

Simulación de escenarios financieros en asignación
de carteras de inversión mexicanas

L.C.C Meliza Contreras González

August 21, 2006

Contenido

1	Escenarios	7
1.1	Escenarios y Clasificaciones	8
1.1.1	Árboles de Escenarios	8
1.1.2	Clasificaciones	9
1.2	Métodos de Generación de Escenarios	9
1.2.1	Generador de Escenarios	9
1.2.2	Enfoques de Optimización	10
1.2.3	Enfoques Estadísticos	10
1.2.3.1	<i>Arranque con Datos Históricos Previos</i>	10
1.2.3.2	<i>Muestreo Aleatorio</i>	10
1.2.3.3	<i>Muestreo Aleatorio Ajustado</i>	11
1.2.3.4	<i>Enlace de Momentos</i>	11
1.2.3.5	<i>Métodos basados en Caminos</i>	11
1.2.3.6	<i>Conglomerados</i>	11
1.2.3.7	<i>Componentes Principales</i>	12
2	Carteras de Inversión Mexicanas	15
2.1	Conformación de la Cartera de Inversión	16
2.1.1	Búsqueda de Fuentes Fidedignas de Información	17
2.1.2	Recopilación de Información de los Activos	18
2.2	Análisis Exploratorio de la Cartera de Inversión	20
2.2.1	Determinación de Activos Significativos	21
2.2.1.1	Obtención de Activos esenciales con las Componentes Principales	21
2.2.2	Comprobación del Comportamiento Lognormal de la Cartera	21
2.2.3	Comportamiento de los Conglomerados de la Cartera	24
3	GeneArEsce	27
3.1	Modelo en Espiral de GeneArEsce	27
3.2	Análisis de las Funcionalidades de GeneArEsce	29
3.2.1	Requerimientos esenciales en la conformación de un Árbol de Escenario	30
3.2.2	Simulación del Comportamiento de los Árboles de Escenario	32

3.3	Diseño de GeneArEsce	34
3.3.1	Interfaz de Usuario	36
3.4	Algoritmos de GeneArEsce	41
3.4.1	Householder, QR y Cholesky	46
3.4.1.1	Householder	46
3.4.1.2	QR	46
3.4.1.3	Cholesky	46
4	GeneArEsce en Carteras de Inversión	47
4.1	Validación de Resultados	47
4.1.1	Comprobación de la correspondencia de la simulación con la realidad	47
4.1.2	Variabilidad de los Resultados de los periodos del Árbol de Escenario	48
4.1.3	Resultados de los periodos con el método de Conglomerados	48
4.1.3.1	Variabilidad entre los periodos	48
4.1.3.2	Variabilidad dentro de los periodos	50
4.1.4	Resultados de los periodos con el método de Conglomerados y Componentes Principales	51
4.1.4.1	Variabilidad entre los periodos	51
4.1.4.2	Variabilidad dentro de los periodos	51
4.1.5	Resultados del comportamiento del Árbol de Escenario con los métodos de generación de Árboles de Escenario	52
5	Conclusiones	55

Desde el inicio de la ciencia la única certidumbre que puede asegurarse es el cambio, tanto en el tiempo como en el espacio, así para estudiar los fenómenos que ocurren en la naturaleza y en particular en la sociedad, es necesario recolectar toda la información histórica, posteriormente analizarla y encontrar patrones de comportamiento que contribuirán a conocer el futuro de tales fenómenos.

Una de las ramas de la ciencia que presenta procesos altamente inestables, es la Economía, cuyos movimientos bursátiles son motivo de estudios minuciosos con el objetivo de determinar los mayores beneficios sin comprometer el capital a riesgos innecesarios. Desafortunadamente esta realidad resulta utópica, pues en cada transacción intervienen aún de forma implícita demasiados factores interrelacionados entre sí que impiden establecer con certeza la mejor opción de inversión [Ale03].

Ante esta problemática se han realizado modelos matemáticos que la representen, así para la toma de decisiones financieras es necesario resolver programas estocásticos que proporcionen un abanico de posibilidades de los pros y contras de invertir en determinadas opciones de cartera [Kau03] [Kou01].

Una alternativa para crear este abanico se le denomina generador de escenarios, que es un proceso que construye distribuciones discretas para las variables de decisión involucradas. A la estructura resultante de su aplicación se le designa como Árbol de Escenarios o Árbol de Eventos, donde cada nodo representa el valor de esa variable de decisión en ese instante de tiempo y las ramas representan las probabilidades de ocurrencia de éstos valores; de este modo las ramas que inician en el nodo raíz y culminan hasta los nodos hojas se les llama escenarios, cuya probabilidad se determina por el producto de cada rama a lo largo del camino. Los árboles se clasifican, dependiendo del tiempo de planificación, como árboles monoperiodo o multiperiodo y dependiendo del número de variables de decisión involucradas, como árboles univariados o multivariados [Kau03].

En este trabajo se presenta una alternativa para generar escenarios mediante las metodologías de simulación y análisis multivariado con las Componentes Principales, los Conglomerados (Cluster) y el híbrido de la combinación de los dos anteriores. Los árboles resultantes se sometieron a distintos análisis de calidad para determinar qué árboles son más aptos para utilizarse en problemas de carteras de inversión mexicanas. La ventaja de emplear la simulación radica en que se puede obtener una gran cantidad de datos sin realizar demasiadas observaciones, basta un generador de variables aleatorias adecuado, mientras que con el análisis multivariado se logra reducir la redundancia de las muestras de los datos; la aplicación de estas técnicas se comentan en [Lor97] [Gul02].

Así este documento se constituye de los siguientes apartados: el capítulo 1 presenta los antecedentes de mayor importancia de las metodologías de generación de Escenarios y también se expone la problemática a resolver de manera detallada. El capítulo 2 detalla los requisitos previos para la elaboración de los Árboles de Escenario incluyendo la conformación de la Cartera de Inversión y el Análisis Exploratorio de los datos resultantes. El capítulo 3 presenta la metodología en espiral para la aplicación estadística GeneArEsce. Como etapa inicial se contempló la simulación del comportamiento y estructura de los Árboles de

Escenario mediante S-Plus. Con los resultados obtenidos se conformó la especificación de requerimientos. Una vez identificados éstos se procedió al diseño de las funcionalidades para plantear el modelo en la Plataforma .NET cuyo resultado es la aplicación GeneArEsce que incluye opciones estadísticas fundamentales como la obtención de matrices de distancias, correlación, varianza y covarianza, opciones avanzadas de análisis multivariado como generación de Dendrogramas, Conglomerados, Componentes Principales y opciones de generación de Árboles de Escenarios con los métodos de Conglomerados, Conglomerados y Componentes Principales y Escenarios mediante las Componentes Principales. El capítulo 4 muestra los árboles resultantes arrojados por cada método de generación de Escenarios y un análisis detallado de sus propiedades estadísticas y comportamiento histórico. Finalmente, en el capítulo 5 se establecen las conclusiones y los detalles que hemos considerado relevantes y suficientes para la defensa de esta Tesis.

Capítulo 1

Escenarios

En cada suceso de la vida cotidiana se genera un abanico de posibilidades hasta el momento en que se elige una opción, que una vez llevada a cabo, resulta imposible invertir sus efectos. Sin embargo, los eventos ya acontecidos constituyen una base de conocimiento que otorga experiencia y acota el universo de opciones hacia tendencias que pueden ser beneficiosas o perjudiciales.

Este tipo de fenómenos se manifiestan en varias áreas del conocimiento. En particular, la Economía se caracteriza por la incertidumbre de su comportamiento a pesar de que existan registros históricos almacenados segundos antes del movimiento bursátil actual. Los enfoques para estudiar este comportamiento oscilan desde algunos planteamientos complejos mediante modelos teóricos hasta estrategias de generación de incertidumbre como la simulación, que es el enfoque fundamental de esta disertación.

Aunque surge la interrogante sobre el número de posibilidades a considerar de tal forma que resulten suficientes para desempeñar un análisis satisfactorio sobre las tendencias de estos movimientos económicos, lo ideal es generar el mayor número de opciones que las herramientas computacionales provean y aplicarle a este amplio conglomerado técnicas multivariadas que clasifiquen las tendencias y el orden de importancia de las variables de decisión involucradas, enfoque primordial para el tratamiento de la información de este trabajo.

Además de las metodologías y los conjuntos de información, un elemento indispensable es la estructura de acceso a la información. Mientras más sencilla menos elementos de variación considerará, si se vuelve más compleja se acercará al comportamiento real. Si a esto le agregamos un factor de decisión como la evolución de las variables de decisión en el tiempo, la estrategia de solución más factible que abarca todos los elementos por sus características de operación son los Árboles de Escenario cuya generación es la motivación esencial de este proyecto.

1.1 Escenarios y Clasificaciones

Los modelos de optimización de carteras de inversión plantean la generación de un conjunto extenso y heterogéneo de posibilidades para obtener la frontera de eficiencia de carteras de inversión eficientes. En particular, los planteamientos se resuelven para modelos multietapas con un alto nivel de incertidumbre, estos planeamientos establecen sus metodologías en modelos de programación estocástica como los Modelos Factoriales y de Escenarios[Jac05].

Los Modelos Factoriales asumen que los retornos de las variables de decisión en la cartera de inversión dependen linealmente del movimiento de uno o más *factores* comunes de varios activos como: el factor de mercado global y los factores industriales entre otros. El uso de estos modelos además de acelerar los cálculos, reducen los requerimientos de entradas que son más fácilmente ajustados para reflejar las condiciones de cambio, en lugar de considerar los coeficientes de la matriz de covarianza [Jac05].

Otra enfoque de solución son los Modelos de Escenario que describen las relaciones entre las variables de decisión involucradas; esto se realiza mediante la enumeración de distintos escenarios que reflejan el comportamiento futuro y estiman tanto la media como la varianza de cada variable de decisión en cada uno de los escenarios obtenidos[Jac05].

