 1987 es considerado como el "*fuzzy boom*" debido a la gran cantidad de productos basados en Lógica Borrosa que se comercializan.

En 1993, *Fuji* aplica la Lógica Borrosa para el control de inyección química en plantas depuradoras de agua por primera vez en Japón. Ha sido precisamente aquí, donde más apogeo ha tenido la Lógica Borrosa, creándose estrechas colaboraciones entre el gobierno, las universidades y las industrias, estableciendo proyectos llevados a cabo por el Ministerio de Industria y Comercio (MITI) y la Agencia de Ciencia y Tecnología (STA) en consorcio con el *Laboratory for International Fuzzy Engineering Research (LIFE)*.

De forma paralela al desarrollo de las aplicaciones de la lógica difusa, *Takagi* y *Sugeno* desarrollan la primera aproximación para construir reglas fuzzy a partir de datos de entrenamiento. [6]

Otro factor decisivo para continuar con la investigación de este campo es el interés en las redes neuronales y su semejanza con los sistemas fuzzy. Se buscan relaciones entre las dos técnicas obteniéndose como resultado los sistemas neuro-fuzzy, que usan métodos de aprendizaje basados en redes neuronales para identificar y optimizar sus parámetros. Para finalizar, aparecen los algoritmos genéticos que sumados a las redes neuronales y los sistemas *fuzzy* son herramientas de trabajo muy potentes en el campo de los sistemas de control. Para terminar esta breve reseña sobre la historia de dicha disciplina, así se refería *Zadeh (1977)* [2]

***“El mundo actual es incierto e impreciso, los actos de los hombres y las relaciones entre ellos están afectadas por la vaguedad, es por esto que la realidad no puede estudiarse en términos absolutos con técnicas aplicables a situaciones ciertas ni aun aleatorias. Casi toda la lógica del razonamiento humano no es la lógica clásica de dos valores, o incluso de varios, sino una lógica de verdades borrosas, de conjunciones borrosas, de reglas de deducción borrosas”***

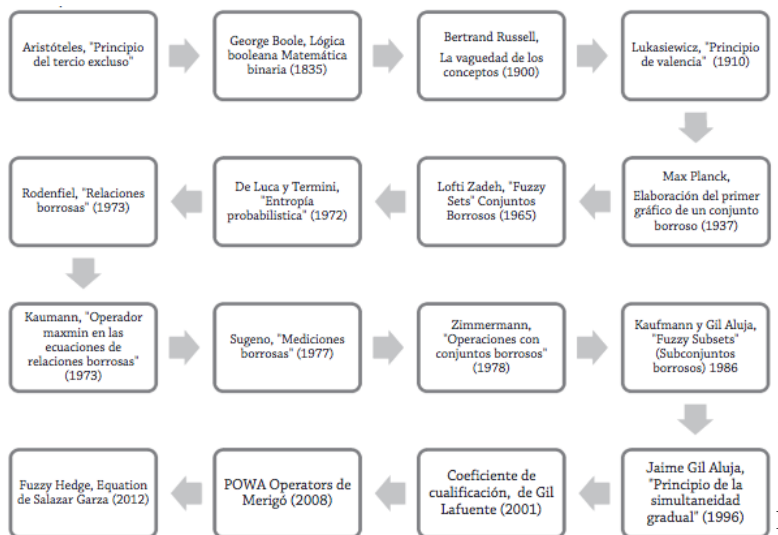


Figura 1

<sup>1</sup>Así lo cuenta Gumaro Álvarez Vizcarra en su trabajo: Lógica Borrosa, Efectos Olvidados y exposición al riesgo cambiario.



#### 4. Contextualización.

– ¿Pero y qué ocurre en el mundo de las finanzas? –

Esta es la pregunta que ha motivado la escritura del presente estudio. Y si se hallase la manera de conectar la toma de decisiones en un contexto de incertidumbre aplicando la metodología de la lógica borrosa, con el fin de predecir los valores financieros, destinados a ser el motor del modelaje en la mencionada toma de decisiones.

La manera de tener en cuenta las sabidas variables inciertas, que a su vez pesan notablemente en la evolución de los valores financieros, es la valoración de los conceptos *fusca*.

Desarrollar la contextualización mediante lo que Gil Aluja y Kaufmann establecieron en *Kaufmann y Gil Aluja (1993, 117 – 118)* acerca de los datos es un buen comienzo. Se cuenta con “datos aleatorios” cuando se conocen los estados y únicamente se pueden afectar los valores asociados a estos estados a través de leyes de probabilidad. El proceso de tratamiento se convierte en un proceso estocástico. También pueden encontrarse sistemas en donde los propios estados son aleatorios. Se está en presencia de “datos inciertos” cuando no pueden obtenerse leyes de probabilidad debido al conocimiento parcial o al desconocimiento de la forma de dicha distribución. [4]

Se pasa de lo medible a lo valuable, ya que las valuaciones son captadas a través de estimaciones numéricas subjetivas, puesto que en estas situaciones, el camino hacia el conocimiento de los problemas de gestión económico -financiera, parte de la percepción de elementos inherentes al proceso bajo estudio. Se ingresa entonces en el ámbito de aplicación de la matemática de la incertidumbre: intervalos de confianza, tripletas de confianza, números borrosos, técnicas de regresión borrosa, métodos de optimización borrosa, números híbridos o variables borroso aleatorias (donde convergen aleatoriedad y borrosidad), grafos neuronales, alfa cortes, etc.

Con esto, notamos la incipiente falta de capacidad de la probabilidad de procesar toda aquella información basada en la percepción o en la sensación. El punto que debemos atacar y que concierne al mundo de las finanzas es aquel dónde lo incierto y lo aleatorio convergen.

De entre muchas de las razones por las cuales relacionar los mecanismos de la lógica borrosa con el mundo de las finanzas, se destacan tres de suma relevancia para entender aquello en lo que se está profundizando. [4]

- *Razón probabilística.* La teoría estadística es frecuentista, y no aborda el cambio y la dependencia estructural futura de las variables que inciden sobre la economía y los negocios. La lógica borrosa evita la pérdida de información que se produce con el uso de métodos estadísticos que realizan manipulación numérica, puesto que trabajan con promedios, ocultando los valores extremos.
- *Razón informacional.* Los objetos probabilísticos de la gran mayoría de modelos económicos y financieros prevalecientes no incorporan información cualitativa de

gran valor. La lógica borrosa permite abandonar la noción preconcebida respecto a la manera en que los datos deben comportarse, permitiendo entender que las decisiones económicas y financieras no están afectadas por la diferencia entre el valor actual de la variable que se observa y algún valor teórico normativo de equilibrio. La lógica borrosa permite superar el simplista enfoque aristotélico para la noción de equilibrio (estás o no en equilibrio), para concebirlo como una paradoja que presenta grados de pertenencia hacia las esquinas del equilibrio y desequilibrio absolutos, los extremos ideales que no siempre se alcanzan. Esto permite entender que la realidad económica y financiera transcurre entre estos dos opuestos la mayor parte del tiempo.

- *Razón conductual.* Los seres humanos poseen racionalidad limitada y son proclives al error, lo cual separa las predicciones de los modelos “exactos” y de “equilibrio general”, de las observaciones de los hechos. A diario, los datos económicos y financieros muestran los patrones de conducta de los agentes, los cuales sintetizan el grado de borrosidad y asimetría de los impulsos neuronales de los agentes humanos.

En este sentido, los modelos financieros borrosos aplicados a la planificación financiera y a la administración de portafolios de inversión, se descubren como interesantes, básicamente por su utilidad para establecer rangos de incertidumbre, así como intervalos de riesgo – rentabilidad que permitan vincular razonablemente la optimización financiera con la formación de expectativas –

Una vez situado el lector en el contexto financiero, no es algo banal el entender que tras la aplicación metódica de la lógica borrosa, existe un profundo trasfondo. Precisamente, la profundidad de dicho trasfondo reside en la teoría de la decisión, en este contexto de incertidumbre severa. Podemos adentrarnos pues, en la enumeración y contenido de estos elementos fundamentales para la decisión.

Hablar de estos elementos primordiales, es hablar obligatoriamente de 4 componentes: [3]

### ***1. Relación***

Las relaciones son la vida social, es el alfa de la socialización, el vínculo, la expresión, el interactuar con diversos elementos, también a efectos empresariales es vinculante. Afectando a todo ser inteligente, animales e incluso objetos físicos o mentales.

La vida se desarrolla en círculos, los cuales representan las relaciones existentes, pasando durante el desarrollo de un círculo a otro, incrementando o no la intensidad de la relación contenida en mencionado círculo. Es exactamente, la palabra intensidad, la que introduce el concepto de grado, y precisamente, el grado de intensidad es lo que personifica la importancia de la lógica borrosa en todo esto. Históricamente, la lógica booleana ha tratado de discernir la existencia o no de relación entre elementos. La lógica borrosa busca un grado de relación entre ellos, el mero ejemplo de las relaciones humanas lo clarifica; no siempre tenemos la misma relación con diferentes personas, puede haber una relación de amistad con una intensidad enorme, y puede haber otra con una intensidad menor, lo cual no significa que dicha relación sea inexistente. Es decir, degradación o fortalecimiento de los vínculos.

Con la base expuesta, la creación, ampliación o disminución de las redes de relación entre elementos se construye a medida que transcurre el tiempo y las etapas, a la vez que se pueden generar nuevas relaciones dentro de un vínculo en concreto, estableciéndose así los conocidos subvínculos.

## **2. Asignación**

Proceso conocido también como afectación o adscripción, se basa en la calificación de las relaciones. Basándose pues en la consideración fundamental de un conjunto de elementos a asignar y de otro conjunto de elementos que recibirá asignación.

El planteamiento de la asignación descansa sobre tres conjuntos, normalmente finitos, de objetos físicos o mentales. El primero recoge los elementos a asignar, el segundo recoge los elementos que deben recibir asignación y, por último, el tercero recaba las características o *inputs* en los cuales se basa el proceso de asignación.

El procedimiento se tecnifica mediante descriptores, es decir, se construye un subconjunto borroso para cada uno de los objetos a asignar tomando como referencia el conjunto de criterios de asignación. Lo mismo se hace para el conjunto de criterios, obteniendo entonces los objetos receptores. Con esto, se adecuan los conceptos de “alejamiento” o “acercamiento” entre objetos.

## **3. Agrupación**

Si nos ceñimos en la actividad financiera como tal, la agrupación de instrumentos de financiación y productos es algo que se encuentra en la mente de los decisores en tal campo. El hecho de poder decidir sin agrupar homogéneamente ha sido y es uno de los factores desencadenantes de la imprecisa toma de decisiones.

Para hablar de ello, se toma en consideración el concepto de semejanza, contando que la no existencia de la propiedad transitiva le resta eficacia. Los subgrafos transitivos a partir de un grafo de semejanzas forman subconjuntos no disjuntos a partir de las propiedades de reflexividad y simetría, disponer de estos elementos teóricos cobra máxima importancia en el proceso. No cabe duda, que todo se construye a raíz del concepto de afinidad y técnicamente, a partir de una matriz cuadrada en la que los elementos de las filas coinciden con los de la columna en cantidad y en esencia.

## **4. Ordenación**

Este último paso culmina la pirámide teórica de la decisión en un contexto de incertidumbre, constatándose el problema base expuesto en la línea argumental de todo lo explicado hasta ahora; seguimos sin poder fijar un número a cada fenómeno susceptible. Debemos usar las herramientas borrosas que nos conduzcan a algoritmos aptos para la ordenación a través de las relaciones entre grafos de formas matriciales y sagitadas. Tratándose de encontrar grafos suficientemente convexos (no permiten hallar un orden en sus vértices, representados por matrices de una sola equivalencia) es el modo de proceder a la ordenación de clases de equivalencia.

En tal contexto actual, en los sistemas sociales y económicos en los que nos encontramos, la capacidad de diferenciar por grupos a la vez que tener claros los grupos de diferenciación de todo elemento o componente significativo en una decisión es la antesala a una toma de decisión racional.

## 5. Descripción teórica de los elementos.

### 5.1 Propiedades

En la siguiente sección, el lector va a disponer de una breve descripción de los elementos que posteriormente se usarán en la aplicación de la lógica difusa. Especialmente los elementos matemáticos de la incertidumbre. Cuando se dispone de un par de conjuntos representados en un grafo en forma de matriz o sagitada, se tienen una relación  $R(x, y)$   $x, y, \in E$ , generalmente el conjunto de conexiones establecidas entre los elementos de cada conjunto se manifiesta de diferente manera, lo que origina el surgimiento de propiedades que deben cumplirse para establecer el tipo de relación existente: [5]<sup>2</sup> y [8]

**Reflexividad:** Cuando se presenta el caso de matrices donde hay coincidencia entre filas y columnas en número y esencia, se tiene la diagonal principal llena de 1; esto resulta, en consecuencia, de la relación de un elemento de un conjunto  $x \in E$  consigo mismo, es decir,  $x \in E$  es total<sup>3</sup>

$$\forall \alpha_i \in E (i=1,2,\dots,n) : \mu_{\alpha_i, \alpha_j} = 1 \text{ si } (i = j)$$

$$\mu_{\alpha_i, \alpha_j} \in [0,1], \text{ si } (i \neq j)$$

Esto es válido en todos los contextos, tanto en las matrices borrosas rectangulares como en las matrices booleanas cuadradas. Una matriz borrosa quedaría de la manera mostrada en la figura siguiente.

$$\tilde{R} = \begin{array}{c|cccc} & \alpha_1 & \alpha_2 & \dots & \alpha_v \\ \hline \alpha_1 & 1 & 0 & \dots & \mu_{1v} \\ \alpha_2 & \mu_{21} & 1 & \dots & \mu_{2v} \\ \dots & \dots & \dots & 1 & \dots \\ \alpha_v & 0 & \mu_{v2} & \dots & 1 \end{array}$$

Figura 2

<sup>2</sup> La primera parte de la sección: relaciones, son extraídas de [5]

<sup>3</sup> Gil Aluja, 1999:44

**Antisimetría:** La relación entre dos elementos del referencial E da lugar, en determinados casos, a una cierta, pero no a una total reciprocidad, en el sentido de que cuando  $\alpha_i$  está relacionada con  $\alpha_j$ , también  $\alpha_j$  está relacionada con  $\alpha_i$ , y esto para todo  $i, j = 1, 2, \dots, n$ , el grado de intensidad de la relación no es el mismo, se puede escribir de la siguiente manera:

$$\forall \alpha_i, \alpha_j \in E, \alpha_i \neq \alpha_j; \mu_{\alpha_i, \alpha_j} \neq \mu_{\alpha_j, \alpha_i}$$

No se rompe la antisimetría, si cuando alguno de los valores de  $\mu_{\alpha_i, \alpha_j}$  son iguales a cero, también los de  $\mu_{\alpha_j, \alpha_i}$  son iguales a cero; es decir,  $\mu_{\alpha_i, \alpha_j} = \mu_{\alpha_j, \alpha_i} = 0$ . Esta característica es muy amplia y acumula una gran variedad de relaciones, aunque basta con que haya un solo caso en que  $\mu_{\alpha_i, \alpha_j} = \mu_{\alpha_j, \alpha_i} \neq 0$  para que esto se rompa<sup>4</sup>. Esta antisimetría, a través de una matriz, se observa de la siguiente manera:

$$\tilde{R} = \begin{array}{c|cccc} & \alpha_1 & \alpha_2 & & \alpha_n \\ \hline \alpha_1 & \mu_{11} & \mu_{12} & \dots & 0 \\ \alpha_2 & \mu_{21} & \mu_{22} & \dots & \mu_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_n & 0 & \mu_{n2} & \dots & 0 \end{array}$$

Figura 3

$$\mu_{ij} \in [0,1], (i, j = 1, 2, \dots, n)$$

$$\mu_{ij} \neq \mu_{ji} \text{ o bien } \mu_{ij} = \mu_{ji} = 0, i \neq j$$

**Simétricas:** Las relaciones simétricas entre elementos de un referencial son aquellas en las que la intensidad de la relación de  $\alpha_i$  hacia  $\alpha_j$  es la misma que la de  $\alpha_j$  hacia  $\alpha_i$ . Se escribe de la siguiente manera:

$$\forall \alpha_i, \alpha_j \in E, \alpha_i \neq \alpha_j, \mu_{\alpha_i, \alpha_j} = \mu_{\alpha_j, \alpha_i}$$

Se consideran excluidas de la condición las relaciones de los elementos consigo mismos, que puedan ser intensas, reducidas e incluso nulas. Matricialmente, se expresa:

<sup>4</sup> Gil Aluja, 1999:46

$$\tilde{R} = \begin{array}{c|cccc} & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \hline \alpha_1 & \mu_{11} & \mu_{12} & \dots & \mu_{1n} \\ \alpha_2 & \mu_{21} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_n & 0 & 0 & \dots & 0 \end{array}$$

Figura 4

$$\mu_{ij} \in [0,1], (i, j = 1, 2, \dots, n)$$

$$\mu_{ij} = \mu_{ji}$$

**Transitividad:** Este tipo de relación requiere por lo menos la intervención de tres elementos del referencial E. Si se consideran tres elementos del referencial E;  $a_i, a_j, a_k$  de tal manera que  $a_i$  tiene relación con  $a_j$ ,  $a_j$  tiene relación con  $a_k$ ,  $a_i$  tendrá relación con  $a_k$ , entonces hay una relación de transitividad. Considerando esta relación de transitividad en la Lógica Borrosa, es conveniente suponer que la relación transitiva entre dos elementos  $a_i$  y  $a_k$  puede tener lugar a través de todos los elementos del referencial, es decir, a través de  $a_j, j= 1, 2, \dots, n$

Cada relación posee un grado en  $[0,1]$ ; el grado de relación entre  $a_i$  y  $a_j$  será  $\mu_{a_i, a_j}$  y el de la relación entre  $a_j$  y  $a_k$ , será  $\mu_{a_j, a_k}$ , y así sucesivamente. La regla para que esta relación de transitividad no sea obstruida es que « las relaciones indirectas entre  $a_j$  y  $a_k$  deben ser menores que las relaciones directas entre  $a_i$  y  $a_k$  », el nivel de relación directa viene dado por  $\mu_{a_i, a_k}$ <sup>5</sup>. Para asegurar que las valuaciones indirectas sean menores a las directas, se opta por el valor mínimo de cada relación,  $(\mu_{a_i, a_1} \wedge \mu_{a_1, a_k}), (\mu_{a_i, a_2} \wedge \mu_{a_2, a_k}), \dots, (\mu_{a_i, a_n} \wedge \mu_{a_n, a_k})$  en un proceso conocido como *composición maxmin*, siendo pues la expresión general que define la transitividad:

$$\forall a_i, a_j, a_k \in E : \mu_{a_i, a_k} \geq \bigvee_{a_j} (\mu_{a_i, a_j} \wedge \mu_{a_j, a_k}), j = 1, 2, \dots, n$$

Las relaciones entre los elementos de un referencial E; son de cuatro tipos y están determinadas por la presencia de algunas propiedades denominadas de preorden y orden, de semejanza o proximidad y de similitud o equivalencia. Para cada tipo es necesario que se cumplan las propiedades, como se presentan en la tabla siguiente<sup>6</sup>

<sup>5</sup> Gil Aluja, 1999:57

<sup>6</sup> Álvarez Vizcarra, Lógica borrosa, efectos olvidados y exposición al riesgo cambiario

a. Relaciones de preorden y orden	b. Relaciones de semejanza o proximidad	c. Las relaciones de similitud (Zadeh, 1971)
<i>Reflexividad</i>	<i>Reflexividad</i>	<i>Reflexividad</i>
$\forall \alpha_i \in E (i = 1, 2, \dots, n): \mu_{\alpha_i \alpha_i} = 1,$ si $i = j \mu_{\alpha_i \alpha_j} \in [0, 1],$ si $i \neq j$	$\forall \alpha_i \in E (i = 1, 2, \dots, n): \mu_{\alpha_i \alpha_i} = 1,$ si $i = j \mu_{\alpha_i \alpha_j} \in [0, 1],$ si $i \neq j$	$\forall \alpha_i \in E (i = 1, 2, \dots, n): \mu_{\alpha_i \alpha_i} = 1,$ si $i = j$ $\mu_{\alpha_i \alpha_j} \in [0, 1],$ si $i \neq j$
<i>Transitividad</i>	<i>Simetría</i>	<i>Transitividad</i>
$\forall \alpha_i \alpha_j \alpha_k \in E (j = 1, 2, \dots, n): \mu_{\alpha_i \alpha_k} \geq \nu_{\alpha_j} (\mu_{\alpha_i \alpha_j}, \mu_{\alpha_j \alpha_k})$	$\forall \alpha_i \alpha_j \in E (i = 1, 2, \dots, n):$ $\mu_{\alpha_i \alpha_j} = \mu_{\alpha_j \alpha_i}$ si $\alpha_i \neq \alpha_j$	$\forall \alpha_i \alpha_j \alpha_k \in E (i = 1, 2, \dots, n): \mu_{\alpha_i \alpha_k} \geq \nu_{\alpha_j} (\mu_{\alpha_i \alpha_j}, \mu_{\alpha_j \alpha_k})$
<i>Antisimetría</i>		<i>Simetría</i>
(*) <i>Antisimetría perfecta</i> $\forall \alpha_i \alpha_j \in E \alpha_i \neq \alpha_j, \mu_{ij} > 0 \Rightarrow \mu_{ji} = 0$ (*) <i>Antisimetría borrosa</i> $\forall \alpha_i \alpha_j \in E \alpha_i \neq \alpha_j, \mu_{ij} \neq \mu_{ji}$ o bien $\mu_{ij} = \mu_{ji} = 0$		$\forall \alpha_i \alpha_j \in E (i = 1, 2, \dots, n):$ $\mu_{\alpha_i \alpha_j} = \mu_{\alpha_j \alpha_i}$ si $\alpha_i \neq \alpha_j$

Figura 5

En el primer caso, al no exigirse la propiedad de simetría, es factible proponer que haya una relación entre  $a_i$  y  $a_j$ , y no entre  $a_j$  y  $a_i$ , lo cual permite cierto orden (Klimova, 2009). Dado que en una misma relación [R] pueden existir unos lazos con simetría y otros sin simetría, por lo cual, cuando no determinan ni similitud ni orden, son denominados de preorden<sup>7</sup>

Si a una relación de preorden se le impone la condición de antisimetría, aparece la relación de orden. De esta manera, se dice, hay una relación de orden cuando se cumplen las propiedades de reflexividad, transitividad y antisimetría. Una relación de similitud está formada por una o varias subrelaciones de similitud que forman las «clases de similitud». Cuando las subrelaciones de similitud son máximas, pero no disjuntas, se conocen con el nombre de subrelación máxima de similitud<sup>8</sup>

Uno de los elementos principales de la teoría borrosa son los **intervalos de confianza**, un intervalo de confianza recoge un rango de valores en el cual va a encontrarse el valor buscado o especificado de la variable en estudio.

Similares a los intervalos de confianza encontramos las **tripleas y cuádruplos de confianza**; La tripleta de confianza es un dato incierto definido por dos extremos que acotan el valor previsional de la variable bajo estudio y un valor central indicativo del valor con mayor posibilidad de obtenerse. Así mismo, el cuádruplo de confianza es un dato incierto definido por dos extremos que acotan el valor previsional de la variable bajo estudio y un valor central indicativo de mayor presunción, definido dentro de otro que acota el dato provisional entre un valor mínimo y uno máximo.

Continuando con las obras de Gil Aluja y Kaufmann (1993, 89 – 90), la definición de las escalas semánticas se antoja fundamental: [4]

<sup>7</sup> Gil Aluja, 1999.

<sup>8</sup> Klimova, 2009.

**Escala binaria**

0: falso  
1: verdadero

**Escala pentaria**

0: falso  
0.25: más falso que verdadero  
0.5: ni verdadero ni falso  
0.75: más verdadero que falso  
1: verdadero

**Escala endecadaria**

0: falso  
0.1: prácticamente falso  
0.2: casi falso  
0.3: bastante falso

**Escala ternaria**

0: falso  
0.5: ni verdadero ni falso  
1: verdadero

**Escala escaria**

0: falso  
0.2: casi falso  
0.4: más falso que verdadero  
0.6: más verdadero que falso  
0.8: casi verdadero  
1: verdadero

## 5.2 Subconjuntos borrosos

Si  $\Omega$  es un conjunto, un subconjunto borroso de  $\Omega$  está definido por su función de pertenencia, escrita  $(x)$ , la cual produce valores en  $[0,1]$  para todo  $x$  en  $\Omega$ . Luego,  $(x)$  es una función que mapea  $\Omega$  en  $[0,1]$ . Si  $(x_0)=1$ , se dice que  $x_0$  pertenece a  $\Omega$ , si  $(x_1) = 0$ , se dice que  $x_1$  no pertenece a  $\Omega$ , y si  $(x_2)= 0.3$  se dice que el valor de pertenencia de  $x_2$  en  $\Omega$  es 0,3. Cuando  $(x)$  es igual a uno o a cero se obtiene un subconjunto ordinario (no borroso) de  $\Omega$ . La función de pertenencia  $(x_i)$  también puede representarse por  $\mu_i$ .

Es muy importante determinar que los subconjuntos borrosos se asocian a distribuciones de posibilidad, las cuales son diferentes de las distribuciones de probabilidad.

Definidos los subconjuntos borrosos, podemos introducir los **subconjuntos borrosos continuos**; En el ámbito continuo, la variable puede tomar cualquier valor comprendido en el intervalo de referencia y todo valor tiene asociado un nivel de presunción (posibilidad) determinado. [2]

Se observa que un alto grado de posibilidad no implica un alto grado de probabilidad. Es decir, lo que es probable debe ser posible, en cambio lo que es posible no siempre es probable. La teoría de probabilidades y la teoría de posibilidades (subconjuntos borrosos) no son sustitutivas entre sí, pero son complementarias.

Ahora, desde el punto de vista de la lógica borrosa, el carácter borroso del conjunto  $A$  se cumple si y sólo si se incumplen las leyes de Aristóteles:

$$A \cap A^c \neq \emptyset \text{ y } A \sqcup A^c \neq X$$



Es decir, lo probabilístico excluye lo borroso, pero lo borroso incluye lo probabilístico porque se trata de un caso especial (Buckley, 2002, 5– 17).

### **5.2.1 Subconjuntos Borrosos Discretos**

En el campo discontinuo, los valores posibles de la variable objeto de estudio son datos concretos distanciados entre sí. [11]

### **5.2.2 Subconjuntos Borrosos Continuos**

A diferencia de lo mostrado en el apartado anterior, en el escenario continuo, la variable puede tomar cualquier valor comprendido en el intervalo de referencia, valor al cual se le asigna un nivel de presunción determinado. [11]

## **5.3 Función de Pertenencia**

La función de pertenencia determina el grado de pertenencia del elemento al conjunto. La pertenencia de elementos a un conjunto difuso, no es cuestión de todo o nada, sino que son diferentes grados de pertenencia dentro del intervalo  $[0,1]$ . Las funciones comúnmente utilizadas para describir el grado de pertenencia de un elemento al conjunto borroso son: Gamma ( $\Gamma$ ) o S, Lambda o triangular y L o Z.

## **5.4 Números Borrosos**

Un número borroso se define como el subconjunto borroso que satisface tres propiedades:

- La variable objeto de estudio toma valores en el campo de los números reales.
- La función de pertenencia es normal. La normalidad significa que por lo menos un valor de la variable estudiada tiene asociada el máximo nivel de presunción.
- La función de pertenencia es convexa. Esto significa que desde el extremo inferior al valor de mayor posibilidad los niveles de presunción crecen o se mantienen, pero en ningún momento disminuyen. Del mismo modo, desde el valor de máxima posibilidad hasta el extremo superior, los niveles de presunción disminuyen o se mantienen, pero nunca aumentan.

De esta forma, un número borroso asocia dos conceptos: El de intervalo de confianza que se halla ligado a la noción de incertidumbre y el de nivel de presunción ligado a la percepción del individuo, es decir, a la noción de valuación. Por analogía, el concepto de variable aleatoria está ligado a la noción de aleatoriedad y por ley de probabilidad a la concepción de medida. [9] y [12]

Es posible definir un número borroso sobre cualquier conjunto totalmente ordenado. Por ejemplo: los números reales o los números enteros.<sup>9</sup>

#### 5.4.1 Números Borrosos Triangulares

Formalmente se denomina número borroso triangular al número borroso real, continuo, cuya función de pertenencia es lineal a derecha y a izquierda, y determina con el eje horizontal un triángulo. Los números borrosos triangulares (NBT) se utilizan fundamentalmente cuando sobre una determinada magnitud se conocen únicamente tres valores: el mínimo, el máximo y el valor de mayor nivel de presunción. Mediante estos valores representamos el número borroso triangular, como [12]:

$$\tilde{A} = (a_1, a_2, a_3) \text{ donde } a_1 \leq a_2 \leq a_3$$

Dados dos números reales  $a_1$  y  $a_2$  tales que  $a_1 \leq a_2$ , se llama intervalo cerrado de extremos  $a_1$  y  $a_2$  a  $A = [a_1, a_2]$  al conjunto de números reales que satisfacen  $A = \{x/x \in \mathbb{R} \wedge a_1 \leq x \leq a_2\}$ . Los intervalos de confianza son un proceso lógico práctico para tratar la incertidumbre. Si consideramos una situación donde “x” es un dato incierto, y lo único conocido con cierto grado de confianza es que “x” es mayor o igual a un número  $a_1$  y menor o igual a otro número  $a_2$ . En otros términos sabemos que  $x \in [a_1, a_2]$  con un nivel de confianza para  $\alpha \in [0, 1]$ <sup>10</sup>

### 6. Exposición de los datos.

Los mercados de valores son sistemas complejos dentro de los cuales un gran número de participantes convergen, interactuando entre sí con el objetivo de maximizar sus ganancias mediante la negociación de derivados y valores. El mercado de acciones del primer mundo fue creado en 1460 en Amberes, Bélgica, pero no fue hasta principios del siglo XVII cuando el primer mercado de valores importante se estableció en Ámsterdam, Países Bajos, gracias a la *Dutch East India Company*, para hacer tratos con acciones y bonos.

Hoy, entre los mercados bursátiles más importantes del mundo, podemos encontrar: NYSE (Bolsa de Nueva York), **NASDAQ** (Asociación Nacional de Comerciantes de Valores Citas automatizadas), TSE (Bolsa de valores de Toronto), NYSE-Euronext, también como los mercados bursátiles de París, Ámsterdam, Londres, Madrid, Milán y Frankfurt.

A pesar de que el principal axioma para los mercados bursátiles es realmente simple: comprar bajo y vender alto; la decisión de cuándo y cuánto comprar o vender no es tan simple. Para superar esta dificultad, han surgido un conjunto de técnicas como puede ser el análisis especulativo, el cual puede considerarse como el estudio de información de mercado con el objetivo de predecir tendencias de ascenso y caída. Un caso particular de este análisis es el análisis técnico, que es uno de los mecanismos más ampliamente utilizados en la decisión haciendo soporte debido a su efectividad y relativa simplicidad.

<sup>9</sup> Véase en el anexo 1, una descripción más técnica de explicados elementos

<sup>10</sup> La presente sección descriptiva ha sido extraída en gran parte de [2]

Modeling Chaotic Behavior of Stock Indices Using Intelligent Paradigms [6]

Este es el nombre del paper que se ha usado como base de la comparativa estática en el tratamiento y estudio de los datos. Este artículo investiga el desarrollo de una técnica confiable y eficiente para modelar lo aparentemente caótico que es el comportamiento de los mercados bursátiles. Considera una red neuronal artificial probada y testada usando el Algoritmo *Levenberg-Marquardt*, *Support Vector Machine* (SVM), el *Takagi-Sugeno* neuro-modelo borroso y una red neuronal que impulsa la diferencia (DBNN).