Los algoritmos con mayor eficiencia empleando Modelos de Escenario son similares a los utilizados en Modelos Factoriales. Respecto a la cantidad de información recopilada para el modelado a partir de datos históricos, se recomienda considerar una cantidad de variables de decisión mayor en comparación de la cantidad de observaciones obtenidas, por ejemplo, casos encontrados en la práctica incluyen miles de variables de decisión pero sólo docenas de meses o cientos de días consideradas como una cantidad de observaciones aceptables [Jac05].

1.1.1 Árboles de Escenarios

Al emplearse los Modelos de Escenario se genera una estructura conocida como Árbol de Escenarios o Árbol de Eventos, que muestra el comportamiento de las variables de decisión sobre un horizonte de planeación vinculando el tiempo y la incertidumbre[Jac05].

La interpretación del árbol resulta interesante: como cualquier árbol tradicional se constituye de nodos y ramas, en este caso cada nodo representa el valor de la variable de decisión en un instante de tiempo, mientras que las ramas representan las probabilidades de ocurrencia de éstos valores

La altura del árbol indica el horizonte de planeación, es decir, en el nodo raíz se considera el instante actual (tiempo=0) y conforme se recorren en profundidad los nodos hijos se realiza un avance en la unidad de tiempo respectiva, culminando con los nodos hojas que representan el periodo final(horizonte de planeación).

Así la complejidad del árbol radica en el número de nodos hojas ya que se expande exponencialmente, lo que indica el número de Escenarios conformados.

Éstos últimos son cada uno de los caminos que van desde el nodo raíz hasta alguno de los nodos hojas y su probabilidad se determina mediante el producto de las probabilidades condicionales de los nodos a lo largo del camino en el caso de periodos dependientes o como el producto de las probabilidades de los nodos en el caso de periodos independientes [Hoy01].

1.1.2 Clasificaciones

Los Árboles de Escenario se clasifican respecto a los siguientes criterios[Kau01]:

- **Variables de Decisión:** si sólo existe una variable de interés, el Árbol es univariado, pero si al menos hay un par de estas variables, el Árbol se considera multivariado.
- **Tiempo de Planificación:** si sólo se considera el siguiente periodo de tiempo a partir de la información histórica se le denomina Árbol Mono-periodo. Si el punto de interés es construir una Base de Datos con el comportamiento futuro en varias etapas se denota como Árbol Multiperiodo.

1.2 Métodos de Generación de Escenarios

Una alternativa para crear el abanico de posibilidades para la toma de decisiones se le denomina Generador de Escenarios. Dado el interés que representa, por la variabilidad y heterogeneidad de los resultados, se han realizado extensos estudios con metodologías estadísticas y de optimización para el planteamiento de algoritmos eficientes. Algunos de ellos se detallan a continuación.

1.2.1 Generador de Escenarios

La finalidad de un Generador de Escenarios es producir un conjunto de valores de las variables de decisión involucradas bajo un determinado horizonte de planeación, cuya salida es un escenario o el conjunto de ellos. Además, dado que es un proceso complejo, necesita un conjunto de entradas con información suficiente y válida, que permita conocer el comportamiento de la distribución de los datos. Los dos elementos presentes son:

- **Base de Datos Histórica:** contiene el comportamiento histórico de las variables de decisión involucradas.
- **Horizonte de Planeación:** representa el número de niveles o unidades de tiempo para el comportamiento futuro, a partir del comportamiento actual dado por la Base de Datos Histórica.

1.2.2 Enfoques de Optimización

Desde el punto de vista de la Optimización se resuelven problemas de optimización no lineal, donde las variables de decisión son los retornos y las probabilidades, mientras que la función objetivo y las restricciones obligan a satisfacer las propiedades estadísticas incluyendo momentos centrales y penalizando las desviaciones de los momentos deseados en la función objetivo. [Gul02] [Hoy01] [Kau03]

La resolución del problema plantea dos directivas:

- **Optimización Secuencial:** Se aplica en el caso de que el Árbol de Escenario se construya considerando la ramificación en cada nodo separadamente. Así, para cada nodo se resuelve un problema de optimización no lineal.
- **Optimización Total:** Se aplica cuando se consideran todos los nodos del árbol y se genera el árbol completo en un problema de optimización no lineal demasiado amplio.

1.2.3 Enfoques Estadísticos

Desde el punto de vista estadístico, pueden citarse las siguientes técnicas de generación de escenarios:

1.2.3.1 *Arranque con Datos Históricos Previos*

Este enfoque genera escenarios empleando los datos disponibles sin implicar técnicas de modelado. Cada escenario se considera una muestra de las variables de decisión obtenidas mediante un muestreo de los valores observados en el pasado. Las fechas de los periodos históricos son seleccionadas aleatoriamente y para cada fecha en la muestra se revisan los retornos de todas las variables involucradas o los factores de riesgo durante el mes anterior a la fecha seleccionada, si estos escenarios son generados en meses. Si se desea un horizonte de planeación amplio, por ejemplo un año, se realizan muestras de los doce retornos mensuales desde distintos puntos en el tiempo [Kou01].

1.2.3.2 *Muestreo Aleatorio*

Este método parte de estimar los coeficientes y la matriz de covarianza del modelo de vectores autoregresivos, que pudieran ser los retornos anuales de activos como bonos, depósitos y acciones. Por ejemplo, si se desea un árbol de eventos con diez nodos después de un año (cada etapa se encuentra en la misma unidad de tiempo), simplemente se repite este procedimiento diez veces, muestreando vectores independientes de los retornos para cada nodo. Cabe señalar que los nodos en la segunda etapa del árbol además de ser muestreados aleatoriamente, presentan una distribución condicional de la primera etapa hacia la segunda, dado que dependen de las salidas de la primera. Sin embargo, este

método se vuelve intratable si se incrementa el número de nodos en cada etapa debido al crecimiento exponencial del árbol en cuestión [Kau01].

1.2.3.3 *Muestreo Aleatorio Ajustado*

Este método intenta reducir las carencias del método anterior, asumiendo un número impar de nodos, y ajustando los momentos pares de la distribución que presentan los datos. Por ejemplo, si se consideran diez nodos sucesores en cada etapa, se muestrean cinco vectores con términos de error del modelo de vector autoregresivo, los términos de error del resto de los nodos son idénticos a los anteriores con signos opuestos. Posteriormente se reescalan los valores muestreados en el orden de ajuste de la varianza y, finalmente, los valores ajustados para los términos de error son substituidos en las ecuaciones del modelo de vector autoregresivo para generar un conjunto de nodos para el árbol de eventos. Cabe señalar que el esfuerzo computacional para ajustar las muestras aleatorias no es factible [Kau01]

1.2.3.4 *Enlace de Momentos*

En este caso el método estima los retornos enlazando sus primeros momentos(media, varianza, coeficiente de asimetría y kurtosis) bajo la distribución muestreada. Esto se realiza mediante la resolución de un modelo de optimización no lineal propuesto por Hoyland y Wallace [Hoy01]. Las variables de decisión en el modelo son los retornos y las probabilidades del árbol, mientras que las funciones objetivo con sus restricciones forzan al cumplimiento de las propiedades estadísticas deseadas.

1.2.3.5 *Métodos basados en Caminos*

El proceso se inicia generando caminos completos y el resultado no es un Árbol de Escenarios, sino un conjunto de caminos. Para transformarlo a un árbol de escenarios, los escenarios se someten a técnicas de conglomerados [Kou01].

1.2.3.6 *Conglomerados*

El Análisis de Conglomerados es la búsqueda de grupos en los datos, de tal forma que los objetos pertenecientes al mismo conglomerado posean características comunes, mientras que en los objetos pertenecientes a distintos conglomerados sus características carecen de patrones de correspondencia.

Existen dos tipos de algoritmos de conglomeración:

- **Métodos de Partición:** describen procesos que dividen los datos dentro de k número de conglomerados, donde k es un parámetro proporcionado por el usuario.
- **Métodos Jerárquicos:** describen procesos de construcción de árboles para establecer una estructura a los conglomerados formados a partir de los datos proporcionados y se dividen en:

- *Métodos Divisivos*: inician con el criterio de conglomerado global y lo descomponen parcialmente hasta que todas las observaciones forman conglomerados simples.
- *Métodos Aglomerativos*: inician considerando a cada observación como un conglomerado simple y sucesivamente se anidan hasta conformar un conglomerado global (todas las observaciones están contenidas en él)[Peñ02].

Como método de generación de escenarios los Conglomerados Jerárquicos resultan de gran utilidad por las características inherentes de su funcionamiento. Inicialmente, se identifican las dos observaciones más parecidas (cercanas) que no se encuentren en el mismo conglomerado y se combinan; posteriormente se aplica alguna regla de distancias entre grupos tantas veces como sea necesario, hasta que todas las observaciones se encuentren en un solo conglomerado. Es un método aglomerativo dado que los conglomerados se forman a partir de la combinación de conglomerados existentes.

Si aplicamos éste método para la obtención de los nodos del Árbol de Escenario a partir de una Base de Datos histórica, puede determinarse el número de grupos(nodos hijos del árbol) y posteriormente utilizando técnicas de simulación, se generan nuevas Bases de Datos con ciertos parámetros obtenidos de la inicial, obteniéndose todos los nodos del árbol. Para determinar la probabilidad de ocurrencia en cada nivel se contabilizan el número de individuos en cada grupo y se dividen entre el total de ellos. En las etapas sucesivas la probabilidad de ocurrencia de las ramas se determinan dado el total de individuos del nodo padre y el número de observaciones heredadas a los nodos hijos [Gul02][Hoy01].

Las técnicas de simulación pueden utilizarse de varias formas, generando los árboles mediante métodos Montecarlo [Pus98] o con la Simulación Paralela donde se construyen escenarios independientes hasta el periodo final antes de que el proceso de conglomeración se realice o con la Simulación Secuencial donde primero se crean ramas iniciales para posteriormente aplicar la conglomeración, seleccionar el centroide (moda, mediana, u otra propiedad estadística), establecer éste como nuevo nodo del árbol y aplicar el proceso iterativamente hasta que se alcance el horizonte de planeación [Gul02].