El índice Nasdaq-100 refleja las compañías más grandes del Nasdaq en los principales grupos industriales, incluyendo *hardware* y *software*, telecomunicaciones, comercio minorista/mayorista y biotecnología. El índice Nasdaq-100 es un índice con ponderación de capitalización modificada, que está diseñado para limitar que el índice sea dominado por unas pocas acciones y mientras, mantener el ranking de capitalización de las empresas. A través de una inversión en acciones de Nasdaq-100, los inversores pueden participar en el rendimiento colectivo de muchas de los componentes de Nasdaq.

Por lo que a este trabajo concierne, vamos a centrarnos en la aplicación de la red *neuro-fuzzy*, con un alto componente de aplicación metodológica similar a la mostrada y descrita en las primeras secciones del estudio. Tal y como se ha contado, los sistemas de inferencia borrosa son excelentes para la toma de decisiones bajo incertidumbre. La computación neuro-difusa es un marco popular en el que en la formación de redes neuronales, los algoritmos se utilizan para ajustar los parámetros de los sistemas de inferencia difusa. El paper analiza el valor del índice Nasdaq-100 entre el 11 de enero de 1995 y el 11 de enero de 2002, para el estudio se ha tomado una muestra contenida dentro del periodo, desde enero de 1999 a enero de 2002, para estudiar el comportamiento de la serie mediante la metodología *Box-Jenkins* y poder así establecer la comparativa de si la modelización difusa es capaz de estudiar o explicar, mejor o peor, el comportamiento de la serie histórica que un modelo de empirismo contrastado como el mencionado. Esta es, realmente, la aportación del presente estudio.

La computación Neuro Fuzzy (NF) es un marco popular para resolver problemas complejos. Si tenemos conocimiento expresado en reglas lingüísticas, podemos construir una inferencia difusa (Sistema (FIS)), y si tenemos datos, o podemos aprender de una simulación (entrenamiento), entonces puede usar ANNs. Para construir un FIS, tenemos que especificar los conjuntos difusos, los operadores difusos y la base de conocimiento. Del mismo modo, para construir una ANN para una aplicación, el usuario necesita para especificar la arquitectura y el algoritmo de aprendizaje. Un análisis revela que los inconvenientes pertenecientes a estos enfoques parecen complementarios y, por lo tanto, es natural considerar construir un sistema integrado que combine ambos conceptos. Si bien la capacidad de aprendizaje es ventajosa desde el punto de vista de FIS, la formación de la base de la regla lingüística será ventajosa desde el punto de vista de ANN.

– ¿Realmente, qué es un Sistema de Inferencia Neuronal Adaptativo?–

En el *paper* expuesto se encuentra una descripción exhaustiva del funcionamiento de la red, en este estudio se considere muy adecuado realizar una explicación funcional basándose en la génesis de dicha metodología, debido a que nos encontramos con una implementación poco común.

ANFIS es una red adaptativa que representa la implementación del modelo de inferencia difuso *Takagi-Sugeno*. La filosofía de ANFIS consiste en predecir el valor en el tiempo  $x=t+P$  de la serie desde el punto  $x=t$ , a partir de unas entradas que resultan del mapeo de puntos en la serie apartados o rezagados en un espacio de tiempo  $\Delta$ , empleando una red neuronal construida con una arquitectura determinada arbitrariamente. Los resultados de estas predicciones se pueden comparar con otros modelos de redes neuronales y/o otras metodologías estadísticas regresivas como ARMA o ARIMA, que es exactamente lo que se propone en el presente proyecto para la serie escogida. [15]

Las redes adaptativas corresponden a aquellas estructuras de varias capas compuestas de un conjunto de nodos conectados a través de enlaces, donde cada nodo es una unidad de procesamiento que desempeña una función sobre la señal recibida para generar otra de salida. Cada enlace indica la dirección del flujo de la señal desde un nodo a otro, estos son adaptativos (con función parametrizada) cuando además de la señal o señales recibidas, dependen de un conjunto de parámetros o valores asociados para generar la salida, o nodos no-adaptativos (sencillos) cuando solo necesitan las entradas. Esta estructura contiene las ventajas siguientes:

- ✓ Se utilizan algunas metodologías y procedimientos típicos de las redes neuronales para el ajuste de sus parámetros, es decir, se configuran los parámetros de las funciones de pertenencia y consecuencia mediante estrategias de entrenamiento.
- ✓ Se pueden construir modelos a partir de pares de datos entrada/salida, reduciendo considerablemente el tiempo de modelamiento y los requerimientos de expertos.

A diferencia de las redes neuronales artificiales, la red adaptativa no tiene pesos sinápticos y presenta la ventaja de tomar el esquema y arquitectura de un Sistema de Inferencia Difusa para poder aplicar metodologías conocidas de entrenamiento y ajuste de sus parámetros.

ANFIS es considerada una red adaptativa y, como se mencionó, es la adaptación del sistema de inferencia difuso *Takagi-Sugeno*. En nuestro caso, y en la figura 6 a continuación expuesta, tratamos con una arquitectura ANFIS de 2 entradas ( $x,y$ ) cada una de las cuales con dos funciones de pertenencia, formando en total, ocho reglas difusas,  $(A_1, \dots, D_2)$ . En el caso de dos entradas, como en el que se está tratando, ANFIS tiene definidas la estructura siguiente de funcionamiento en cada capa.

**Capa 1.** Cada nodo en esta capa tiene una función de nodo.  $O_i^1 = \mu_{A_i}(x)$ , para  $i=1, 2$  o  $O_i^1 = \mu_{B_i}(y)$ , para  $i=3,4,\dots$ .  $O_i^1$  es el grado de membresía de un conjunto difuso  $A = (A_1, A_2, B_1$  o  $B_2)$  y especifica el grado en que la entrada, dada  $x$  (o  $y$ ) satisface el cuantificador  $A$ . Por lo general, la función de nodo puede ser cualquier función parametrizada, por ejemplo, una membresía gaussiana. La función está especificada por dos parámetros  $c$  (centro de función de membresía) y  $\sigma$  (ancho de la función de membresía gaussiana  $(x, c, \sigma) = e^{-\frac{1}{2}(\frac{x-c}{\sigma})^2}$ )

**Capa 2.** Cada nodo en esta capa multiplica las señales entrantes y envía el producto fuera. Cada salida de nodo representa la potencia de disparo de una regla.  $O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y)$ ,  $i = 1,2,\dots$ , En general, cualquier operador de norma T que realice fuzzy "Y" se puede utilizar como la función de nodo en esta capa. Los nodos en esta capa son no-adaptivos. En esta capa se generan los pesos de disparo. Es posible calcularlos como la relación entre las señales de entrada a este nodo tal y como se define  $O_i^2$

**Capa 3.** Los parámetros *consecuencia* de la regla se determinan en esta capa  $O_i^3 = f_i = xp_i + yq_i + r_i$ , donde  $\{p_i, q_i, r_i\}$  son los parámetros llamados consecuencia de la aplicación de la regla del sistema de inferencia fuzzy de Sugeno, es decir, la combinación lineal de entrada. Si  $x=A_i$  &  $y=B_i$ , entonces  $f_i = xp_i + yq_i + r_i$ , tal y como se ha definido.

También obtenemos  $\bar{w}_i$ , que son los pesos normalizados, a partir del *output* de la capa 2.

**Capa 4.** Cada nodo  $i$  en esta capa tiene una función adaptativa, cuya salida es el producto del nivel de disparo normalizado y la combinación lineal de las entradas, obteniendo:

$$O_i^4 = \sum \bar{w}_i f_i = \sum ((p_i x + q_i y + r_i))$$

**Capa 5.** Cada nodo en esta capa agrega todas las fuerzas de disparo de las capas anteriores resultando lo siguiente:

$$O_i^5 = \sum_i \bar{w}_i$$

y por último, en la última capa se calcula la salida total del sistema.

**Capa 6.** Cada  $i$ -ésimo nodo en esta capa calcula los outputs individuales resultantes.

$$O_i^6 = \frac{\sum \bar{w}_i f_i}{\bar{w}_i}, i = 1,2,\dots$$

La Figura de la página siguiente representa la arquitectura de 6 capas de ANFIS de salida múltiple y cuya funcionalidad se acaba de detallar en las líneas anteriores.

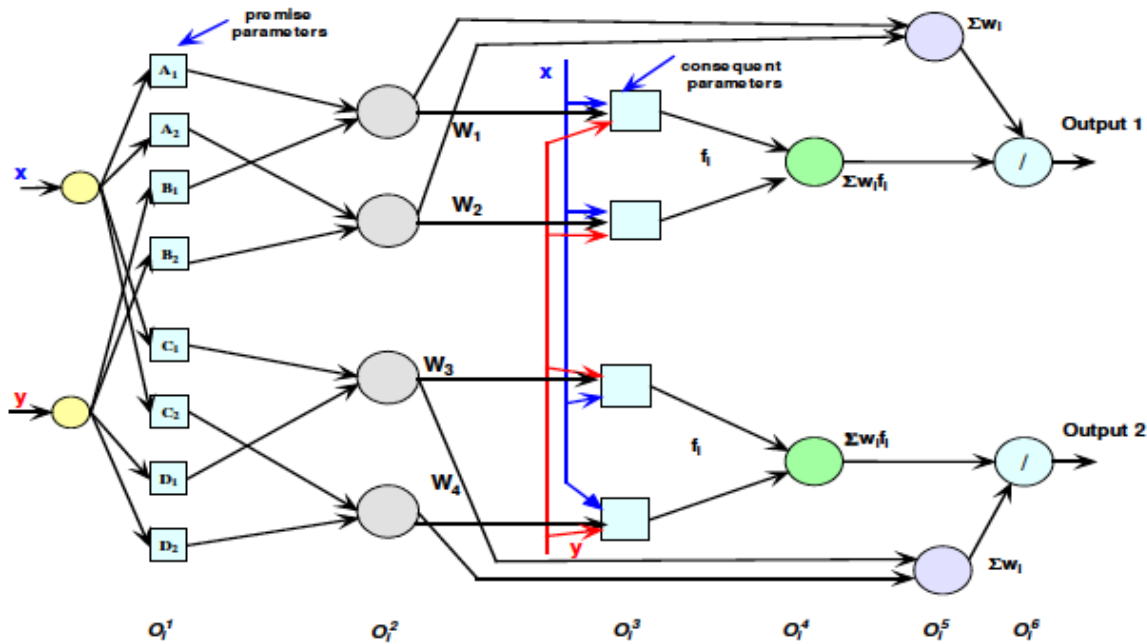


Figura 6

Pues bien, como se ha explicado, la metodología ANFIS involucra un análisis previo de la serie, que consiste en adaptar el modelo de capas, ajustar los parámetros de entrenamiento y la configuración de las entradas para entrenamiento de la red neuro-difusa, con el fin de generar modelos en los que se puede confiar desde el punto de vista del ajuste del modelo. En la predicción del mercado de acciones las formas determinísticas de las series de tiempo se ven alteradas por una componente aleatoria y no muestra valores atípicos ni ninguna componente estacional, es decir que no presenta patrones que se repitan [15]

ANFIS utiliza una regla de aprendizaje híbrida con una combinación del método gradiente descendiente y el estimador de mínimos cuadrados (LSE). Suponiendo una única salida ANFIS representada por:  $output = F(I, S)$  [16]

Donde  $I$  es el conjunto de variables de entrada y  $S$  es el conjunto de parámetros y  $F$  la función que modela la red adaptable. Si existe una función  $H$  tal que la función compuesta  $H \circ F$  es lineal en algunos de los elementos de  $S$ , resulta que estos elementos se pueden identificar por el método LSE. En otras palabras, se puede descomponer en dos componentes  $S = S_1 \oplus S_2$ , donde  $\oplus$  representa una suma directa, tal que  $H \circ F$  es lineal en algunos de los elementos de  $S_2$ , entonces al aplicar  $H$  a la ecuación anterior se obtiene:

$$H(Output) = H \circ F(I, S),$$

que es lineal en los elementos de  $S_2$ . Si se tienen los valores de los elementos de  $S_1$ , se pueden insertar  $P$  datos de entrenamiento en la ecuación que se acaba de escribir y obtener una ecuación de matriz de la forma:

$$U\theta = v \quad (6.1)$$

donde  $\theta$  es un vector desconocido cuyos elementos son parámetros en  $S_2$ . Sea  $|S_2|=M$ , entonces las dimensiones de  $U$ ,  $\theta$  y  $v$  son  $P \times M$ ,  $M \times 1$  y  $P \times 1$ , respectivamente. Puesto que  $P$  (número de pares de datos de entrenamiento) es normalmente mayor que  $M$  (número de parámetros lineales), estamos ante un problema sobre determinado y, generalmente, no hay una solución exacta para (6.1). Entonces, la mejor solución para  $\theta$ , es la que minimiza el error cuadrático  $\|U\theta - v\|^2$ , es el estimador de mínimos cuadrados  $\theta^*$ , definido por:

$$\theta^* = (U^T U)^{-1} U^T v \quad (6.2)$$

donde  $U^T$  es la traspuesta de  $U$ , y  $(U^T U)^{-1} U^T$  es la pseudo inversa de  $U$  siempre que  $U^T U$  no sea singular. Con esto escrito, podemos definir por ultimo dos expresiones que nos van a ayudar a entender el proceso iterativo que realiza la red, que es lo que al lector le va a ayudar a ver, de dónde proviene el error que usamos como medida comparativa.

Sea  $u^T$  el  $i$ -ésimo vector fila de la matriz  $U$  definido en (6.1); y  $v^T$ , el  $i$ -ésimo elemento de  $v$ , entonces se puede calcular iterativamente utilizando las fórmulas siguientes:

$$\begin{aligned} \theta_{i+1} &= \theta_i + C_{i+1} + U_{i+1}(v^T_{i+1} - u^T_{i+1} \theta_i) \\ C_{i+1} &= C_i - \frac{C_i u_{i+1} u^T_{1-i} C_i}{1 + u^T_{1+i} C_i u_{i-1}}, i = 0, 1, \dots, P-1 \end{aligned} \quad (6.3)$$

En lo que concierne a este estudio, y considero de una vinculación severa a la aportación del trabajo, es la explicación de cómo a través de la aplicación de dichas capas y mecanismo adaptativo, se es capaz de valorar o medir la capacidad predictiva de la red, para una serie histórica de valores. Si bien es cierto, que las ecuaciones arriba expuestas son básicas para poder entender dicha funcionalidad, el lector puede encontrar en [16] una exhausta explicación de citada metodología.

Ahora bien, se pueden mezclar, como aquí, el método del gradiente con el de LSE para actualizar los parámetros en una red adaptable. Para poder aplicar el aprendizaje híbrido de forma iterativa, cada época de este procedimiento de aprendizaje se compone de un paso en avance y un paso en retroceso. Empezando con el primero, en el paso en avance, se proporcionan los datos de entrada y las señales funcionales se propagan hacia adelante para calcular cada salida de nodo, capa por capa, hasta obtener las matrices  $U$  y  $v$  de la ecuación. 6.1, y los parámetros en  $S_2$  se identifican mediante la ecuación 6.2 o las fórmulas que se detallan arriba de la ecuación. 6.3.

Tras esto, se calcula la medida del error para cada par de datos de entrenamiento. En el paso en retroceso, las señales de error (la derivada de la medida del error con respecto a cada salida

de nodo), se propagan desde la salida hasta la entrada de la red, y los parámetros de  $S_1$  son actualizados mediante el método gradiente descendiente. Si se leen las páginas anteriores con cautela, se puede observar cómo se hace referencia a que los valores (en este caso las series temporales) se ven alterados por una componente aleatoria que los hace comportarse diferente a lo largo del tiempo, por lo tanto, para que la fórmula recursiva de mínimos cuadrados considere las características variantes en el tiempo de los datos de entrada, los efectos de los pares de datos pasados deben decaer, conforme nuevos pares de datos se hacen presentes.

Una manera simple para ello es formular la medida del error cuadrático como una versión ponderada que asigne mayores factores de peso a los pares de datos más recientes. Esto último dicho se puede realizar si se añade un factor de olvido  $\lambda$  a la fórmula recursiva inicial originaria (6.3)

$$\theta_{i+1} = \theta_i + C_{i+1} + U_{i+1}(v_{i+1}^T - u_{i+1}^T \theta_i)$$

$$C_{i+1} = \frac{1}{\lambda} \left[ C_i - \frac{C_i u_{i+1} u_{i+1}^T C_i}{1 + u_{i+1}^T C_i u_{i+1}} \right]$$

A continuación se muestran los resultados [6], y se han remarcado los dos resultados que ejemplifican la evaluación del método predictivo *neuro-fuzzy* usado para el tratamiento de la serie temporal del Nasdaq-100.

	SVM	Neuro-Fuzzy	ANN-LM	DBNN
<b>Training results (RMSE)</b>				
Nasdaq-100	0.02612	0.02210	0.02920	0.02929
NIFTY	0.01734	0.01520	0.01435	0.0174
<b>Testing results (RMSE)</b>				
Nasdaq-100	0.01804	0.01830	0.02844	0.02864
NIFTY	0.01495	0.01271	0.01227	0.02252

Figura 7

Mostrando un *RMSE* de 0,022 para los resultados del entrenamiento y de 0,018 para los testados.



## 7. Análisis de la serie mediante Metodología *Box-Jenkins*.

### 7.1 Descripción teórica

La metodología *Box-Jenkins* pivota en tres puntos fundamentales que se clasifican del siguiente modo:

- Identificación del modelo (por medio de la FAS y FAP)
- Estimación de parámetros
- Diagnósis (significatividad de parámetros y FAS y FAP de la serie de residuos)

*Máxima Verosimilitud:*

La función de máximo verosimilitud estima aquellos parámetros que maximizan la función máxima verosimilitud y que viene dada por la siguiente expresión:

$$L(\theta) = L(\theta|x_1, \dots, x_n) = \prod_i f_{\theta}(x_i)$$

Siendo,  $\theta$  el parámetro a estimar,  $x_1, \dots, x_n$  la muestra de observaciones independientes y  $f_{\theta}(x_i)$ , la función de densidad.

Finalmente, una vez realizada la estimación de los parámetros, entramos en la fase de diagnóstico de la metodología *Box-Jenkins*, donde para ver si el modelo que hemos utilizado es válido, debemos mirar que los residuos del modelo no estén autocorrelacionados. De ser así, los errores se acumulan en el término de perturbación, lo que indica, que hemos estimado erróneamente el modelo o no hemos introducido todas las variables a explicar.

Se utilizarán los criterios de selección de AIC- Akaike Information Criterion y BIC- Bayesian Information Criterion para seleccionar el modelo que más se ajuste a nuestra serie.

*AIC- Akaike Information Criterion:*

Se trata de una medida de calidad de ajuste de un modelo respecto a los datos utilizados. Esta medida tiene en cuenta el número de parámetros utilizados en el modelo y se define de la siguiente forma:

$$AIC = -2 \log(\hat{\theta}) + 2k,$$

Siendo  $\theta$  el vector de parámetros,  $\log(\hat{\theta})$  el logaritmo de la función verosimilitud que maximiza  $\theta$  y  $k$  es el número de parámetros estimados.

**BIC- Bayesian Information Criterion:**

También es una medida de calidad de ajuste de un modelo para los datos utilizados. Esta medida, a diferencia del AIC, tiene en cuenta el tamaño de la muestra y se define de la siguiente forma:

$$BIC = -2 \log(\hat{\theta}) + k \log(n)$$

Siendo  $\theta$  el vector de parámetros,  $\log(\hat{\theta})$  el logaritmo de la función verosimilitud que maximiza  $\theta$ ,  $k$  es el número de parámetros estimados y  $n$  es el tamaño de la muestra.

**Modelos para la media condicional**

Se consideran tres modelos para cada serie. En primer lugar, un proceso autorregresivo (AR):

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

En segundo lugar, un modelo de media móvil (MA):

$$y_t = \mu + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \dots + \theta_q \mu_{t-q} + \varepsilon_t$$

Finalmente, el proceso puede ser una combinación entre un proceso autorregresivo y media móvil:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \mu_t + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \dots + \theta_q \mu_{t-q} + \varepsilon_t$$

Aun así, identificar delante de qué modelo estamos solo utilizando el correlograma de la FAS y la FAP es de suma dificultad, ya que entramos en una óptica subjetiva y poco empírica. Es por eso, que es necesario utilizar los criterios de información para seleccionar el modelo que más se ajuste a nuestra serie.

**Modelos para la varianza condicional**

*GARCH* ( $p, q$ ):

El modelo GARCH es la abreviatura de *Generalized Autorregresive Conditional Heterocedasticity* y da nombre a la ampliación del modelo ARCH ( $q$ ) propuesto por Engel (1982). El modelo GARCH ( $p, q$ ) se puede definir como:

$$\sigma_{it}^2 = \omega_i + q = 1 Q_i \alpha_{iq} \varepsilon_{i,t-q}^2 + p = 1 P_i \beta_{ip} \sigma_{i,t-p}^2$$

- $\varepsilon_t$  es un idénticamente distribuido con media cero y desviación típica 1.

- Los parámetros  $\omega > 0$  y  $\alpha_i, \beta_i \geq 0$  e  $i=1 \dots q$ , y  $j=1 \dots p$ . Además, para el cumplimiento de la condición de estacionariedad en media, la suma de los parámetros es menor a la unidad.
- La función de distribución marginal no es conocida, pero se pueden calcular los primeros momentos y definir el proceso respecto a su media y a su varianza.

El modelo GARCH, da pie a una gran variedad de modelos que parten de la base explicada anteriormente, como podrían ser los modelos GARCH en Media, el EGARCH, el GJR-GARCH, etc.

## 7.2 Análisis de la Serie Escogida.

Pues bien, arranquemos con la serie temporal de cotizaciones históricas del **Nasdaq-100**. Para importar los datos directamente de la página *Yahoo finanzas* se ha usado el paquete de R *quantmod*. [7]

### *Dibujo de la serie de precios.*

Es importante aclarar que el trabajo está desarrollado con los precios a cierre ajustados, una vez seccionados los datos, la serie temporal de precios a cierre del banco Nasdaq para el periodo temporal 1999-2002 es la siguiente:

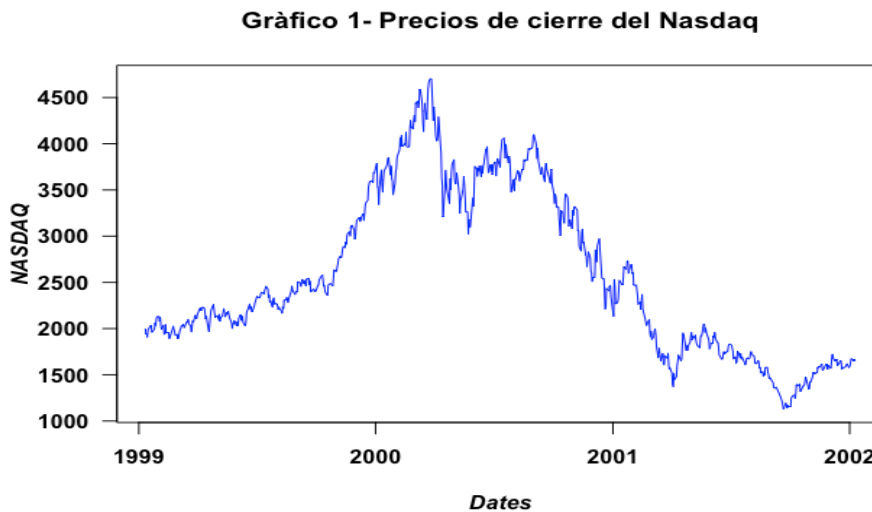


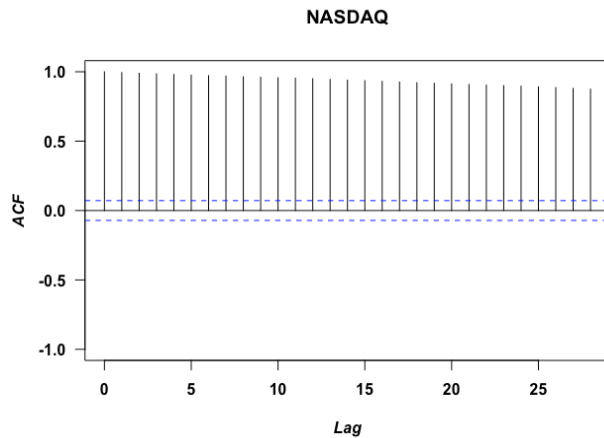
Figura 8

### 7.2.1 Análisis de la Estacionariedad de la Serie

¿Qué podemos decir de esta serie de precios?

Viendo el gráfico anterior, todo hace indicar que se trata de una serie no estacionaria ya que presenta una tendencia estocástica detectable y la media no es constante. La serie pasa temporadas bastante largas fuera de su media.

Si bien es cierto que tanto esto como los gráficos de autocorrelación total y parcial nos dejan pruebas informales sobre la estacionariedad o no de la serie, debemos tenerlas en cuenta siempre.



Aquí se puede observar la función de autocorrelación total de la serie. (FAS) Los coeficientes se encuentran muy cercanos a 1 y decrecen de manera muy lenta, por lo tanto, podemos intuir que se trata de una serie no estacionaria, tal y como hemos avanzado analizando el gráfico de la serie de precios en la página anterior.

Figura 9

### 7.2.2 Análisis de los Estadísticos Básicos de la Serie

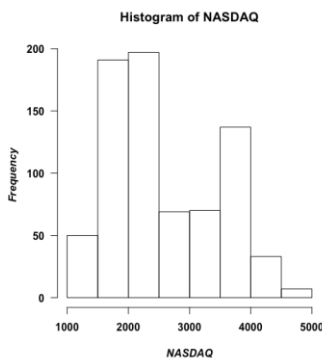
A continuación, previo al desarrollo de los contrastes en el procedimiento, se procede a mostrar los estadísticos básicos de la serie junto con la exposición de su histograma.

<i>NOMBRE</i>	<i>NAME</i>	<i>VALOR</i>
<i>Media de los datos</i>	<i>Mean</i>	<i>2.571108e+03</i>
<i>&amp;</i>		
<i>Desviación Estándar</i>	<i>Sd</i>	<i>880.5977</i>
<i>Valor mínimo</i>	<i>Min</i>	<i>1126.95</i>
<i>Valor máximo</i>	<i>Max</i>	<i>4704.73</i>
<i>Simetría de los datos</i>	<i>Skewness</i>	<i>0.4738885</i>
<i>Curtosis</i>	<i>Kurtosis</i>	<i>-0.9767359</i>
<i>Contraste Normal</i>	<i>Contraste Jarque Bera</i>	<i>P value = 2.573e-13</i>

Figura 10

Para ello, se desarrolla el contraste de Normalidad JB, rechazamos la hipótesis nula de comportamiento de la serie como una normal. Con un p valor < 0,05.

El histograma nos muestra la distribución de frecuencias de los resultados arriba expuestos.



Para confirmar que la serie de precios tiene autocorrelación realizamos el contraste de **Ljung Box** de 1 a 35 retardos. En absolutamente todos los retardos testados, la p valor es inferior a 0,05 ( $2,2e-16$ ), por lo tanto rechazo la hipótesis nula de No autocorrelación por lo que se afirma que la serie de precios presenta autocorrelación.

Los precios presentes se expresan en función de los pasados.

Figura 11

Ahora, para corroborar de manera **formal** lo que las primeras páginas nos han revelado de manera informal mediante el estudio de los gráficos de la serie de precios, se procede a contrastar la estacionariedad de la serie de precios. Se hace mediante los test de raíces unitarias *Dickey Fuller Aumentado* (ADF) y *Phillips Perron*, cuya hipótesis nula expresa que los precios son no estacionarios, al menos un  $I(1)$  contra una alternativa que expresa el opuesto, que los precios se comportan de manera estacionaria  $I(0)$ . Se ha hecho primero, sin constante y posteriormente, con ella. Los resultados han sido los mismos: output de mis resultados en RStudio, de elaboración propia.

```
Residual standard error: 85.11 on 742 degrees of freedom
Multiple R-squared: 0.01303, Adjusted R-squared: 0.005053
F-statistic: 1.633 on 6 and 742 DF, p-value: 0.1351
```

```
Value of test-statistic is: -0.479
```

```
Critical values for test statistics:
```

```
1pct 5pct 10pct
tau1 -2.58 -1.95 -1.62
```

revelaban, no podemos hablar de una serie de precios estacionaria, estos serán al menos un  $I(1)$ . Si añadimos una constante, el resultado es el mismo.

En el test ADF sin constante, vemos como el valor del estadístico queda a la derecha del 10pct, y la p valor  $> 0,05$ . Con todo, podemos NO rechazar la  $H_0$  de comportamiento no estacionario. Abogando a lo que los gráficos nos

Se procede pues, a diferenciar la serie una vez, obtener los rendimientos y trabajar con ellos en caso de conseguir estacionaridad tras los test de nuevo. Para hacerlo se trabaja con la diferencia de los logaritmos eliminando la primera observación.

```
Residual standard error: 0.03119 on 748 degrees of freedom
Multiple R-squared: 0.5272, Adjusted R-squared: 0.5259
F-statistic: 417 on 2 and 748 DF, p-value: < 2.2e-16
```

```
Value of test-statistic is: -21.8487 238.6827
```

```
Critical values for test statistics:
```

```
1pct 5pct 10pct
tau2 -3.43 -2.86 -2.57
phi1 6.43 4.59 3.78
```

comportamiento No Estacionario, de manera que una vez visto que los precios no son  $I(2)$ , se confirma que la serie de rendimientos es estacionaria.

El mismo procedimiento que se ha hecho para discriminar la serie de precios y rendimientos, a través de los test de *Dickey Fuller Aumentado*, se puede repetir para contrastar resultados con el test de *Phillips Perron*. Las hipótesis son las mismas, el resultado en la serie de precios son los mismos, no rechazamos la hipótesis nula de no estacionariedad. En cambio, el test de los rendimientos rechaza claramente la  $H_0$ , mostrando un comportamiento estacionario de la serie de rendimientos, independientemente del número de periodos (largo o corto). La única diferencia que merece reseña en estas líneas es que el test ADF introduce retardos de la variable dependiente mientras que el de *Phillips Perron* realiza una corrección paramétrica.

Esta vez el output de la izquierda es el mismo contraste ADF, pero con constante (el resultado es el mismo sin) y sobre la serie de precios ya diferenciada, de los rendimientos.

Podemos observar claramente como esta vez sí se rechaza la hipótesis nula de

Ahora bien, los dos contrastes arriba expuestos tienen una baja potencia, de modo que rechazamos la hipótesis nula más veces de lo que debiéramos. La solución a este problema es realizar el procedimiento del contraste mediante el *KPSS Test*. Se ha hecho siguiendo la misma metodología que antes, primero un contraste de la serie de precios, el cual ha rechazado la hipótesis nula. En este test, las hipótesis están intercambiadas respecto a los dos anteriores, en este caso, la  $H_0$  indica estacionariedad de los precios/rendimientos  $I(0)$ , y la alternativa el comportamiento no estacionario.

```
> NASDAQ.kpss<-ur.kpss(NASDAQ, type = c("tau"), lags = c("short"))
> summary(NASDAQ.kpss)

#####
# KPSS Unit Root Test #
#####

Test is of type: tau with 6 lags.

Value of test-statistic is: 2.3558

Critical value for a significance level of:
      10pct 5pct 2.5pct 1pct
critical values 0.119 0.146 0.176 0.216

> rendNASDAQ.kpss<-ur.kpss(diff(log(NASDAQ)), type = c("mu"), lags = c("short"))
> summary(rendNASDAQ.kpss)

#####
# KPSS Unit Root Test #
#####

Test is of type: mu with 6 lags.

Value of test-statistic is: 0.3317

Critical value for a significance level of:
      10pct 5pct 2.5pct 1pct
critical values 0.347 0.463 0.574 0.739
```

Tal y como intuíamos previamente, cuando se aplica el test a los rendimientos no rechazamos, dicha hipótesis nula. La serie de precios diferenciada una vez, da lugar a la serie de rendimientos, la cual es estacionaria.