1.2.3.7 Componentes Principales

Este método surge para resolver un problema central en el análisis de datos multivariante: la reducción de la dimensión. Su objetivo es transformar el cúmulo de variables originales del fenómeno estudiado por un subconjunto de variables ficticias bajo el costo de la pérdida mínima de información.

La importancia del Análisis de Componentes Principales radica en el siguiente planteamiento: dadas n observaciones de p variables, se analiza si es posible representar adecuadamente esta información con un número menor de variables construidas como combinaciones lineales de las originales. Esta técnica es debida a Hotelling (1933), aunque sus orígenes se encuentran en los

ajustes ortogonales por mínimos cuadrados introducidos por K. Pearson (1901), sus principales directivas de solución son [Peñ02]:

- Representar óptimamente en un espacio de dimensión reducida, observaciones de un espacio general p -dimensional, por lo que se consideran el paso preliminar para identificar las variables latentes.
- Transformar las variables originales generalmente correlacionadas en nuevas variables sin correlación, facilitando la interpretación de los datos.
- Generar nuevas variables no correlacionadas que expresan la información contenida en el conjunto original de datos para reducir la dimensión (número de variables) del problema estudiado.
- Eliminar algunas de las variables originales si aportan mínima información del fenómeno planteado.
- Sintetizar en cada una de sus componentes la máxima variabilidad residual contenida en los datos.

Ante estas ventajas surge la idea de plantear una estructura que genere comportamientos de posibles escenarios. Estos conceptos fueron retomados por Loretan [Lor97], para construir en lugar de Árboles, escenarios simples aplicando el proceso siguiente: dada la Base de Datos inicial se obtienen las Componentes Principales de las variables de decisión involucradas, se escogen las componentes más significativas y con las nuevas variables obtenidas (combinaciones lineales de las anteriores) se ordenan dada su frecuencia acumulada para la posterior obtención de los cuantiles de sus eventos cola y así conocer los escenarios extremos tanto favorables como desfavorables.

Las metodologías empleadas en este trabajo se orientan hacia la generación de Árboles de Escenario Multivariados-Multiperiodo y contemplan las técnicas de Simulación Secuencial, Conglomerados y Componentes Principales y la generación de Escenarios mediante las Componentes Principales.

Capítulo 2

Carteras de Inversión Mexicanas

La era de la globalización ha conducido a un binomio indivisible: comportamiento de índices financieros y herramientas computacionales potentes tanto para el análisis como para el tratamiento de la información. Dada la importancia de éstos índices es necesario exponer la conceptualización esencial de los procesos financieros que repercuten en la toma de decisiones de inversión.

El proceso de inversión implica la toma de decisiones respecto a un abanico de valores negociables con su respectiva tasa de riesgo, el monto de las inversiones y el periodo de realización de las mismas. Realizar una inversión significa arriesgar el efectivo acumulado actual por una recompensa futura incierta. Esto se realiza mediante la interacción de dos ingredientes fundamentales el tiempo y el riesgo, ocasionando distintos tipos de activos de mercado dependiendo del porcentaje de participación de los mismos, cuya clasificación aparece a continuación [Ale03].

- **Flujo de efectivo(CETES):** la inversión se considera real y corresponde a activos tangibles sin riesgo como los bienes raíces, con la característica de aumentar o disminuir su valor en el tiempo.
- **Bonos:** la inversión se considera financiera, implica la elaboración de contratos de vencimiento, por lo que el riesgo se incrementa parcialmente y el periodo de realización de las operaciones es fijo.
- **Acciones:** la inversión se estima financiera, son activos con el mayor grado de riesgo puesto que el éxito de la operación depende de la decisión de ofertar o adquirir los valores en el instante de tiempo adecuado, lo que no es infactible estimar con seguridad porque la unidad de tiempo de realización es en segundos.

Así la tarea primordial de un análisis de valores es evaluar los activos para determinar sus beneficios futuros posibles, las condiciones bajo las cuales se recibirán las ganancias y la probabilidad de ocurrencia de estas condiciones, es decir

estimar las características de riesgo y rendimiento de los valores en cuestión. Por esta razón los Árboles de Escenario son la metodología idónea, pues su estructura refleja tanto el comportamiento de los activos así como la probabilidad de ocurrencia de los mismos, por lo que es necesario construir una cartera de inversión donde se encuentren inmersos los activos fundamentales[Ale03].

Así el proceso de inversión se compone de los siguientes pasos:

1. **Establecimiento de la política de inversión:** En esta etapa se determinan los objetivos del inversionista y el monto a invertir. El objetivo es concientizar al interesado que se obtendrán ganancias con el riesgo de incurrir en pérdidas significativas.
2. **Análisis de los valores:** En este paso se realiza un estudio minucioso de los precios del mercado para predecir su comportamiento futuro. Se buscan patrones recurrentes en los precios pasados y se examinan los precios recientes para identificar patrones emergentes similares a los del pasado.
3. **Construcción de la cartera:** En esta fase se establece la cartera de inversión que es el proceso de selección de activos determinando su porcentaje de participación.
4. **Revisión de la cartera:** Este paso se considera de actualización de las nuevas necesidades del inversionista o porque las transacciones realizadas en el presente afectan a la cartera anterior o debido a que los precios de los activos cambian significativamente.
5. **Evaluación del desempeño de la cartera:** En esta etapa se precisa el rendimiento ganado por la cartera y el riesgo de pérdida.

2.1 Conformación de la Cartera de Inversión

La elección del tipo de activo a participar en la inversión depende de las necesidades del inversionista, sus preferencias de rendimiento y su tolerancia al riesgo, sin embargo para el proceso de toma de decisiones es indispensable mostrarle al interesado los rendimientos y riesgos de todos los activos. Con esta información se procede a la diversificación cuyo objetivo es combinar distintos tipos de activos en una cartera de tal forma que el nivel de riesgo disminuya a diferencia de una cartera de un sólo tipo de activo debido a que se reducen los extremos en los rendimientos de la cartera por lo que se decrementa la fluctuación en el valor de la misma [Ale03].

Para determinar una cartera de inversión adecuada es importante considerar el comportamiento del mercado. Un mercado eficiente se caracteriza por la aleatoriedad de los rendimientos de los valores. Los cambios aleatorios en el precio son el resultado de la reevaluación de los inversionistas acerca de los prospectos de un valor y del ajuste apropiado de su compra y venta [Ale03].

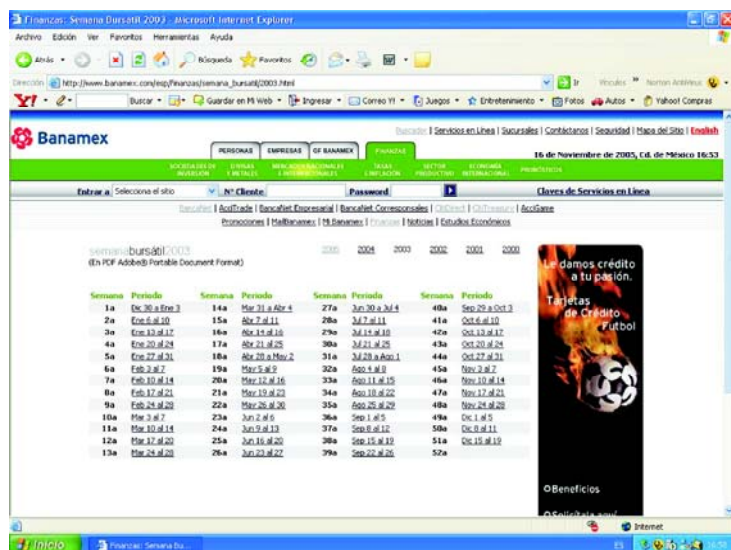


Figura 2.1: Fuentes de Banamex

2.1.1 Búsqueda de Fuentes Fidedignas de Información

Una vez identificados los tipos de activos es necesario encontrar datos fidedignos que representen todos los sectores posibles del fenómeno. Con el propósito de construir una Base de Datos se realizó una intensa búsqueda de información en las principales fuentes de información financiera.

En primer lugar se revisaron los expedientes en línea de Banamex® (ver figura 2.1) en particular la sección de la Semana Bursátil que proporciona información sobre el comportamiento de las Acciones y Bonos Mexicanos por semanas y se encuentra disponible desde el año 2000 hasta el año 2005. Cada uno de los reportes (ver figura 2.2) contiene información del comportamiento de todos los sectores económicos, información de los cierres semanales de la Bolsa Mexicana y gráficos sobre las variaciones de los activos entre otros.

Sin embargo el contenido era demasiado denso y para la conformación de la cartera era indispensable determinar los activos más significativos en el Mercado Mexicano. Así se examinaron periódicos, en particular las secciones financieras del periódico el Universal®, donde se encontró una clasificación preliminar de las principales variables financieras. Los apartados ms importantes abarcaban la sección bursátil con el contenido de las principales Acciones de la Bolsa Mexicana de Valores, las cuotas de Fondo de Renta Fija (Bonos), las cuotas de Fondo de Renta Variable, entre otros índices.

Aunque se contaba con información valiosa, la recuperación y manejo de la misma para su posterior procesamiento resultaba costosa en tiempo, por lo que se buscaron bases de datos financieras que proporcionaran información suficiente

Emisora	Precio al 03-Jan-03	Soporte	Resistencia	Recomendación	Comentario	PMO
BMV	6253.34	6000.00	6300.00		Cerca de importante resistencia	32
Alfa a	16.98	16.40	17.75	R	Continúa consolidando	34
Amel a1	6.28	6.00	6.85	R/Ce	Indicadores con margen de alza	40
Amx1	7.67	7.25	7.90	R/Ce	Compra si rompe resistencia	35
Apaso	62.75	59.00	65.00	RV	Toma de utilidades	40
As	15.93	14.70	16.70	R	En canal lateral	40
Bimbo a	14.97	14.50	16.00	R	Alarma con promedio móvil	38
Calé	9.91	9.55	10.45	R	Mantene tendencias de baja	40
Cemex cpo	46.15	44.00	48.65	R/Ce	Indicadores con margen de alza	40
Cie B	17.20	16.60	17.90	R	Indicadores frenan baja	31
Comercial	5.44	5.00	5.70	R	Alarma con promedio móvil	24
Costal	17.00	15.50	18.00	RV	Toma de utilidades	21
Decc B	3.80	3.50	3.80	R	Soporte en mínimo 12 meses	28
Elektra cpo	24.80	24.20	27.00	R	Indicadores en sobre venta	40
Fernsukid	37.99	37.00	39.60	R	Importante resistencia	22
Comas a1	25.98	25.00	26.55	R	En resistencia	34
Gcc B	7.80	7.50	8.00	RV	Importante resistencia en SO	16
Gra B	19.35	18.50	20.00	R	Indicadores mixtos	40
GMa B	8.26	7.80	8.45	R	Estabilizado en sobre compra	28
Grhorte a	26.69	24.30	27.90	R	Nuevo máximo histórico en 26.50	37
Gisa	14.71	14.00	15.90	R	En consolidación	40
Gracato B	11.94	10.90	12.20	R	Lateral	31

Figura 2.2: Reportes Financieros de Banamex

sobre el comportamiento de los tipos de activos encontrados con anterioridad. La Base de Datos de Economatica®(ver figura 2.3) fue una solución eficaz en este proceso, por lo que se decidió considerarla como la fuente de información de los registros históricos de los activos analizados.

2.1.2 Recopilación de Información de los Activos

Economatica®proporciona las siguientes opciones de consulta (ver figura 2.4):

- **Stock Guide:** Contiene todos los registros históricos de todos los tipos de activos, incluso considera opciones de filtrado por país, tipo de activo, entre otros.
- **Gráficos:** Muestra la gráfica de crecimiento(alza) o decrecimiento(caída) de los activos en cierto periodo de tiempo.
- **Radiografía:** Despliega mediante reportes detallados, la información histórica más significativa de un activo.
- **Indicadores Financieros:** Ofrece resultados del comportamiento de la variable financiera solicitada por el usuario.
- **Indicadores de Mercado:** Proporciona resultados del comportamiento de la variable técnica solicitada por el usuario.
- **Cotizaciones:** Muestra la variabilidad de las Acciones con sus respectivas operaciones de cierre, mínimo, máximo, apertura, volumen y promedio.



Figura 2.3: Base de Datos Economática

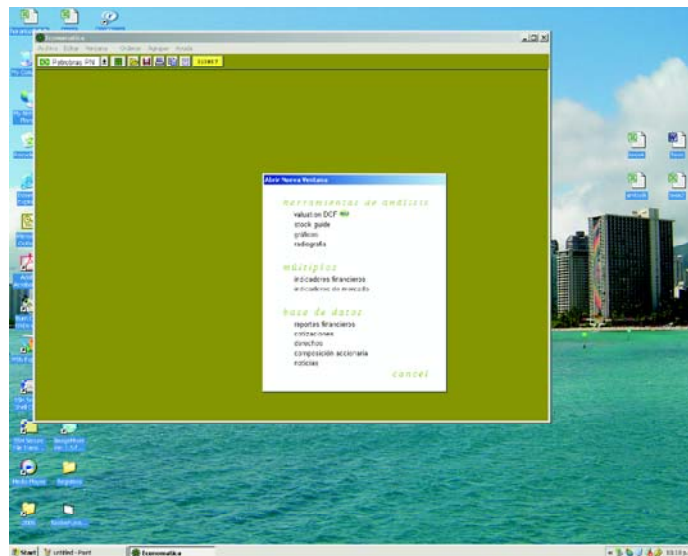


Figura 2.4: Opciones Financieras de Economática

Para la recopilación de los registros de los Activos Accionarios se revisó la opción Stock Guide y Cotizaciones, resultando una Base de Datos consistente cuyos valores correspondían al precio y volumen diario de las Acciones.

En la obtención de los Bonos y Cetes se realizó un filtrado en la opción Stock Guide. Para la recuperación de los Bonos se encontró que el tipo de activo que lo representa en una Base eran las Cuotas de Fondo de Renta Fija mientras que en el caso de los Cetes el tipo de activo fue el Índice de Precios. Sin embargo la opción Cotizaciones no arrojaba resultados para estos activos pero la opción Stock Guide ofrecía información específica de la variable financiera: porcentaje de variación del cierre en un intervalo de tiempo determinado para todo tipo de activos.

Sin embargo, una cartera de inversión se conforma por activos con distintos niveles de riesgo, las acciones sólo representaban el mayor nivel de riesgo, más era necesario proporcionar activos con un riesgo medio como los Bonos y libres de riesgo como los CETES (flujo de efectivo). Como paso adicional para establecer una cartera de inversión fidedigna y su posterior simulación, un factor proporcional es la unidad de medida de los activos. Dado que las unidades de las acciones de la Bolsa Mexicana son los precios y los Bonos y Cetes se miden respecto a su porcentaje de variación de cierre en el tiempo, se invirtió tiempo en el estudio de las opciones de Economatica para unificar criterios de medición.

Después de una revisión minuciosa se decidió como unidad de unificación la variación del precio de los tres activos.

El último aspecto a considerar fue la unidad de tiempo de estas variaciones. Inicialmente se consideraron los días, sin embargo, las variaciones de Bonos y Cetes perdían validez por sus características económicas inherentes. La siguiente unidad evaluada fue la semana sin embargo los avances fueron insuficientes y se perdía información valiosa. Finalmente la unidad considerada fue el mes dado que para los tres activos el orden de magnitud fue aceptable.

2.2 Análisis Exploratorio de la Cartera de Inversión

Una vez conformada la cartera de inversión se procede a la aplicación de metodologías efectivas para comprobar el funcionamiento deseado, los datos atípicos y la información fidedigna del fenómeno estudiado. Una de las herramientas empleadas para cumplir estos objetivos es el Análisis Exploratorio de los Datos para garantizar la calidad de los mismos.

Los objetivos fundamentales de este análisis fueron los siguientes:

- Determinación de Activos Significativos:
- Comprobación del Comportamiento Lognormal de la Cartera:
- Comportamiento de los Conglomerados de la Cartera:

2.2.1 Determinación de Activos Significativos

Los datos obtenidos en Economatica contenían valores nulos para algunos de los activos involucrados. La información recopilada inicialmente abarcaba el comportamiento de los activos del año 2000 hasta el 2005, contando con 192 Acciones, 5 CETES y 362 Bonos, pero al ser sometida la información fue sometida a un proceso de filtrado para la eliminación de datos no disponibles en años específicos, queda reducido a los siguientes activos: 40 Acciones, 236 Bonos, y 5 Cetes, para un total de 281 activos. Posteriormente fue necesario reducir el número de activos con un criterio de calidad que sólo ofreciera las variables más significativas.

2.2.1.1 Obtención de Activos esenciales con las Componentes Principales

La técnica multivariada que brindó un criterio de solución fueron las Componentes Principales pues mediante éstas se determinan las variables más importantes. En particular, la finalidad era conformar la cartera con tan sólo 50 activos y al aplicar la técnica se obtuvieron los siguientes resultados:

- Las primeras 47 componentes explicaban la variabilidad del fenómeno en un 99%.
- Las primeras tres componentes clasificaron la importancia de los tipos de activos.
- Las Acciones presentaron la mayor ponderación.
- Una cantidad mínima de Bonos se encontraban en las Componentes importantes.
- Los Cetes se encontraban ausentes en las Componentes de mayor relevancia.

Dado que estaba definido de antemano que los tres tipos de activos deberían considerarse, la cartera de inversión se conformó bajo la siguiente ponderación: desde el punto de vista de la proporción de la importancia se consideraron 35 acciones y 10 Bonos y por otro, lado en proporción de la representatividad de los activos, se mantuvieron los 5 CETES originales.

2.2.2 Comprobación del Comportamiento Lognormal de la Cartera

Después de conocer los activos esenciales se procede a comprobar la Lognormalidad de la muestra, dado que en la mayoría de las fuentes de información [Jam01] se menciona el comportamiento de los activos como lognormales, por lo que se procedió a realizar Pruebas de Hipótesis para su evaluación.

En una primera aproximación se consideró aplicar la función logaritmo a cada uno de los elementos de la muestra y así obtener una nueva muestra, si ésta

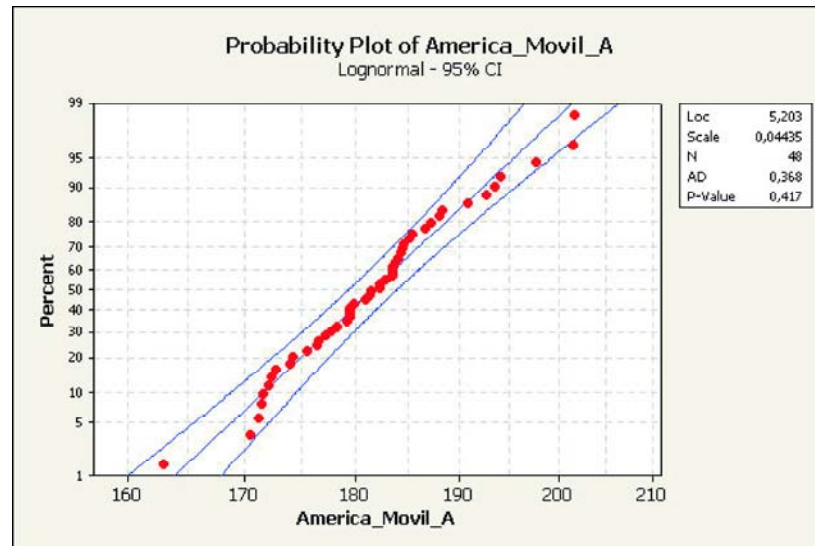


Figura 2.5: Prueba de Hipótesis para las Acciones

presentaba comportamiento normal entonces la muestra original por ende era lognormal. Sin embargo algunas valores de las observaciones contenían valores que indeterminaban la función logaritmo. Así que se consideraron dos criterios de solución:

- **Eliminación de las variables o de las observaciones que impiden la conversión de los datos:** Este caso es demasiado radical y provocaría pérdida de información significativa y no fue aplicado en la investigación.
- **Conversión de los valores a datos válidos para la Transformación Normal:** En este caso la solución impide perder la información, y consiste en realizar una traslación de los datos agregando un valor constante mayor en valor absoluto a todos los elementos de la muestra con lo que se evitan las inconsistencias en la conversión, éste enfoque fue considerado en este trabajo.