### 7.3 Identificación, estimación y diagnosis de un Modelo para Media y Varianza

Llegados a este punto, nos encontramos con el primer conflicto del trabajo. Una vez contrastada la estacionariedad de los rendimientos mediante el test *KPSS*, si graficamos la correlación parcial y total de los rendimientos, vemos que sí lo están. Todo hace indicar que no se trata de un ruido blanco y tendremos modelo para la media. A continuación, se grafican la FAS y la FACP para los rendimientos de la serie.

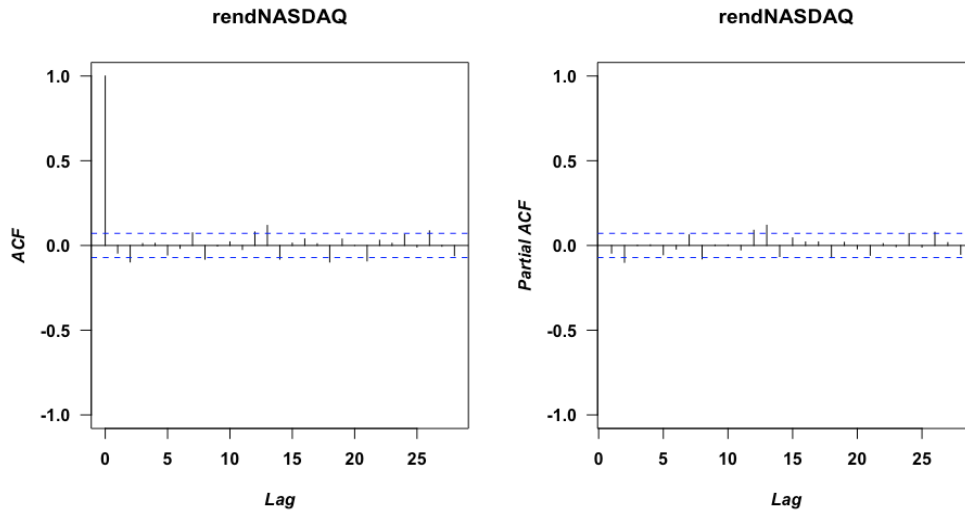


Figura 12

Se puede observar, claramente que existe autocorrelación en los rendimientos. Los gráficos nos vuelven a dar otra idea intuitiva, pero esta afirmación debemos confirmarla con los contrastes de autocorrelación, esta vez para los rendimientos.

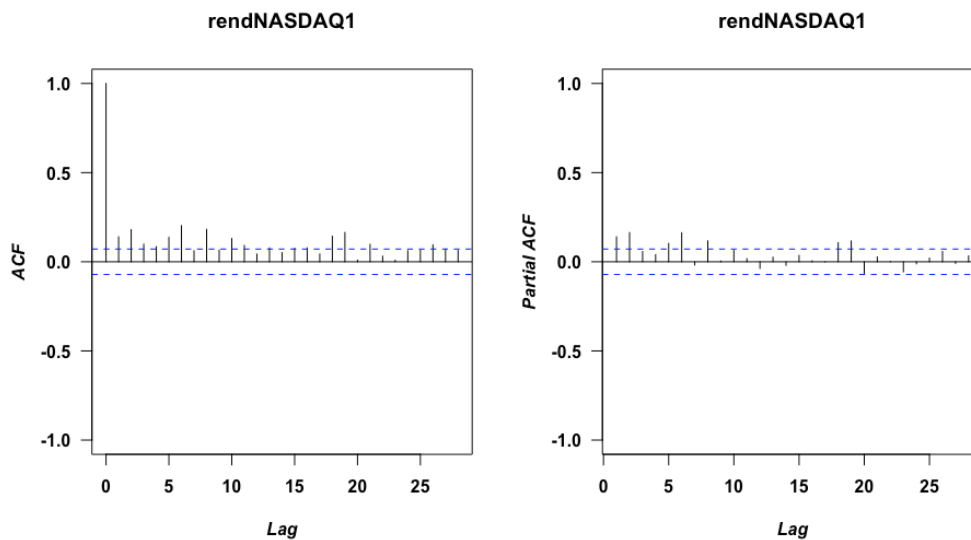


Figura 13

Como bien se ha señalado, podemos disponer de una serie temporal o bien de un ruido blanco, en todo caso debemos mirar la autocorrelación de los rendimientos al cuadrado para poder establecer una explicación de la varianza cambiante en el tiempo, es decir, debemos saber si podemos establecer un modelo para la varianza. Y esto es precisamente lo que indica el último gráfico. Gráfico en el cual sí observamos, a priori, autocorrelación de los rendimientos al cuadrado (rendNASDAQ1). Pevio a realizar los test de autocorrelación, podemos establecer que tendremos un modelo para la media y podremos modelizar la varianza cambiante a lo largo del tiempo.



Para asegurar la autocorrelación entre los rendimientos, usaremos el contraste *Ljung Box* desde 1 a 20 retardos. Este contraste propone la hipótesis nula de no autocorrelación contra la alternativa de existencia de ella. Los resultados son tajantes:

Lags	P value	Resultado
5	0.04573	Rechazo Ho
10	0.02031	Rechazo Ho
15	0.000197	Rechazo Ho
20	0.0001102	Rechazo Ho

Como se puede ver en la tabla adjunta arriba, rechazamos la hipótesis de no existencia de autocorrelación en los rendimientos, confirmamos pues, todos aquellos indicios que los gráficos nos han dado un poco atrás. Podemos modelar la media.

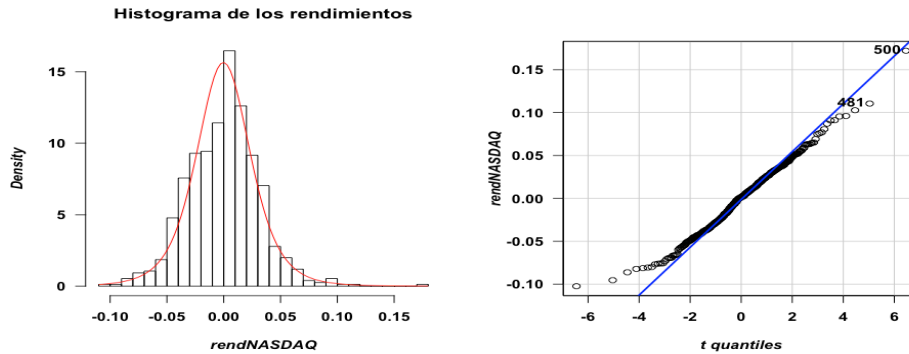
Si en cambio, trabajamos con los rendimientos al cuadrado, el test no varía: seguimos rechazando.

Lags	P value	Resultado
5	6.191e-13	Rechazo Ho
10	2,2e-16	Rechazo Ho
15	2,2e-16	Rechazo Ho
20	2,2e-16	Rechazo Ho

Figuras 14 y 15

En este caso, la autocorrelación entre los rendimientos al cuadrado denota la existencia de varianza cambiante a lo largo del tiempo, hecho que debemos modelizar a través de los modelos Garch, herramienta que nos permitirá ajustar un modelo de varianza y volatilidad que elimine la autocorrelación en los rendimientos.

Antes de profundizar en la estimación del modelo de la media y del Garch, optamos por estudiar a qué distribución se ajusta más nuestra serie de rendimientos. Para ello, superponemos al histograma de los rendimientos la trayectoria teórica perfecta de una distribución, primero Normal y luego T-Student. Consideramos que la serie de rendimientos Nasdaq se ajusta muchísimo mejor a una distribución T-Student.



Figuras 16 y 17

### 7.3.1 Modelo para la Media

Tal y como se ha visto en las páginas anteriores, se rechaza la no autocorrelación de los residuos, por lo tanto se debe estudiar un modelo para la media capaz de eliminar la correlación en dichos rendimientos. Para ello, hemos discernido entre las diferentes opciones AR, MA y ARMA hasta encontrar, que el modelo que más se nos adecua es el modelo ARMA (1, 1).

```
> #arma(1,1)
> model3 = arima(rendNASDAQ, order = c(1,0,1),include.mean = TRUE)
> model3

Call:
arima(x = rendNASDAQ, order = c(1, 0, 1), include.mean = TRUE)

Coefficients:
      ar1      ma1  intercept
  0.6093 -0.6805   -2e-04
s.e. 0.1758  0.1611    9e-04

sigma^2 estimated as 0.0009714: log likelihood = 1543.21, aic = -3078.43
> pnorm(c(abs(model3$coef)/sqrt(diag(model3$var.coef))), mean=0, sd=1, lower.tail=FALSE)
      ar1      ma1  intercept
2.637024e-04 1.204111e-05 3.966819e-01
>
> p1=c(1,-model3$coef[1:2]);p1
      ar1      ma1
1.0000000 -0.6093354  0.6805022
> roots=polyroot(p1)
> Mod(roots)
[1] 1.212231 1.212231
```

Cómo se concreta esta afirmación...

Como punto diferencial debemos poner los ojos en los coeficientes ar1 y ma1, **no rechazamos** el modelo vista su significación, cosa que nos ocurre con todos los demás modelos que hemos probado en el test. Todo esto son primeras aproximaciones a lo que intuimos que es el modelo adecuado, no perdamos de vista que nuestro objetivo principal es

eliminar la correlación entre los rendimientos, y esto lo debemos testar repitiendo esta vez, el test de autocorrelación ya visto.

Como se puede ver, no rechazamos la hipótesis de no autocorrelación, de los residuos, modelizados bajo un ARMA (1,1).

```
Box-Ljung test

data: model3$residuals
X-squared = 0.26177, df = 1, p-value = 0.6089

> Box.test(model3$residuals, lag = 2, type = c("Ljung-Box"))

Box-Ljung test

data: model3$residuals
X-squared = 2.4377, df = 2, p-value = 0.2956

> Box.test(model3$residuals, lag = 3, type = c("Ljung-Box"))

Box-Ljung test

data: model3$residuals
X-squared = 3.2707, df = 3, p-value = 0.3517

> Box.test(model3$residuals, lag = 4, type = c("Ljung-Box"))

Box-Ljung test

data: model3$residuals
X-squared = 3.7763, df = 4, p-value = 0.4371

> Box.test(model3$residuals, lag = 5, type = c("Ljung-Box"))

Box-Ljung test

data: model3$residuals
X-squared = 5.2451, df = 5, p-value = 0.3867

> Box.test(model3$residuals, lag = 10, type = c("Ljung-Box"))

Box-Ljung test

data: model3$residuals
X-squared = 13.883, df = 10, p-value = 0.1784

> Box.test(model3$residuals, lag = 15, type = c("Ljung-Box"))
```

Hemos conseguido explicar el comportamiento de la media de la serie mediante el modelo explicado. Para corroborar dicho hallazgo, volemós a representar la FAS y la FAP de los precios diferenciados (rendimientos) una vez modelizados. Tal y como podemos apreciar en el comportamiento de estos rendimientos, ya no presentan correlación.

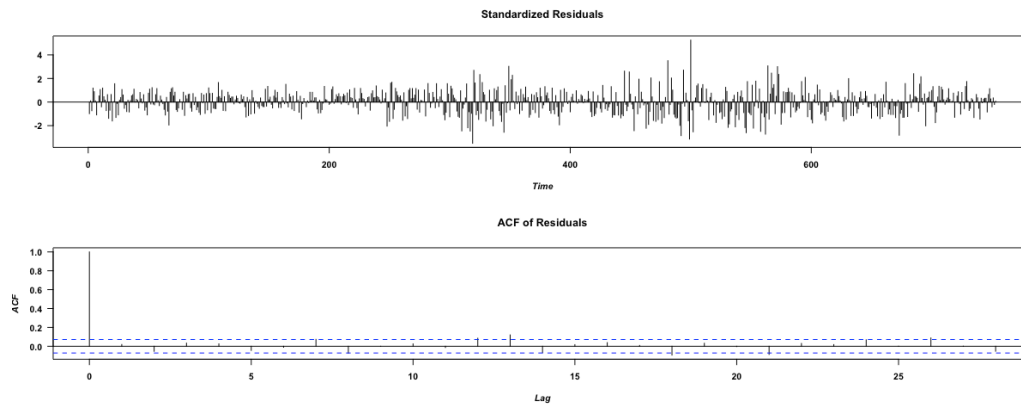


Figura 18

### 7.3.2 Predicción y Evaluación del Modelo

Para realizar la predicción, hemos usado un método recursivo desde la propia muestra. Usando las últimas 100 observaciones de la serie de rendimientos. Con ellas hemos sido capaces de estudiar la capacidad predictiva del modelo y así poder establecer una comparativa con los valores predichos del modelo que incorpora las variables *neuro-fuzzy*, al fin y al cabo es el objetivo de este estudio. Pues bien, los resultados son los siguientes.

```
> serie2<-ts(rendNASDAQ[1:653],)
> model = arima(serie2, order = c(1,0,1),include.mean = TRUE)
> model

Call:
arima(x = serie2, order = c(1, 0, 1), include.mean = TRUE)

Coefficients:
      ar1      ma1  intercept
  0.6414 -0.7243  -3e-04
s.e.  0.1455  0.1293   1e-03

sigma^2 estimated as 0.0009972: log likelihood = 1329.72, aic = -2651.45
> pnorm(c(abs(model$coef)/sqrt(diag(model$var.coef))), mean=0, sd=1, lower.tail=FALSE)
      ar1      ma1  intercept
5.205189e-06 1.058340e-08 3.769989e-01
> predicciones<-matrix(0,nrow=1,ncol=1)
> prediccion<-function(s)
+ {
+   rendNASDAQ<-ts(rendNASDAQ[1:(653+i)],)
+   model = arima (rendNASDAQ, order=c(1,0,1), include.mean=TRUE)
+   rendNASDAQ<-predict(model,n.ahead=1)
+   prediccion<-(rendNASDAQ$pred)
+ }
> for (i in 0:100)
+   predicciones<-rbind(predicciones,prediccion(i))
> predicciones<-predicciones[-1] #quitamos el 0 inicial.
> predicciones<-ts(predicciones, start=653, end=752) #convertimos las predicciones en una time-series para poder graficarlo junto con los valores reales.
> ts.plot(rendNASDAQ[653:752],predicciones,col=c('gray0','dodgerblue'), lwd=3, main='Predicción Recursiva vs. Observación Real')
> legend("bottomleft", c("Obs. Real","Predicción"), lty=c(1,1), col=c('gray0','dodgerblue'), lwd=3)
```

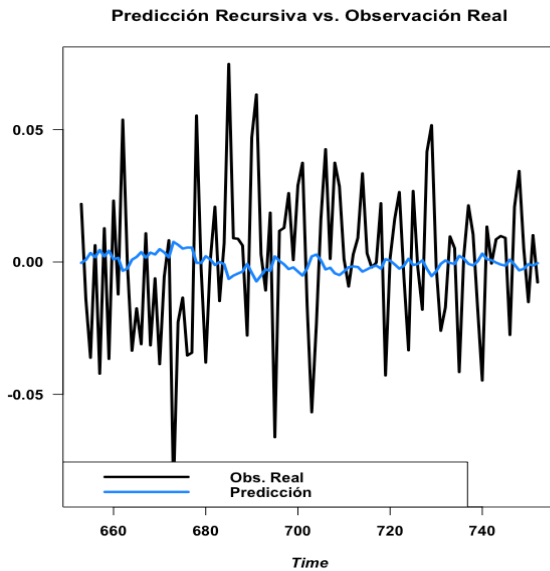


Figura 19

Tal y como se ve en la imagen superior, volvemos a estimar el modelo, predecimos los rendimientos time-series y lo ajustamos a los valores reales, para hacer la comparación recursiva. Podemos observar lo siguiente:

Vemos cómo podemos considerar la predicción a priori correcta, sobre todo porque intuimos la paridad en la volatilidad del comportamiento de los rendimientos, pero no llega a predecir la magnitud de la evolución del todo.