Ante estos resultados se optó por analizar cada una de las variables de la cartera por separado y verificar su Lognormalidad mediante la Prueba de Hipótesis de Anderson, que reporta si las variables de estudio son aceptadas o rechazadas como Lognormales, éstas pruebas fueron realizadas en Minitab.

En el caso de las Acciones (ver Figura 2.5) la prueba aceptó la hipótesis nula en la mayoría de los casos con un valor de $p=0.4$, salvo algunos casos en que la hipótesis nula fue rechazada debido a que algunos valores de sus observaciones salían del rango por una diferencia mínima, aunque la mayoría de los valores se

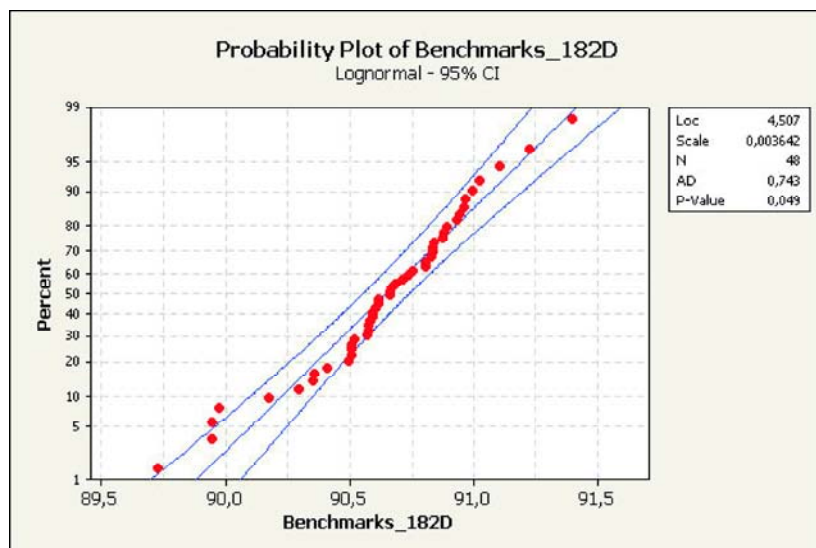


Figura 2.6: Prueba de Hipótesis para los CETES

encontraban inmersos en el rango válido de la distribución, por lo que se asumió considerar a las Acciones como variables Lognormales.

En el caso de los CETES (ver Figura 2.6) la prueba rechazó a las variables con un valor de $p=0.049$, más como se observa en la figura el rechazo se debe a que algunas observaciones salen del rango de la distribución lognormal, más la mayoría de las observaciones si lo cumplen, por lo que los CETES también se consideraron lognormales.

En el caso de los Bonos (ver Figura 2.7) la prueba rechazó la hipótesis nula con un valor de $p=0.005$, sin embargo, a diferencia de los casos anteriores, la figura muestra que más del 90% de las observaciones se encuentran fuera del rango de la distribución lognormal, por lo que se concluye que los Bonos presentan un comportamiento totalmente distinto al esperado. Así que se consideraron dos criterios de solución:

- **Eliminación de las variables:** Este caso es demasiado radical y sería válido en el caso de considerar sólo riesgos altos y bajos, por lo que no fue aplicado para la investigación.
- **Conversión de los valores a datos válidos para la Transformación lognormal:** En este caso la solución impide perder la información, y consiste en realizar una transformación de los datos a una distribución normal y posteriormente a una Lognormal.

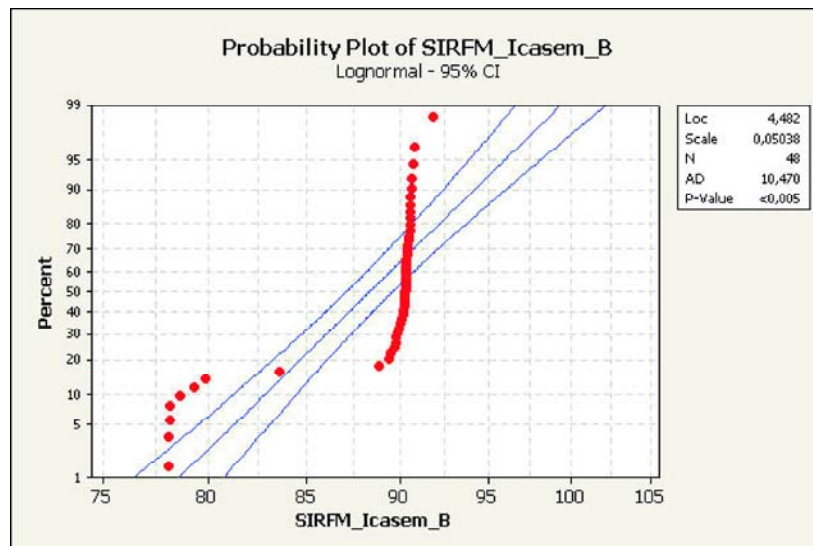


Figura 2.7: Prueba de Hipótesis para los Bonos

2.2.3 Comportamiento de los Conglomerados de la Cartera

El Análisis de Conglomerados arrojó que algunas observaciones generaban grupos aislados, como el grupo 1 y 19 que aparecen en la figura 2.8 rodeados por un círculo verde. Así que se consideraron dos criterios de solución:

- **Eliminación de las observaciones aisladas:** sin embargo estas observaciones aportan información que puede suceder en la vida real, por ejemplo una devaluación, por lo que no fue considerada una buena solución.
- **Preservación de las observaciones aisladas:** En este caso estas observaciones se consideran casos extremos y fue la solución considerada en este trabajo

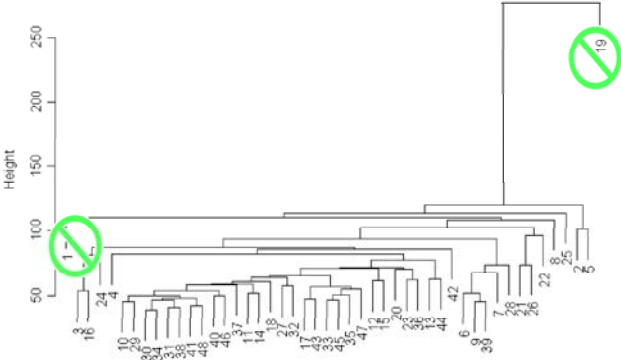


Figura 2.8: Conglomerados de la Cartera

Capítulo 3

GeneArEsce

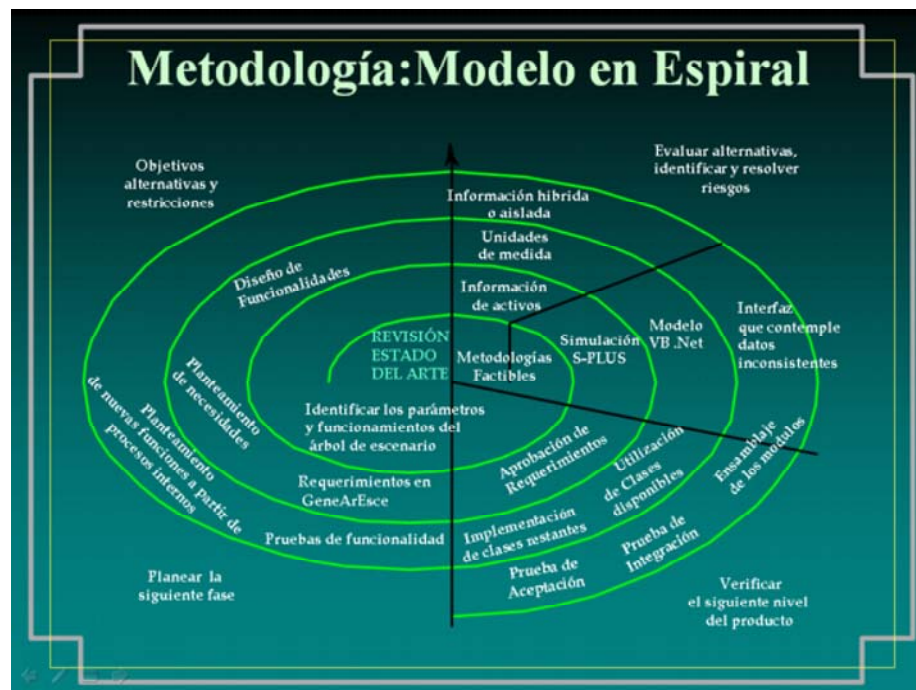
En este apartado se presenta la metodología en espiral para la aplicación estadística GeneArEsce®. Como etapa inicial se contempló la simulación del comportamiento y estructura de los Árboles de Escenario mediante S-Plus®. Con los resultados obtenidos se conformó la especificación de requerimientos. Una vez identificados éstos se procedió al diseño de las funcionalidades para plantear el modelo en la Plataforma .NET® cuyo resultado es la aplicación GeneArEsce® que incluye opciones estadísticas fundamentales como la obtención de matrices de distancias, correlación, varianza y covarianza, opciones avanzadas de análisis multivariado como generación de Dendrogramas, Conglomerados, Componentes Principales y opciones de generación de Árboles de Escenarios con los métodos de Conglomerados, Conglomerados y Componentes Principales y Escenarios mediante las Componentes Principales.