Con todo, pasamos a evaluar la predicción y el ajuste del modelo para la media a través de indicadores de errores. En especial, el *root mean square deviation*. El cual usaremos para comparar con el modelo de lógica difusa.

```
> eqm1<-sum(errors*errors)/100 #error cuadratico medio
> reqm1<-eqm^(.5) #root error cuadratico medio
> eam1<-sum(abs(errors))/100 #error absoluto medio
> epam1<-sum(abs(errors)/abs(rendNASDAQ[653:752]))/100 #error porcentual absoluto medio
> eqm1
[1] 0.0009210209
> reqm1
[1] 0.02813775
> eam1
[1] 0.02338734
> epam1
[1] 1.209673
```

Dónde se observa que el RMSE (reqm1) es muy similar al del paper estudiado (0,018)

### 7.3.3 Modelo para la Varianza y estudio de la Volatilidad de la serie

Informar al lector sobre cómo se ha escogido un modelo para la varianza requerirá de una introducción previa sobre qué modelos se han estudiado, para precisamente así, encontrar aquel que mejor cumpla los requisitos necesarios para explicar la varianza cambiante a lo largo del tiempo.

Se ha optado por un orden Garch (1,1). Tras esto, se han comparado los modelos simétrico “sGarch” y los asimétricos “eGarch” y “GJRGarch”, respetando el orden del modelo (1,1), el modelo Arma (1,1) y ajustado a una distribución T-Student. Usando el paquete rugarch.

Para tomar una decisión, se ha hecho hincapié en los siguientes *inputs*:

- Significación de los parámetros óptimos del modelo.
- Comparación en el criterio de información de Akaike. (AIC)
- Cumplimiento de No Rechazo de la Ho de No autocorrelación en los test: *Weighted Ljung-Box Test on Standardized Squared Residuals* y *Weighted Ljung-Box Test on Standardized Residuals*.

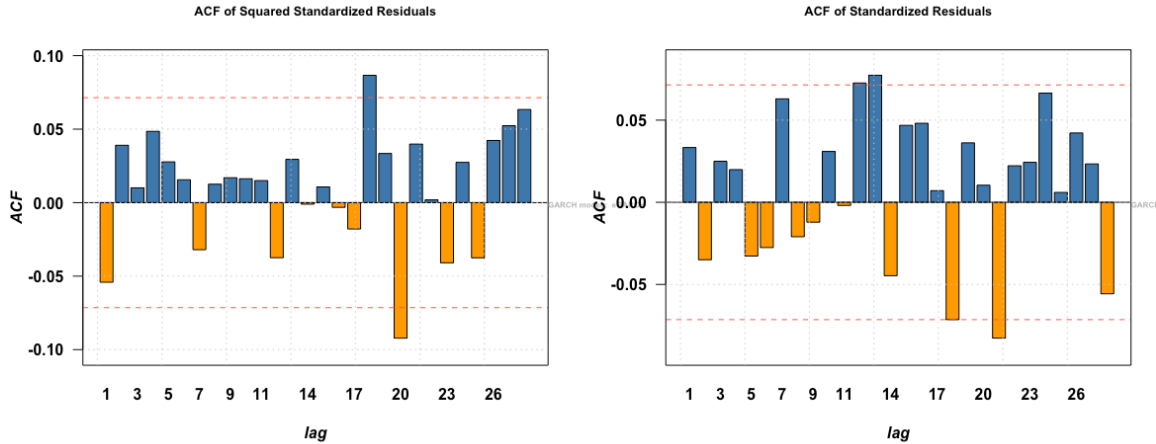
#### ● *MODELO eGARCH (1,1)*

<pre> *-----* *           GARCH Model Fit           * *-----* Conditional Variance Dynamics ----- GARCH Model      : eGARCH(1,1) Mean Model       : ARFIMA(1,0,1) Distribution      : std  Optimal Parameters -----       Estimate Std. Error  t value Pr(&gt; t ) mu      0.000212   0.000825   0.25717 0.79705 ar1     0.595779   0.069655   8.55328 0.00000 ma1    -0.662736   0.065010  -10.19433 0.00000 omega  -0.297112   0.047606   -6.24099 0.00000 alpha1 -0.095323   0.017078   -5.58176 0.00000 beta1   0.958444   0.006646  144.22344 0.00000 gamma1  0.128076   0.031563   4.05781 0.00005 shape  99.996800  92.998313   1.07525 0.28226 </pre>	<pre> Information Criteria ----- Akaike      -4.2361 Bayes       -4.1869 Shibata     -4.2363 Hannan-Quinn -4.2171  Weighted Ljung-Box Test on Standardized Residuals -----                                 statistic p-value Lag[1]                          0.8405  0.3592 Lag[2*(p+q)+(p+q)-1][5]        2.1483  0.9234 Lag[4*(p+q)+(p+q)-1][9]        4.0350  0.6828 d.o.f=2 H0 : No serial correlation  Weighted Ljung-Box Test on Standardized Squared Residuals -----                                 statistic p-value Lag[1]                          2.218  0.1364 Lag[2*(p+q)+(p+q)-1][5]        4.018  0.2518 Lag[4*(p+q)+(p+q)-1][9]        5.212  0.3989 d.o.f=2 </pre>
---	--

Con un Criterio Akaike: -4,2361

Este es el modelo que se ha escogido para modelizar la varianza, significación de todos los parámetros óptimos, como hemos indicado antes a excepción de la constante. Criterio de AIC más pequeño y cumplimiento del no rechazo en ambos test de contraste de autocorrelación ADF para los residuos.

Una vez hecho el análisis descriptivo de los parámetros, complementamos el estudio con el análisis gráfico, para ello, se grafica mediante RStudio, los dos test de autocorrelación de los residuos del modelo. Los resultados son los siguientes: Figura 20



Vemos que para ningún retardo, ni los residuos, ni los residuos al cuadrado sobrepasan los límites. Denota que no existe autocorrelación entre ellos, el modelo ha sido capaz de eliminarla. Los retardos que salen de las bandas de confianza en la función de autocorrelación total de los residuos al cuadrado se encuentran en el periodo 18 y 20 aproximadamente, muy lejos como para tenerlos en cuenta en la interpretación de los resultados.

Si complementamos el análisis con un estudio de la predicción de la volatilidad, podemos observar resultados reveladores. Si realizamos una estimación condicional, la haremos usando el conocido método llamado *Rolling*.

La única forma de analizar una previsión es compararla con los datos reales, es por ello que haremos un recorte de 100 observaciones al final de la serie y trabajaremos sobre ésta. Una vez hecho el recorte, estimaremos estas últimas diez observaciones recortadas y compararemos su resultado con el original.

Con el *Rolling Windows* iremos prediciendo periodo a periodo a través de las predicciones anteriores. Es decir, comenzaremos prediciendo “t+1” desde el momento “t”, “t+2” desde la predicción de “t+1” y, así sucesivamente hasta llegar a “t+n” observaciones (En este caso en concreto hablaremos de 100 observaciones). Al final del recorrido, habremos hecho 100 predicciones, es por ello que tendremos un error calculado para cada una de ellas menos los cinco últimos, en el que el mismo programa interrumpe su cálculo.

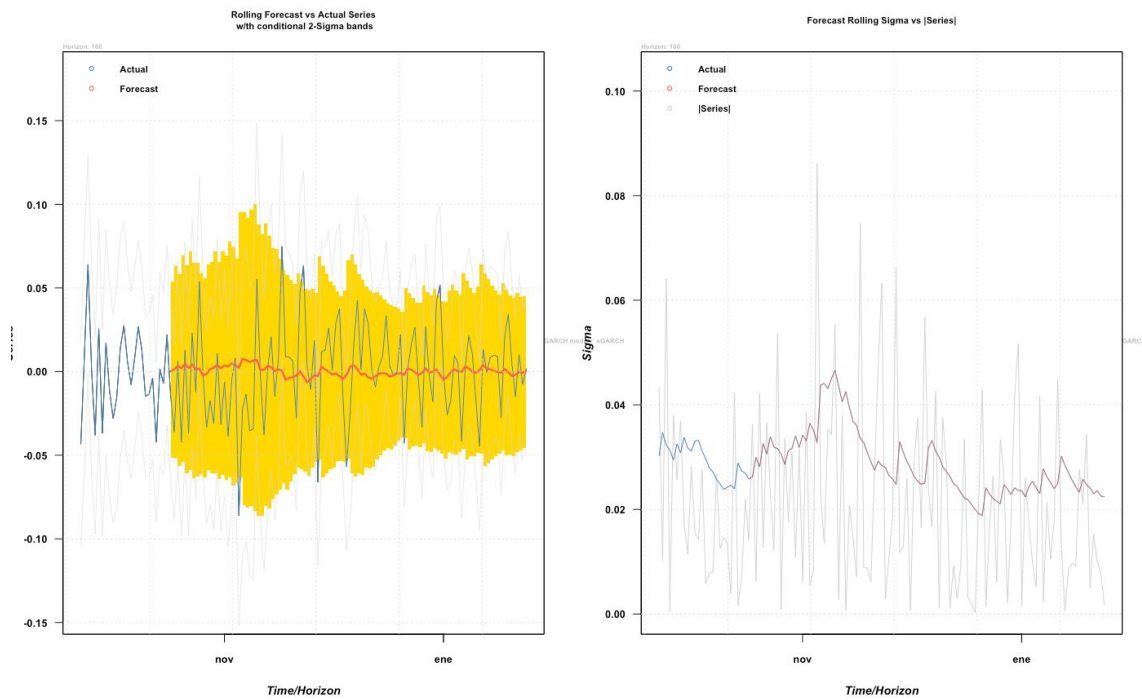


Figura 23

En la figura 23 podemos ver los resultados, la línea roja nos indica la tendencia de la predicción de los rendimientos con sus respectivos niveles de significación y lo amarillo el intervalo de confianza. Si lo observamos, vemos cómo se mezcla con los valores absolutos de los rendimientos con el objetivo de ver si existe una relación causal entre las variaciones de los rendimientos futuros y la volatilidad predicha. Volvemos a ver cómo la predicción es capaz de seguir una dinámica pareja a la realidad, captando los picos, pero no de la misma intensidad o magnitud, es interesante ver cómo se perfilan ambas siluetas (evolución serie y predicción)



## **8. Conclusiones.**

Con todo lo visto en las páginas del estudio, el lector debería cerciorarse de que finalmente se ha sido capaz de establecer una relación comparativa entre dos metodologías en la toma de decisión, aplicadas a una serie de datos históricos pertenecientes al índice Nasdaq – 100.

En primer lugar, la metodología *Neuro-Fuzzy* con bases establecidas en la Lógica Difusa, es capaz, una vez testado y comprobado, de predecir con un *RMSE* de 0,018. La misma serie, tratada mediante la metodología de modelos autorregresivos *Box-Jenkins*, la cual no incorpora ningún mecanismo, ni elemento borroso, lo hace con un error de 0,028.

La conclusión es clara a la vez que reveladora, se puede ver que los modelos de predicción basados en la neuro-difusa, concretamente ANFIS, la red adaptativa que representa la implementación del modelo de inferencia difuso *Takagi-Sugeno*, será capaz de establecer una mejor predicción futura del comportamiento de dicha serie, que la que es capaz de hacer la metodología *Box-Jenkins*.

Por otro lado, considero que esta comparativa estática debería probarse con muchas más series, ya que adjetivo dicho resultado como una afirmación inductiva. Con lo dicho, el trabajo cumple con las expectativas iniciales, en las que se intuía que la incorporación de relaciones borrosas al estudio de series temporales, ayudaría a la predicción de éstas, mejorándolas.

Complementariamente, la predicción en media de la serie mediante *Box-Jenkins* no capta del todo la magnitud del periodo (lo mismo ocurre con el estudio de su volatilidad) pese a realizar una predicción bastante adecuada a los hechos reales, además, consigue eliminar por completo la autocorrelación en los residuos, tras el hallazgo de un modelo adecuado para media y varianza.

Teniendo en cuenta que la función de pertenencia en ANFIS no puede cambiar durante el entrenamiento (adecuándose así a la problemática cambiante) y que la predicción mediante *Box-Jenkins* (pese a considerar el efecto del impacto de las noticias) no capta la magnitud total de la volatilidad de la serie, considero de mejor aplicación de dichos modelos a periodos pertenecientes al corto o medio plazo.

Para terminar, considero de total interés futuro el consiguiente estudio de la *neuro-fuzzy*, debido a que la lógica difusa adhiere un mecanismo de inferencia sobre la incertidumbre, mientras que las redes neuronales permiten grandes ventajas computacionales, representando los sistemas de inferencia difusa y optimizando el proceso de codificación del conocimiento a través de etiquetas lingüísticas. Todo esto, escenario adecuado y propicio para la predicción, valoración y toma de decisiones.

## Anexo 1: Teoría de Conjuntos Borrosos

Tal y como se ha indicado en el desarrollo del trabajo, en este apéndice se va a poder entender de manera más técnica el origen de la interacción y teoría de los conjuntos borrosos.

Al considerar un conjunto A, habitualmente "x es un elemento de A " o bien "x pertenece a A " se indica " $x \in A$ ". Si todo elemento de A pertenece al conjunto E, se dice que A está incluido en E o que A es un subconjunto de E. Simbólicamente  $A \subset E$ . E se denomina conjunto referencial o universal. La representación clásica mediante el diagrama de Venn, hace que se vea de manera mucho más conceptual [10]

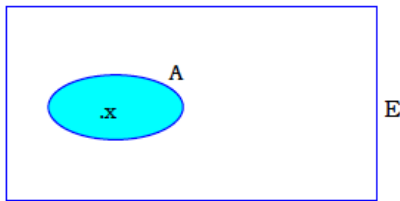


Figura 21

Si matizamos el concepto de función de pertenencia podemos escribir que el concepto de pertenencia o no de un elemento a un conjunto A puede expresarse numéricamente mediante la función de pertenencia, también llamada a veces función característica. Esta función asigna a cada elemento x del universo de discurso un dígito binario (1 o 0) según x pertenezca o no al conjunto A: [11]

$$\mu_A : X \rightarrow \{0,1\} \mid \mu_A(x) = \begin{cases} 1 & \text{cuando } x \in A \\ 0 & \text{cuando } x \notin A \end{cases}$$

cualquier conjunto  $A \subseteq X$  se puede definir por los pares que forman cada elemento x del universo y su función de pertenencia, expresándose de a la siguiente forma:

$$A = \{(x, \mu_A(x)) \mid \forall x \in X\}$$

En un conjunto clásico se asigna el valor 0 o 1 a cada elemento para indicar la pertenencia o no a dicho conjunto. Esta función puede generalizarse de forma que los valores asignados a los elementos del conjunto caigan en un rango particular, y con ello indicar el grado de pertenencia de los elementos al conjunto en cuestión. A dicha función, como arriba se ha definido, se le denomina función de pertenencia y el conjunto por ella definida conjunto borroso. Esta función de pertenencia es definida por  $\mu_A$  para un conjunto borroso A, determinado por el rango de los números entre [0,1], definido de la siguiente manera:

$$\mu_A = X \rightarrow [0,1]$$

Así, cualquier elemento x en X tiene grado de pertenencia  $\mu_A(x) \in [0,1]$ . A queda completamente determinado por:

$$A = \{(x, \mu_A(x)) | x \in X\}$$

A diferencia de un conjunto clásico, los elementos pertenecen o no a él totalmente, por ejemplo, un número puede pertenecer o no a un conjunto determinado, pero no pertenecerá con un determinado grado al conjunto, como se hace en los conjuntos borrosos en donde hay grados de pertenencia en referencia a un universo local.