3.1 Modelo en Espiral de GeneArEsce

Para la generación de aplicaciones de calidad, las Ciencias de la Computación dieron origen a la Ingeniería de Software, cuya aportación más importante es el establecimiento de metodologías efectivas para realizar una planificación adecuada de los recursos, el tiempo y las mejoras de un Sistema de Software. En particular el Modelo en Espiral es ampliamente utilizado en la programación de etapas como alternativa de solución de proyectos complejos.

En este proyecto se empleo el Modelo en Espiral para cada una de las etapas tanto de elaboración de los requerimientos, la obtención de la Base de Datos histórica y la implementación de las funcionalidades.

Como se muestra en la figura 3.1, como primera etapa de la Investigación se realizó un extenso estudio sobre el estado del arte de los métodos de generación de escenarios. Luego se decidió que metodologías eran factibles para el proyecto. Con estas consideraciones se identificaron los parámetros y algoritmos de generación de los Árboles de Escenario, terminando así la primera espiral de nuestro modelo que se refleja en el primer capítulo de este documento.

Figura 3.1: *Modelo en Espiral*

En la siguiente espiral, como etapa inicial se realizó una extensa búsqueda de los activos que conformarían la Base de Datos Histórica, referida en el Capítulo 2 de este documento, para generar su simulación posterior mediante la plataforma S-PLUS®, que permitió identificar los requerimientos básicos de GeneArEsce® y establecer el planteamiento de las necesidades que debería cubrir el sistema.

La tercera espiral intensificó el refinamiento de los datos logrando la unificación de las unidades de medida de los tres tipos de activos seleccionados (ver Capítulo 2). Posteriormente se planteó el modelo de generación de escenarios mediante la plataforma Visual Basic .NET®; se realizó una búsqueda exhaustiva de las clases y métodos disponibles para la aplicación. Luego se implementaron las bibliotecas matemáticas esenciales para el funcionamiento de GeneArEsce® y cada uno de los módulos desarrollados fueron sometidos a pruebas de funcionalidad, cotejando con paquetes matemáticos como MathLab® y S-PLUS® que los resultados obtenidos en GeneArEsce fueran similares. Ante estos resultados y dado que los requerimientos de GeneArEsce tan sólo consistían en la generación de los Árboles de Escenario se decidió agregar como funcionalidades los algoritmos estadísticos de Conglomerados, Componentes Principales, Matrices de Varianzas y Covarianzas, Matrices de Correlación disponibles para el usuario y así aprovechar las técnicas estadísticas previamente estudiadas para la implementación de los Árboles de Escenario.

En la siguiente espiral se realizó la validación de la Base de Datos Histórica mediante un Análisis Exploratorio de los datos obtenidos (ver Capítulo 2). Los resultados del análisis arrojaron inconsistencia de la información por lo que fue necesario diseñar alternativas para la manipulación previa de estos datos de forma tal que el módulo de generación de Árboles de Escenario fuera modificado mínimamente. Posteriormente los módulos se integraron formando la aplicación GeneArEsce® que se detalla a continuación.

3.2 Análisis de las Funcionalidades de GeneArEsce

Como funcionalidades fundamentales en GeneArEsce se deducieron en orden de prioridad las siguientes:

- **Generación de Árboles de Escenario:** Este apartado encapsula todos los algoritmos para obtener Árboles de Escenario.
- **Generación de Escenarios:** Este apartado abarca los métodos de Componentes Principales y percentiles para generar Escenarios
- **Generación de Estadísticas:** Este apartado incluye los Algoritmos Estadísticos y de Álgebra Lineal empleados para la generación de los Árboles.
- **Exportación de Resultados:** Este apartado abarca los mecanismos de exportación de las Bases de Datos resultantes y los reportes sobre información estadística.

El diagrama de Casos de Uso de GeneArEsce aparece en la figura 3.2, donde se plantean dos actores que interactúan con el sistema, el *Investigador*, un usuario que interactúa con todas las opciones estadísticas y el *Optimizador* que es un sistema externo cuyo interés radica en la obtención de Árboles de Escenario. El *Investigador* puede solicitar los siguientes Casos de Uso:

- **Imprimir Reportes:** Este apartado encapsula todos los reportes tanto de la Matriz de Covarianzas y Varianzas (CovVar), la Matriz de Correlación (Correl), Conglomerados y Componentes Principales (ComponentesP).
- **Exportar Árbol:** Este apartado guarda en una Base de Datos los Árboles obtenidos con los métodos de Conglomerados y Conglomerados con Componentes Principales.
- **Proporcionar Estadísticas:** Este apartado incluye la obtención de la Matriz de Varianzas y Covarianzas, la Matriz de Correlación, la elaboración de Conglomerados y el cálculo de las Componentes Principales.
- **Generar Escenario:** Este apartado abarca los mecanismos de Componentes Principales con percentiles para generar Escenarios
- **Generar Árbol:** Este caso es compartido con el actor *Optimizador* abarca los mecanismos de generación de Árboles de Escenario con Conglomerados y el híbrido Componentes Principales y Conglomerados.

3.2.1 Requerimientos esenciales en la conformación de un Árbol de Escenario

Los generadores de Árboles de Escenario presentan como resultado un Árbol representado con una Base de Datos cuyos registros contienen los siguientes elementos:

- **id_nodo:** Número del nodo actual.
- **Activos:** Cada una de las variables de decisión involucradas, en este caso el conjunto de Acciones, Cetes y Bonos.
- **Probabilidad de ocurrencia:** La probabilidad de que ocurran los valores del nodo actual en el nivel actual.
- **Padre:** Nodo del que desciende el nodo actual.
- **Nivel:** Periodo de tiempo en donde transcurre el comportamiento de los activos en el nodo actual.

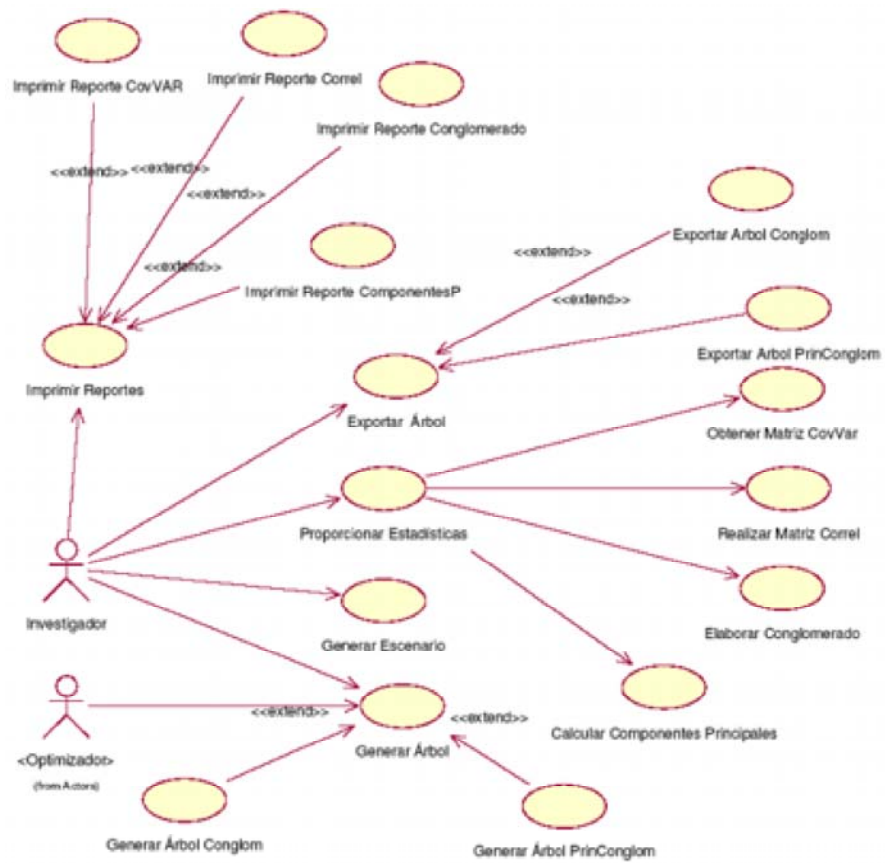


Figura 3.2: Diagrama de Casos de Uso

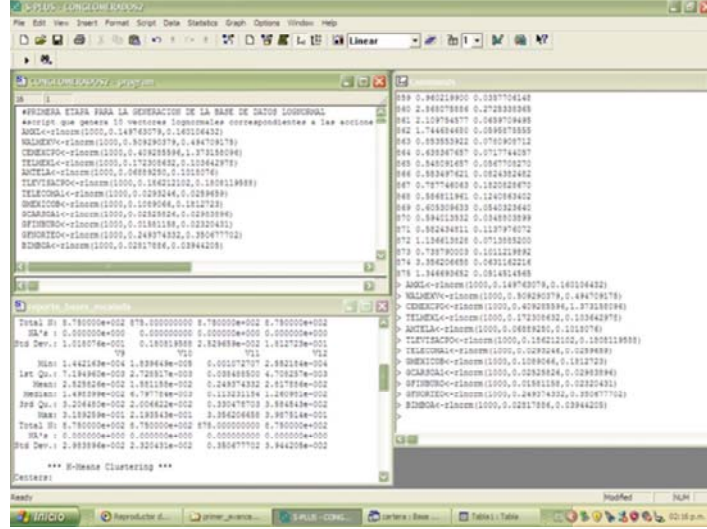


Figura 3.3: Plataforma S-PLUS

3.2.2 Simulación del Comportamiento de los Árboles de Escenario

Una vez identificados los requerimientos se procedió a simular el funcionamiento del Árbol de Escenario con la plataforma estadística S-PLUS®.

En la figura 3.3 se muestra la plataforma S-PLUS, donde se realizaron los primeros ejemplos de los Árboles de Escenario utilizando el método de Conglomerados y el híbrido Conglomerados-Componentes Principales para su posterior implementación en .NET®. Las funciones estadísticas empleadas para la simulación fueron:

- **hclust:** Función que realiza conglomerados jerárquicos de una muestra de datos multivariada.
- **princomp:** Función que genera las Componentes Principales de una muestra de datos multivariada.
- **rlnorm:** Función que simula variables aleatorias lognormales.

En la figura 3.4 aparece la estructura de un Árbol de Escenario aplicando la técnica de Conglomerados. Inicialmente el nodo padre consiste en un grupo conformado por los valores medios de los activos a estudiar, luego se propone el número de grupos candidatos y a partir de ellos se obtienen Bases de Datos

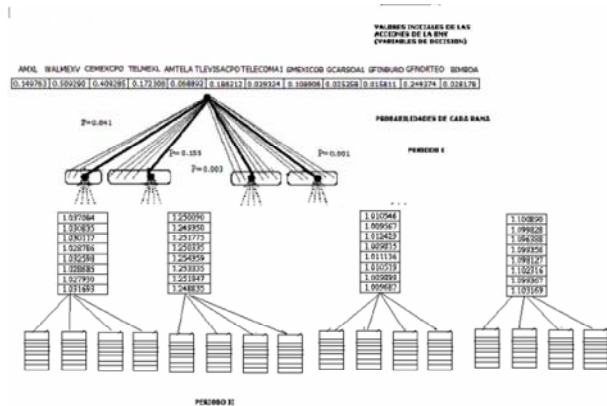


Figura 3.4: *Árbol de Escenario con Conglomerados*

Lognormales. La probabilidad de ocurrencia se genera a partir del número de individuos contenidos en cada uno de los grupos (muestra) respecto a la población total. Así para cada uno de los grupos se almacenan los nuevos valores medios de los activos y se aplica nuevamente el proceso anterior para cada uno de los grupos lo que implica que la complejidad del *Árbol de Escenario* es exponencial, por lo que su representación gráfica al incrementar el número de niveles es infactible visualizar.

En la figura 3.5 aparece la estructura de un *Árbol de Escenario* aplicando la técnica de Conglomerados y Componentes Principales. Inicialmente el nodo padre consiste en un grupo conformado por los valores medios de las Componentes Principales de los activos a estudiar, luego se procede con la técnica de Conglomerados como en el caso anterior, por lo que su representación gráfica es infactible visualizar al incrementar el número de niveles.

El proceso de simulación anterior resulta de gran utilidad en la visualización del proceso de Generación de *Árboles de Escenario*, arroja los detalles mínimos a considerar en el diseño de GeneArEsce y refleja la complejidad de su implementación con librerías que sólo devuelven los resultados más no explican a detalle los algoritmos empleados, por lo que fue necesario realizar una investigación minuciosa sobre métodos de simulación, Análisis Multivariado y Bases de Datos para llevar a cabo exitosamente la aplicación GeneArEsce.

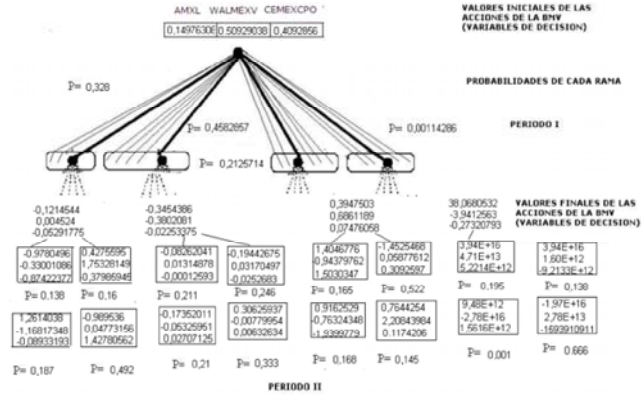


Figura 3.5: *Árbol de Escenario con Conglomerados y Componentes Principales*

3.3 Diseño de GeneArEsce

Al identificar las funcionalidades se procedió a diseñar las clases con sus respectivos métodos y atributos. Como clases iniciales se crearon *Estat*, *AlgeLin* y *Dendrograma*. Las clases *Estat* y *AlgeLin* son utilizadas en las clases *GenMuLog*, *PrinC* y *Conglomerate*. Por otro lado la clase *Dendrograma* es empleada en la clase *Conglomerate*. Además las clases *Conglomerate* y *GenMuLog* se usan en las clases de generación de Árboles como las clases *T.Conglom*, *T.PrinConglom*, mientras que *PrinC* es empleado por *PrinCEs*. Además las clases *T.Conglom* y *T.PrinConglom* son descendientes de la clase *T.Escenario*.

Estat contiene como atributo esencial un Objeto Matriz y cuyos métodos primordiales considera:

- **Gen_MxCovVar:** Función que genera la Matriz de Varianzas y Covarianzas de una muestra de datos multivariada.
- **Gen_MxCorrel:** Función que genera la Matriz de Correlaciones de una muestra de datos multivariada.
- **Gen_AxMedias:** Función que genera el vector de Medias de una muestra de datos multivariada.

AlgeLin contiene como atributo esencial un Objeto Matriz y cuyos métodos importantes considera:

- **Householder:** Función que genera las Transformaciones de Householder
- **QR:** Función que aplica la función QR.
- **Gauss:** Función que genera la Sustitución hacia atrás de Gauss
- **Cholesky:** Función que produce la Factorización de Cholesky.

Dendrograma contiene como atributo esencial un Objeto de tipo Base de Datos y cuyos métodos importantes considera:

- **Aplicar_Criterio_Clasificación:** Función que emplea el criterio de En-cadenamiento Simple en el ajuste de distancias
- **Actualizar_Dendrograma:** Función que inserta registros a la Base de Datos Dendrograma.
- **Generar Grupo:** Función que genera cada uno de los grupos para el Dendrograma

GenMuLog contiene como atributos esenciales un Objeto Arreglo de Medias y un objeto Matriz de Varianza y Covarianza, cuyos métodos importantes considera:

- **Gen_Normales:** Función que genera variables aleatorias normales
- **Gen_Multivariate_Normales:** Función que genera variables aleatorias multivariadas normales
- **Gen_Multivariate_Lognormales:** Función que genera variables aleatorias multivariadas lognormales

PrinC contiene como atributo esencial un Objeto Matriz y cuyos métodos importantes considera:

- **Gen_PrinC:** Función que genera las Componentes Principales
- **Gen_Scores:** Función que genera los scores dadas las Componentes Principales
- **Gen_Loadings:** Función que genera el porcentaje de variación acumulado

Conglomerate contiene como atributos esenciales un Objeto Matriz de Distancias, el criterio de distancia y el nivel del Dendrograma a seleccionar y cuyos métodos importantes considera:

- **Importar_Datos:** Función que importa la Base de Datos seleccionada por el usuario
- **Gen_Mx_Distance:** Función que genera la Matriz de Distancias

T_Escenario contiene como atributos esenciales un Objeto Base de Datos y un objeto Cabecera para personalizar los campos de la Base de Datos que contendrá el Árbol de Escenario, cuyos métodos importantes considera:

- **Importar_Datos:** Función que importa la Base de Datos seleccionada por el usuario
- **Generar_Ramas:** Función que genera con un recorrido a lo ancho las ramas del Árbol
- **Construir_Árbol:** Función que actualiza los registros de los nodos en la Base de Datos

Las clases restantes son: PrinCes que cuenta con el método Gen_Escenario a partir de la clase PrinC y la clase T_PrinConglom que utiliza el método Gen_Arbol_CP para generar Árboles con las Componentes Principales, mientras que T_Conglom es una clase heredada de T_Escenario.

3.3.1 Interfaz de Usuario

La interacción del usuario con GeneArEsce es un punto de atención en el desarrollo de la aplicación, como se observa en la figura 3.7, donde se presenta la plataforma de inicio de GeneArEsce. Dentro de las opciones más significativas en la figura 3.8 se muestra el Dendrograma de la opción Conglomerados.

Por otro lado en la figura 3.9, aparece la estructura del Árbol de Escenario visto como una Base de datos mediante la opción de *Generar Árbol con Conglom.* Mientras que en la figura 3.10 aparece la estructura del Árbol de Escenario con la opción *Generar Árbol con Conglom y CP.*