### Subconjuntos Borrosos

La teoría de subconjuntos borrosos nace en 1965 de la mano de *Lotfi Zadeh* (1965). [12] Su sólida formulación matemática, así como su gran aplicabilidad en problemas de ingeniería o de ciencias sociales, la han convertido en una teoría de gran potencial. Su acierto radica en admitir cualquier valor dentro del intervalo [0,1] como valor de verdad en una oración enunciativa, tal y como hemos podido entender en el trabajo. Considerados dos conjuntos borrosos, A y B, podemos determinar que, A será un subconjunto borroso de B, cuando:

$$\mu_A(x) \leq \mu_B(x), \forall x \in X$$

En la teoría de los conjuntos clásicos, existen operaciones entre ellos, que se extienden a los conjuntos borrosos, como: [13]

### Negación

La negación, también denominada complemento a uno, de un conjunto borroso A con función de pertenencia  $\mu_A(x)$ , es denotada por  $\bar{A}$  y se define con

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$$

### Unión

La unión de dos conjuntos borrosos A y B, sus funciones de pertenencia son  $\mu_A(x)$  y  $\mu_B(x)$ , respectivamente, el cual puede definirse como un conjunto borroso  $C = A \cup B$ , definida por la expresión

$$\mu_C(x) = \mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}, \forall x \in X$$

Que se expresa de manera reducida de la siguiente manera

$$\mu_C = \mu_A \vee \mu_B$$

### **Intersección**

La operación de intersección de dos conjuntos borrosos A y B, con funciones de pertenencia  $\mu_A(x)$  y  $\mu_B(x)$ , respectivamente, se define como un conjunto borroso  $C=A \cap B$  cuya función de pertenencia se relaciona con A y B de la siguiente forma

$$\mu_C(x) = \mu_{A \cap B}(x) = \min \{ \mu_A(x), \mu_B(x) \}, \forall x \in X$$

Existe también otra forma de escribirla de manera abreviada como:

$$\mu_C = \mu_A \wedge \mu_B$$

Las otras construcciones de la teoría que son esenciales son: [14]

### **Fijar la inclusión**

$A \sqsubset B$  si y solamente si  $\forall x$  (por todo x)  
 $\mu_A(x) \leq \mu_B(x)$

### **Fijar la igualdad**

$A = B$  si y solamente si  $\forall x$  (por todo x)  
 $\mu_A(x) = \mu_B(x)$ .

Para terminar la especificación sobre los conjuntos borrosos, en el trabajo se hace notar la diferencia entre subconjuntos borrosos **discretos y continuos**. Con ello se quiere explicar lo siguiente; se puede escribir cualquier universo discreto en la siguiente forma: [11]

$$A = \sum_{x \in X} \mu_A(x) / x$$

En el campo discontinuo, los posibles valores de la variable objeto de estudio son datos concretos distanciados entre sí. En cambio, si hacemos referencia a los subconjuntos borrosos continuos:

$$A = \int \mu_A(x) / x$$

Pudiendo la variable en este escenario, tomar cualquier valor comprendido en la referencia estudiada, evidentemente.

Una vez entendida la base de los conjuntos borroso y subconjuntos, podemos seguir profundizando en un apartado esencial: **los números borrosos**. [12]

Se define un número borroso como un subconjunto borroso del referencial de los reales ( $E = R$ ) tal que la función de pertenencia cumple las siguientes condiciones:

- Existe como mínimo un valor  $x$  tal que  $\mu_{\tilde{A}}(x) = 1$
- $\mu_{\tilde{A}}(s) \geq \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)) \forall x_1, x_2 \in R$  i  $\forall s \in [x_1, x_2]$
- $\mu_{\tilde{A}}$  es una función continua en casi todos los puntos de  $R$
- $\lim_{x \rightarrow -\infty} \mu_{\tilde{A}}(x) = 0$  y  $\lim_{x \rightarrow \infty} \mu_{\tilde{A}}(x) = 0$

Una forma de visionar un número borroso consiste en representar gráficamente su función de pertenencia en unos ejes de coordenadas donde en el eje de abscisas situamos los valores  $x$  reales y en el de ordenadas los valores de  $\mu_{\tilde{A}}(x)$

En la siguiente figura se muestra un ejemplo una representación gráfica de un número borroso cuya forma se parece a la de un triángulo. Cuando hablamos de un número borroso que puede ser representado por cuatro semirrectas, dos de las cuales serán horizontales en el cero, hablamos de un número borroso triangular y son números que permiten imaginar rápidamente un intervalo de valores, así como de un intervalo de confianza para cada uno de los valores.

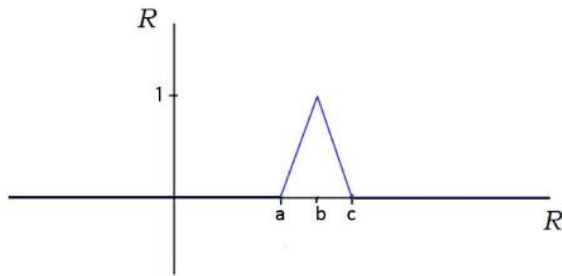


Figura 22

Continuando con los números borrosos triangulares, se demuestra que tienen una función característica del siguiente tipo:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & \text{si } (x < a) \\ \frac{x-a}{b-a} & \text{si } (a \leq x \leq b) \\ \frac{c-x}{c-b} & \text{si } (b \leq x \leq c) \\ 0 & \text{si } (x > c) \end{cases}$$

Vemos así que el hecho de que los valores  $a$ ,  $b$ , y  $c$  sean  $a \leq b \leq c$ , hacen caracterizar el número borroso triangular y permiten escribir el subconjunto borroso triangular de la siguiente manera:

$\tilde{A} = (a, b, c)$ , tal como señalábamos en las páginas correspondientes en el trabajo, pero con notación diferente, recordamos;  $\tilde{A} = (a_1, a_2, a_3)$ •

---

• Tal y como se puede observar y relacionar con el contenido de la página 15

## Anexo 2: Código R

```
1. #NASDAQ-100#
2. install.packages("tseries")
3. library(tseries)
4. install.packages("car")
5. library(car)
6. install.packages("urca")
7. library(urca)
8. install.packages("forecast")
9. library(forecast)
10.   install.packages("fGarch")
11.   library(fGarch)
12.   install.packages("rugarch")
13.   library(rugarch)
14.   install.packages("quantmod")
15.   library(quantmod)
16.   install.packages("stochvol")
17.   library(stochvol)
18.   install.packages("mvtnorm")
19.   library(mvtnorm)
20.
21.   setwd("~/RStudio")
22.   datos_macro<-read.table("^NDX.csv", header=TRUE, sep=",")
23.   getSymbols("^NDX",src='yahoo', from="1999-01-11", to="2002-
  01-11")
24.
25.   NASDAQ<-NDX[,6]
26.   rendNASDAQ<-diff(log(NASDAQ))
27.   length(rendNASDAQ)
28.   names(NASDAQ)
29.   Dates<-as.Date(rownames(zoo(NASDAQ)))
30.   NASDAQ<-as.numeric(NDX$NDX.Adjusted)
31.   names(NASDAQ)
32.   length(NASDAQ)
33.   par(mfrow=c(1,2), font=2, font.lab=4, font.axis=2, las=1)
34.   plot(Dates,NASDAQ,type="l", col="blue",main="Gráfico 1-
  Precios de cierre del Nasdaq" )
35.
36.   #Estadísticos básicos
37.   mean(NASDAQ)
38.   sd(NASDAQ)
39.   min(NASDAQ)
40.   max(NASDAQ)
41.   skewness(NASDAQ)
42.   kurtosis(NASDAQ)
43.   summary(NASDAQ)
44.   #JARQUE BERA
45.   normalTest(NASDAQ,method="jb")
46.   basicStats(NASDAQ)
47.
48.   #Histograma para NASDAQ
49.   par(mfrow=c(1,2), font=2, font.lab=4, font.axis=2, las=1)
50.   hist(NASDAQ)
51.   hist(NASDAQ,nclass=50,col="steelblue")
```

```
52.
53.     par(mfrow=c(1,2), font=2, font.lab=4, font.axis=2, las=1)
54.     acf(NASDAQ, ylim=c(-1,1), main="NASDAQ")
55.     pacf(NASDAQ, ylim=c(-1,1), main="NASDAQ")
56.
57.     Box.test(NASDAQ, lag = 1, type = c("Ljung-Box"))
58.     Box.test(NASDAQ, lag = 5, type = c("Ljung-Box"))
59.     Box.test(NASDAQ, lag = 10, type = c("Ljung-Box"))
60.     Box.test(NASDAQ, lag = 15, type = c("Ljung-Box"))
61.     Box.test(NASDAQ, lag = 20, type = c("Ljung-Box"))
62.
63.     #Contraste de Dicker y Fuller
64.     help(ur.df)
65.     NASDAQ.df<-ur.df(NASDAQ, type = c("none"), selectlags =
66.     c("BIC"))
67.     summary(NASDAQ.df)
68.     NASDAQ.df<-ur.df(NASDAQ, type = c("none"), lags=5)
69.     summary(NASDAQ.df)
70.     #Contrastes de raiz unitaria con el test de Phillips-Perron
71.     NASDAQ.pp<-ur.pp(NASDAQ, type = c("Z-tau"), model =
72.     c("constant"), lags = c("short"))
73.     summary(NASDAQ.pp)
74.     NASDAQ.pp<-ur.pp(NASDAQ, type = c("Z-tau"), model =
75.     c("constant"), lags = c("long"))
76.     summary(NASDAQ.pp)
77.     #Contraste KPSS
78.     NASDAQ.kpss<-ur.kpss(NASDAQ, type = c("tau"), lags =
79.     c("short"))
80.     summary(NASDAQ.kpss)
81.     NASDAQ.kpss<-ur.kpss(NASDAQ, type = c("tau"), lags =
82.     c("long"))
83.     summary(NASDAQ.kpss)
84.     length(NASDAQ)
85.     length(diff(log(NASDAQ)))
86.     diff(log(NASDAQ))
87.     rendNASDAQ.df<-ur.df(diff(log(NASDAQ)), type=c("none"),
88.     selectlags = c("BIC"));
89.     summary(rendNASDAQ.df) #El segundo contraste lo haremos sin
90.     constante
91.     rendNASDAQ.df<-ur.df(diff(log(NASDAQ)),
92.     type=c("drift"), selectlags = c("BIC"))
93.     summary(rendNASDAQ.df)
94.     plot(NASDAQ)
95.     plot(NASDAQ.df) #vemos si existe autocorrelacion en los
96.     residuos.
97.     plot(rendNASDAQ.df)
98.     #Aplicación PP a lo rendimientos
99.     rendNASDAQ.pp<-ur.pp(diff(log(NASDAQ)), type = c("Z-tau"),
100.     model = c("constant"), lags = c("short"))
101.     summary(rendNASDAQ.pp)
```



```
99.      rendNASDAQ.pp<-ur.pp(diff(log(NASDAQ)), type = c("Z-tau"),
      model = c("constant"), lags = c("long"))
100.     summary(rendNASDAQ.pp)
101.
102.     #Aplicación KPSS a los rendimientos
103.     rendNASDAQ.kpss<-ur.kpss(diff(log(NASDAQ)), type = c("mu"),
      lags = c("short"))
104.     summary(rendNASDAQ.kpss)
105.     rendNASDAQ.kpss<-ur.kpss(diff(log(NASDAQ)), type = c("mu"),
      lags = c("long"))
106.     summary(rendNASDAQ.kpss)
107.     plot(rendNASDAQ.kpss)
108.
109.     rendNASDAQ<-diff(log(NASDAQ));head(rendNASDAQ)
110.     length(rendNASDAQ)
111.
112.
113.     #ESTUDIO DE LOS RENDIMIENTOS O RESIDUOS
114.     sd(rendNASDAQ)
115.     min(rendNASDAQ)
116.     max(rendNASDAQ)
117.     skewness(rendNASDAQ)
118.     kurtosis(rendNASDAQ)
119.     summary(rendNASDAQ)
120.     normalTest(rendNASDAQ,method="jb")
121.     basicStats(rendNASDAQ)
122.
123.     #histograma de los rendimientos de BHP
124.     par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
125.     hist(rendNASDAQ)
126.     hist(rendNASDAQ,nclass=50,col="steelblue")
127.
128.     #Histograma de los rendimientos
129.     par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
130.     help("curve")
131.
132.     #NORMAL
133.     hist(rendNASDAQ,breaks=20,freq=F, main = 'Histograma de los
      rendimientos')
134.     curve(dnorm(x, mean=mean(rendNASDAQ), sd=sd(rendNASDAQ)),
      col=2, add=T)
135.     qqPlot(rendNASDAQ, envelope=F)#normal
136.
137.
138.     #TSTUDENT
139.     hist(rendNASDAQ,breaks=20,freq=F, main = 'Histograma de los
      rendimientos')
140.     curve(dstd(x, mean=mean(rendNASDAQ), sd=sd(rendNASDAQ)),
      col=2, add=T)
141.     qqPlot(rendNASDAQ,distribution="t", df=5, envelope=F)
142.
143.     qqnorm(rendNASDAQ)
144.     qqline(rendNASDAQ,datax = FALSE)
145.
146.     qqnorm(rendNASDAQ)
147.     qqline(rendNASDAQ,datax = FALSE)
148.
```

```
149. #prueba para NASDAQ
150. rendNASDAQ<-diff(log(NASDAQ))
151. acf(rendNASDAQ,ylim=c(-1,1),main="rendNASDAQ")
152. pacf(rendNASDAQ,ylim=c(-1,1),main="rendNASDAQ")
153.
154. rendNASDAQ2<-diff(log(NASDAQ))^2
155. acf(rendNASDAQ2,ylim=c(-1,1),main="rendNASDAQ1")
156. pacf(rendNASDAQ2,ylim=c(-1,1),main="rendNASDAQ1")
157.
158. #Autocorrelación de los rendimientos
159. Box.test(rendNASDAQ, lag = 5, type = c("Ljung-Box"))
160. Box.test(rendNASDAQ, lag = 10, type = c("Ljung-Box"))
161. Box.test(rendNASDAQ, lag = 15, type = c("Ljung-Box"))
162. Box.test(rendNASDAQ, lag = 20, type = c("Ljung-Box"))
163.
164. #Autocorrelación de los rendimientos al cuadrado
165. Box.test(rendNASDAQ2, lag = 5, type = c("Ljung-Box"))
166. Box.test(rendNASDAQ2, lag = 10, type = c("Ljung-Box"))
167. Box.test(rendNASDAQ2, lag = 15, type = c("Ljung-Box"))
168. Box.test(rendNASDAQ2, lag = 20, type = c("Ljung-Box"))
169.
170. #IDENTIFICACIÓN Y DIAGNÓSTIC MODELO PARA LA MEDIA
171. win.graph(width=8,height=5)
172. par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
173. acf(rendNASDAQ,ylim=c(-1,1),main="rendNASDAQ")
174. pacf(rendNASDAQ,ylim=c(-1,1),main="rendNASDAQ")
175.
176. Box.test(rendNASDAQ, lag = 1, type = c("Ljung-Box"))
177. Box.test(rendNASDAQ, lag = 2, type = c("Ljung-Box"))
178. Box.test(rendNASDAQ, lag = 3, type = c("Ljung-Box"))
179. Box.test(rendNASDAQ, lag = 4, type = c("Ljung-Box"))
180. Box.test(rendNASDAQ, lag = 5, type = c("Ljung-Box"))
181. Box.test(rendNASDAQ, lag = 10, type = c("Ljung-Box"))
182. Box.test(rendNASDAQ, lag = 15, type = c("Ljung-Box"))
183. Box.test(rendNASDAQ, lag = 20, type = c("Ljung-Box"))
184.
185.
186. #arma(1,1)
187. model3 = arima(rendNASDAQ, order = c(1,0,1),include.mean =
TRUE)
188. model3
189. pnorm(c(abs(model3$coef)/sqrt(diag(model3$var.coef))),
mean=0, sd=1, lower.tail=FALSE)
190.
191. p1=c(1,-model3$coef[1:2]);p1
192. roots=polyroot(p1)
193. Mod(roots)
194.
195. tsdiag(model3)
196. win.graph(width=8,height=5)
197. par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
198. acf(model3$residuals,ylim=c(-1,1),main="residuosarma(1,1)")
199. pacf(model3$residuals,ylim=c(-1,1),main="residuosarma(1,1)")
200. Box.test(model3$residuals, lag = 1, type = c("Ljung-Box"))
201. Box.test(model3$residuals, lag = 2, type = c("Ljung-Box"))
202. Box.test(model3$residuals, lag = 3, type = c("Ljung-Box"))
203. Box.test(model3$residuals, lag = 4, type = c("Ljung-Box"))
```

```
204.     Box.test(model3$residuals, lag = 5, type = c("Ljung-Box"))
205.     Box.test(model3$residuals, lag = 10, type = c("Ljung-Box"))
206.     Box.test(model3$residuals, lag = 15, type = c("Ljung-Box"))
207.     Box.test(model3$residuals, lag = 20, type = c("Ljung-Box"))
208.
209.     qqnorm(model3$residuals)
210.     qqline(model3$residuals, datax = FALSE)
211.     plot(model3$residuals)
212.     title (main="Gráfico de los residuos")
213.     normalTest(model3$residuals,method="jb")
214.     qqnorm(model3$residuals)
215.     qqline(model3$residuals, datax = FALSE)
216.     plot(model3$residuals)
217.     title (main="Gráfico de los residuos")
218.     normalTest(model3$residuals,method="jb")
219.
220.     #Predicción
221.     serie2<-ts (rendNASDAQ[1:653],)
222.     length(serie2)
223.     length(rendNASDAQ)
224.     #estimamos el modelo de nuevo
225.     model = arima(serie2, order = c(1,0,1),include.mean = TRUE)
226.     model
227.     pnorm(c(abs(model$coef)/sqrt(diag(model$var.coef))), mean=0,
sd=1, lower.tail=FALSE)
228.     # predecimos
229.     serie2f<-predict(model,n.ahead=100)
230.     show(serie2f)
231.     ts.plot(serie2f$pred)
232.     ts.plot(rendNASDAQ[653:752],serie2f$pred,col=1:2)
233.     length(rendNASDAQ)
234.     # Evaluación capacidad predictiva
235.     errors<-rendNASDAQ[653:752]-serie2f$pred
236.     ts.plot(rendNASDAQ[653:752]-serie2f$pred,col=1)
237.     eqm<-sum(errors*errors)/100
238.     reqm<-eqm^(.5)
239.     eam<-sum(abs(errors))/100
240.     epam<-sum(abs(errors)/abs(rendNASDAQ[653:752]))/100
241.     eqm
242.     reqm
243.     eam
244.     epam
245.     predicciones<-matrix(0,nrow=1,ncol=1)
246.     prediccion<-function(s)
247.     {
248.         rendNASDAQ<-ts(rendNASDAQ[1:(653+i)],)
249.         model = arima(rendNASDAQ, order=c(1,0,1),
include.mean=TRUE)
250.         rendNASDAQ<-predict(model,n.ahead=1)
251.         prediccion<-(rendNASDAQ$pred)
252.     }
253.     for (i in 0:100)
254.     {
255.         predicciones<-rbind(predicciones,prediccion(i))
256.     }
257.     predicciones<-predicciones[-1] #quitamos el 0 inicial.
258.     predicciones<-ts(predicciones, start=653, end=752)
```

```
259.     ts.plot(rendNASDAQ[653:752],predicciones,col=c('gray0','dodge
rblue'), lwd=3, main='Predicción Recursiva vs. Observación Real')
260.     legend("bottomleft", c("Obs. Real","Predicción"), lty=c(1,1),
col=c('gray0','dodgerblue'), lwd=3)
261.
262.     # Evaluación capacidad predictiva
263.     errors<-rendNASDAQ[653:752]-predicciones
264.     ts.plot(errors,col=2)
265.     eqm1<-sum(errors*errors)/100
266.     reqm1<-eqm^(.5)
267.     eam1<-sum(abs(errors))/100
268.     epam1<-sum(abs(errors)/abs(rendNASDAQ[653:752]))/100
269.     eqm1
270.     reqm1
271.     eam1
272.     epam1
273.
274.     ##ESTUDIO DE UN MODELO PARA LA VARIANZA
275.     #EGARCH
276.
277.     #distribución t-student
278.     spec1=ugarchspec(variance.model=list(model="eGARCH",
garchOrder = c(1,1)), mean.model=list(armaOrder=c(1,1)),
distribution.model = "std")
279.     m12=ugarchfit(spec=spec1,data=rendNASDAQ)
280.     m12
281.     plot(m12)
282.
283.     #Obtención de la volatilidad estimada y los residuos
284.     #NASDAQ
285.     v = sigma(m12)
286.     v=ts(v,frequency=250,start=c(1999,1)) #start: enero del 2012
287.     plot(v,xlab='year',ylab='vol',type='l',ylim=c(0,.05))
288.     write.table(v,file = "volatility.csv", sep = ";")
289.
290.     v = sigma(m12) # Obtain volatility
291.     resi=residuals(m12,standardize=T) # Standardized residuals
292.     vol=ts(v,frequency=250,start=c(1999,1))
293.     res=ts(resi,frequency=250,start=c(1999,1))
294.     par(mfcol=c(2,1)) # Show volatility and residuals
295.     plot(vol,xlab='year',ylab='volatility',type='l',ylim=c(0,.05)
)
296.     plot(res,xlab='year',ylab='st. resi',type='l')
297.     par(mfcol=c(2,2)) # Obtain ACF & PACF
298.     acf(resi,lag=20)
299.     pacf(resi,lag=20)
300.     acf(resi^2,lag=20)
301.     pacf(resi^2,lag=20)
302.
303.     #DIAGNOSIS
304.     Box.test(resi^2,lag=1,type='Ljung')
305.     Box.test(resi^2,lag=2,type='Ljung')
306.     Box.test(resi^2,lag=3,type='Ljung')
307.     Box.test(resi^2,lag=4,type='Ljung')
308.     Box.test(resi^2,lag=5,type='Ljung')
309.     Box.test(resi^2,lag=10,type='Ljung')
310.     Box.test(resi^2,lag=15,type='Ljung')
```

```
311.     Box.test(resi^2,lag=20,type='Ljung')
312.
313.
314.     #opció 3 (volatilidad y rendimientos juntos)
315.
316.     v = sigma(m12) # Obtain volatility
317.     par(mfcol=c(2,1)) # Show volatility and returns
318.     plot(v,xlab='year',ylab='volatility',type='l',ylim=c(0,.05))
319.     plot(dates,rendNASDAQ,xlab='year',ylab='returns',type='l')
320.
321.
322.     #COMPARARACIÓN VOLATILIDAD CON LOS RENDIMIENTOS
323.     #NASDAQ
324.     returnsabs=abs(rendNASDAQ)
325.     time = data.frame(v)
326.     time2=data.frame(returnsabs)
327.     v = sigma(m12)#para obtener la volatilidad estimada
328.     v=ts(v,frequency=250,start=c(1999,1))
329.     plot(v,xlab='year',ylab='vol',type='l',ylim=c(0,.05))
330.     write.table(v,file = "volatility.csv", sep = ";")
331.     ts.plot(time2,gpars= list(xlab="time", ylab="", col =
1:ncol(time)))
332.     legend("topleft", c("returnsabs","v"), lty=c(1,1),
col=c("black","red"))
333.
334.     #NASDAQ MODELO EGARCH (Asimétrico)
335.     ni=newsimpact(z = NULL, m12)
336.     plot(ni$zx, ni$zy, ylab=ni$yexpr, xlab=ni$xexpr, type="l",
main = "News Impact Curve")
337.
338.     #PREDICCIÓN
339.     #NASDAQ ROLLING
340.     spec1=ugarchspec(variance.model=list(model="eGARCH",
garchOrder = c(1,1)), mean.model=list(armaOrder=c(1,1)),
distribution.model = "std")
341.     m18=ugarchfit(spec=spec1,data=rendNASDAQ, out.sample = 100)
342.     forc = ugarchforecast(m18, n.ahead=100, n.roll= 100)
343.     fpm(forc) summary measures else will return NA
344.     sigma(forc)
345.     fitted(forc)
346.     par(mfcol=c(1,2))
347.     plot(forc)
348.     plot(forc,which="all")
```

### Anexo 3: Leyenda

- Figura [1]: Cronología histórica Fuzzy Logic, extraída de [5]
- Figura [2]: Matriz booleana borrosa, extraída de [5]
- Figura [3]: Matriz asimétrica de relaciones, extraída de [5]
- Figura [4]: Matriz simétrica de relaciones, extraída de [5]
- Figura [5]: Cuadro resumen de relaciones entre elementos del referencial E, extraída de [5]
- Figura [6]: Funcionamiento aplicación Neuro – Fuzzy, extraída de [6]
- Figura [7]: Cuadro resultados tras aplicación de mecanismos de inteligencia, extraída de [6]
- Figura [8]: Precios a cierre de la serie escogida en Nasdaq-100, elaboración propia en R
- Figura [9]: Función de Autocorrelación total de la serie, elaboración propia en R
- Figura [10]: Cuadro resumen de los elementos principales de la serie, elaboración propia
- Figura [11]: Histograma de la serie de Nasdaq-100, elaboración propia en R
- Figura [12]: Funciones de Autocorrelación parcial y total de los rendimientos de la serie, elaboración propia en R
- Figura [13]: Funciones de Autocorrelación parcial y total, al cuadrado, de los rendimientos de la serie, elaboración propia en R
- Figura [14]: Cuadro de resultados del test L-Jung box para los rendimientos de la serie, elaboración propia
- Figura [15]: Cuadro de resultados del test L-Jung box para los rendimientos al cuadrado de la serie, elaboración propia
- Figura [16]: Histograma de la serie de rendimientos, elaboración propia en R
- Figura [17]: Evolución teórica de los rendimientos siguiendo una T-Student, elaboración propia en R
- Figura [18]: Funciones de Autocorrelación parcial y total de los rendimientos, una vez diferenciados y modelizados, elaboración propia en R
- Figura [19]: Gráfico predicción recursiva de la serie de rendimientos, elaboración propia R
- Figura [20]: Resultados tras el test de autocorrelación a los residuos del modelo, elaboración propia en R
- Figura [21]: Diagrama de *Venn*, explicativo conjuntos borrosos extraída de [10]
- Figura [22]: Gráfico explicativo número borroso triangular, extraída de [12]
- Figura [23]: Predicción condicional, método *Rolling*, elaboración propia en R

## 9. Referencias.

- [1] MÚNERA, S., MORENO, J., ESCOBAR. (2013). *A Technical Analysis Indicator Based On Fuzzy Logic*
- <[https://ac.els-cdn.com/S1571066113000054/1-s2.0-S1571066113000054-main.pdf?\\_tid=26480f13-aa13-436f-8ea9-596a9b561b5c&acdnat=1527609885\\_a15c52be626e79689d42b4bcdcab952](https://ac.els-cdn.com/S1571066113000054/1-s2.0-S1571066113000054-main.pdf?_tid=26480f13-aa13-436f-8ea9-596a9b561b5c&acdnat=1527609885_a15c52be626e79689d42b4bcdcab952)> [Consulta: 2018]
- [2] HERNÁN MORELLO, J. (2015). *Lógica Borrosa aplicada a las finanzas. Valuación de proyectos.*
- <[http://www.jatic2015.ucaecemdp.edu.ar/trabajos/1.8.Logica\\_Borrosa\\_aplicada\\_a\\_las\\_finanzas.\\_Valuacion\\_de\\_proyectos.pdf](http://www.jatic2015.ucaecemdp.edu.ar/trabajos/1.8.Logica_Borrosa_aplicada_a_las_finanzas._Valuacion_de_proyectos.pdf)>  
[Consulta: 2018]
- [3] GIL-ALUJA, J. (1999). *Elementos para una Teoría de la Decisión en la Incertidumbre.* Editorial Milladoiro. España
- [4] BETANCUR GUTIÉRREZ, JC. (2006). *Aplicación de los Conjuntos Borrosos a las Decisiones de Inversión.*
- <<https://repository.eafit.edu.co/handle/10784/7689>>  
[Consulta: 2018]
- [5] ÁLVAREZ VIZCARRA, G. (2014). *Lógica Borrosa, efectos olvidados y exposición al riesgo cambiario.*
- <[https://www.researchgate.net/profile/Gumaro\\_Vizcarra/publication/309155771\\_Logica\\_Borrosa\\_Efectos\\_Olvidados\\_y\\_Exposicion\\_al\\_riesgo\\_cambiario/links/58016bd008ae1c5148c9fe77/Logica-Borrosa-Efectos-Olvidados-y-Exposicion-al-riesgo-cambiario.pdf](https://www.researchgate.net/profile/Gumaro_Vizcarra/publication/309155771_Logica_Borrosa_Efectos_Olvidados_y_Exposicion_al_riesgo_cambiario/links/58016bd008ae1c5148c9fe77/Logica-Borrosa-Efectos-Olvidados-y-Exposicion-al-riesgo-cambiario.pdf)>  
[Consulta: 2018]
- [6] ABRAHAM, A., NINAN SAJITH, P., SARATCHANDRAN, P. (2004). *Modeling Chaotic Behavior of Stock Indices using Intelligent Paradigms.*
- <https://pdfs.semanticscholar.org/1513/861593c470827367229c09c500af7a746abb.pdf> >  
[Consulta: 2018]
- [7] BUSTELO DE LA CASA, A. (2017). *Análisis de Modelo en Series Financieras.* Trabajo. Barcelona: Universitat de Barcelona
- [8] KAUFMANN, A., GIL-ALUJA, J. (1991). *Nuevas técnicas para la dirección estratégica.* Universitat de Barcelona Editors. España.
- [9] GIL-ALUJA, J. (1990). *Las matemáticas del Azar y la Incertidumbre.* Centro de Estudios Ramon Areces S.A. España.

[10] L. LAZZARI, L., A. M. MACHADO, E., H. PÉREZ, R. (1999). *Los Conjuntos Borrosos: Una Introducción*.

<<https://dialnet.unirioja.es/servlet/articulo?codigo=2469891>>  
[Consulta: 2018]

[11] BASIL M, AL H., DÍAZ NAFRÍA, JM. (2010). *Fuzzy Logic*.

<<http://glossarium.bitrum.unileon.es/Home/logica-borrosa>>  
[Consulta: 2018]

[12] RONDÓS CASAS, E., FARRERAS NOGUER, MA., LINARES MUSTARÓS, S. (2015). *El Número Borroso Triangular "Ratio Acid-Test Mínima"*.

<[http://bibliotecadigital.econ.uba.ar/download/cuadcimbage/cuadcimbage\\_n18\\_03.pdf](http://bibliotecadigital.econ.uba.ar/download/cuadcimbage/cuadcimbage_n18_03.pdf)>  
[Consulta: 2018]

[13] RAMÍREZ PÉREZ, N., LAGUNA ESTRADA, M. (2012). *La lógica borrosa: conjuntos borrosos, razonamiento aproximado y control borroso*.

<<http://pistaseducativas.itc.mx/wp-content/uploads/2013/01/6-RAMIREZ-PE-100-55-65.pdf>>  
[Consulta: 2018]

[14] WATKINS, T. *Lógica Confusa: La lógica de conjuntos borrosos*.

<<http://www.applet-magic.com/fuzzysetsp.htm>>  
[Consulta: 2018]

[15] CORREA-HENAO, GJ., MONTOYA SUÁREZ, LM. (2012). *Aplicación del Modelo ANFIS para Predicción de Series de Tiempo*.

<<http://www.funlam.edu.co/revistas/index.php/lampsakos/article/view/927/1024>>  
[Consulta: 2018]

[16] Bravo Narvaéz, YL., García Vélez, DA (2002). Estudio y aplicación de los modelos ANFIS de MATLAB.

<<http://bibdigital.epn.edu.ec/handle/15000/5463>>  
[Consulta: 2018]