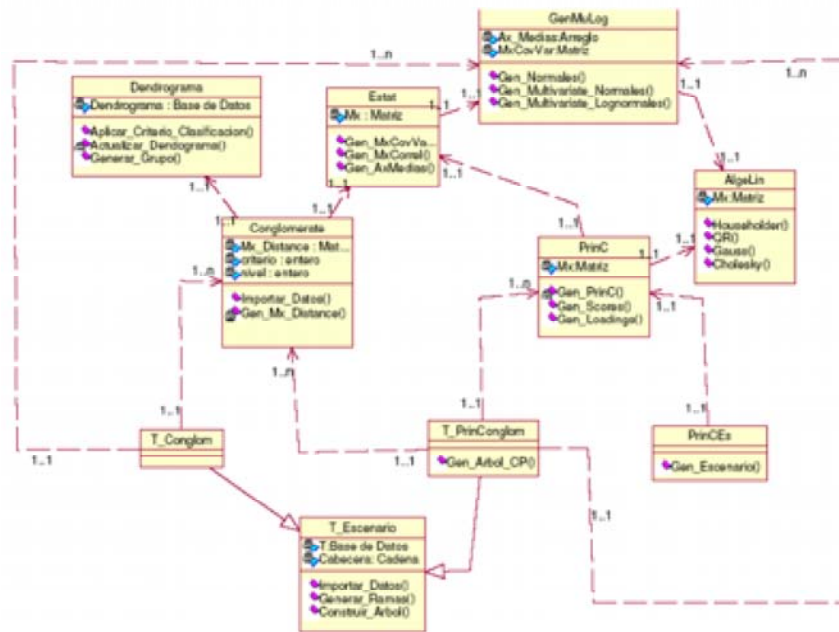
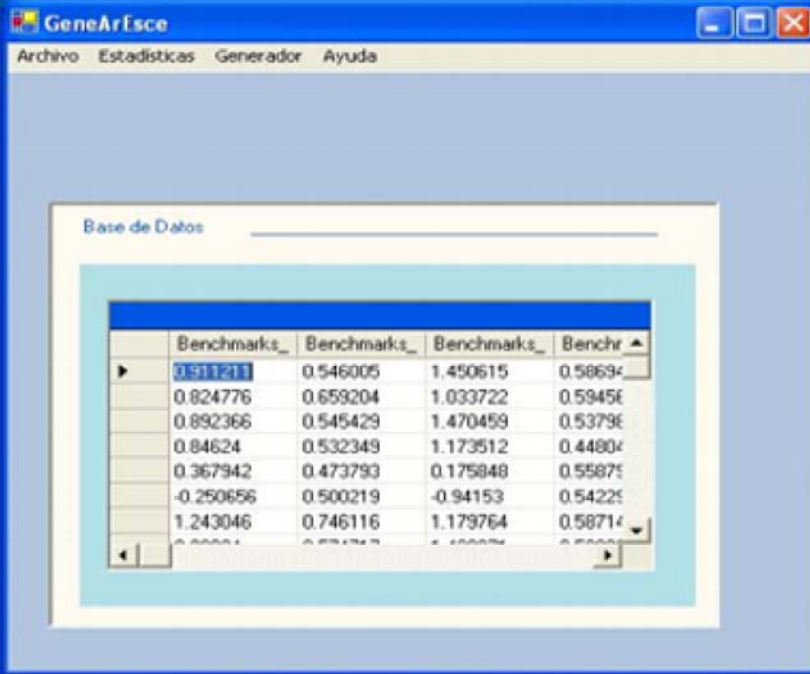


Figura 3.6: Diagrama de Clases de GeneArEsce



The screenshot shows the GeneArEsce application window. The title bar reads "GeneArEsce" and the menu bar includes "Archivo", "Estadísticas", "Generador", and "Ayuda". The main area is titled "Base de Datos" and contains a table with four columns: "Benchmarks_", "Benchmarks_", "Benchmarks_", and "Benchr ▲". The first cell of the first row is highlighted in blue.

	Benchmarks_	Benchmarks_	Benchmarks_	Benchr ▲
▶	0.931711	0.546005	1.450615	0.58694
	0.824776	0.659204	1.033722	0.59456
	0.892366	0.545429	1.470459	0.53796
	0.84624	0.532349	1.173512	0.44804
	0.367942	0.473793	0.175848	0.55875
	-0.250656	0.500219	-0.94153	0.54225
	1.243046	0.746116	1.179764	0.58714

Figura 3.7: *GeneArEsce*

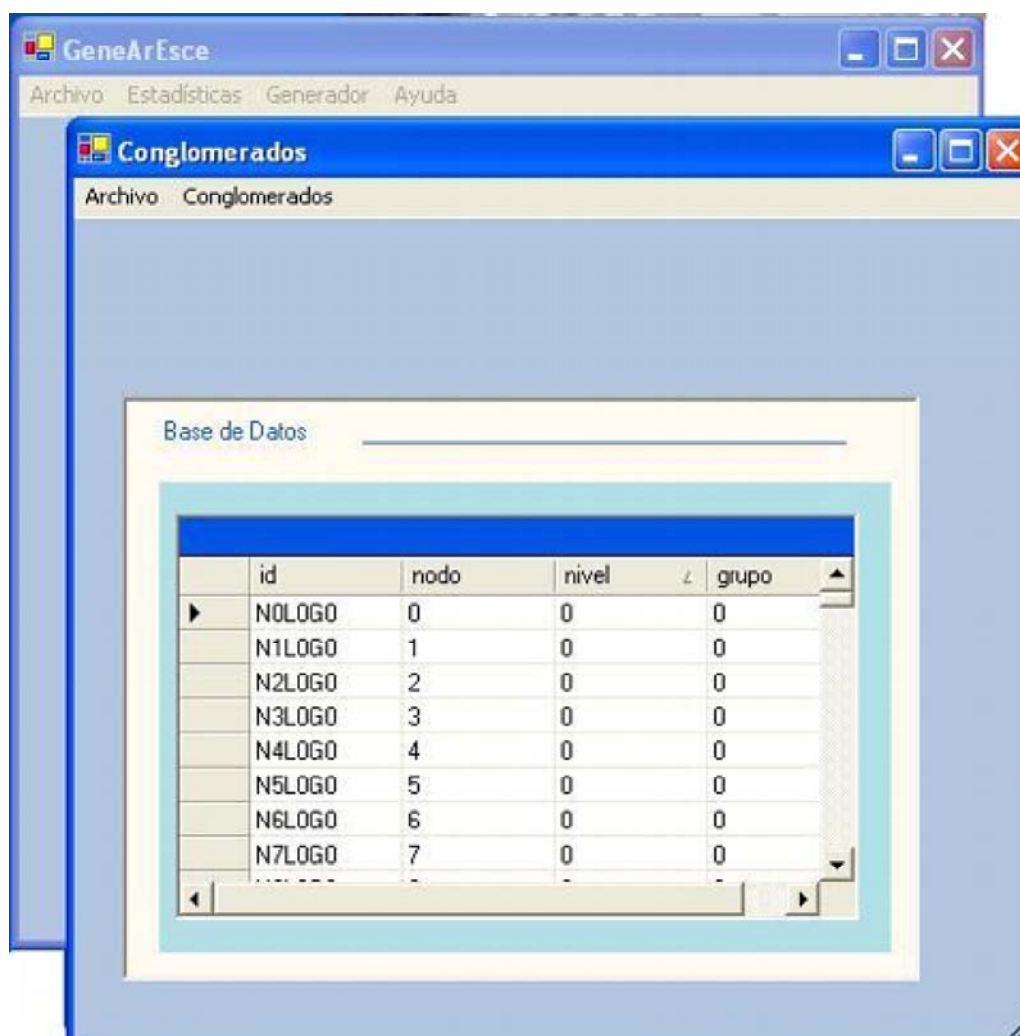
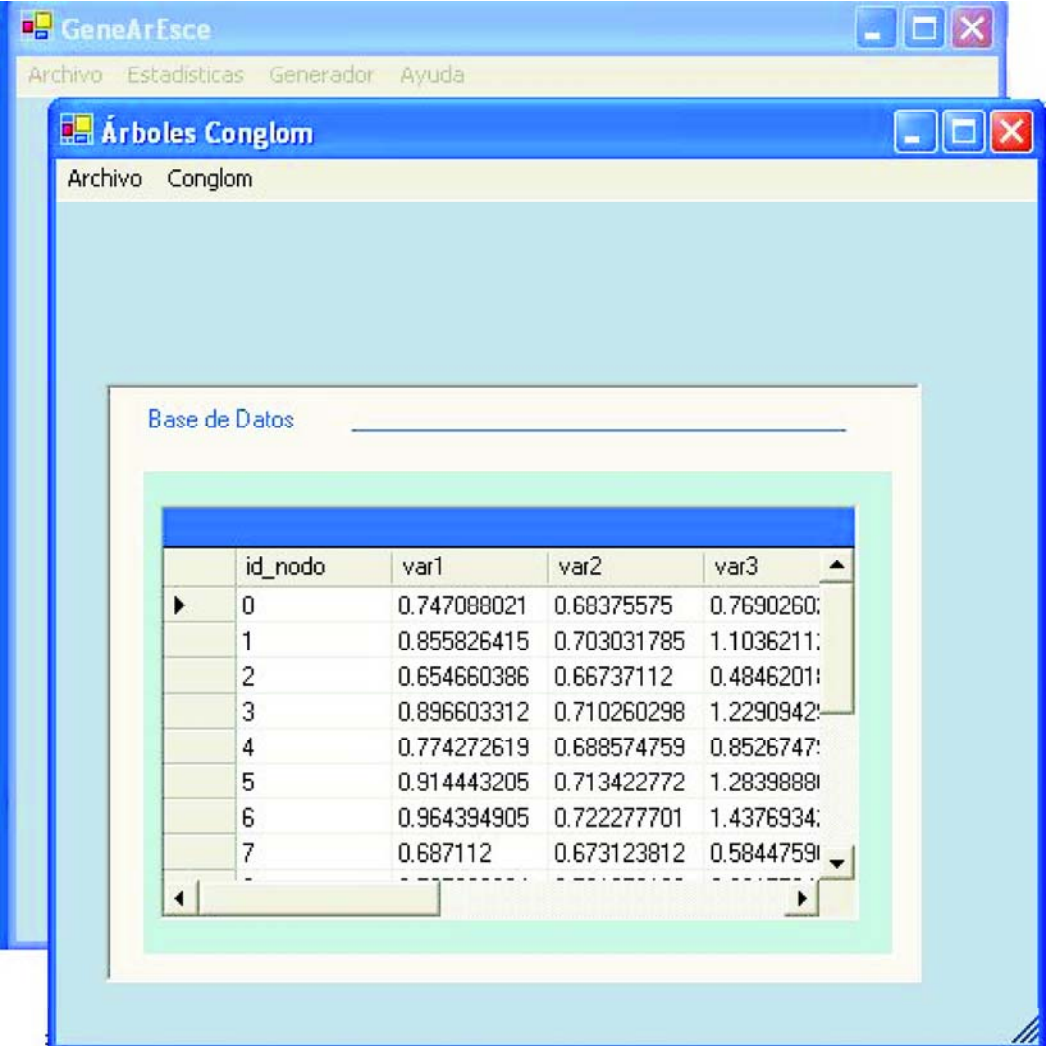


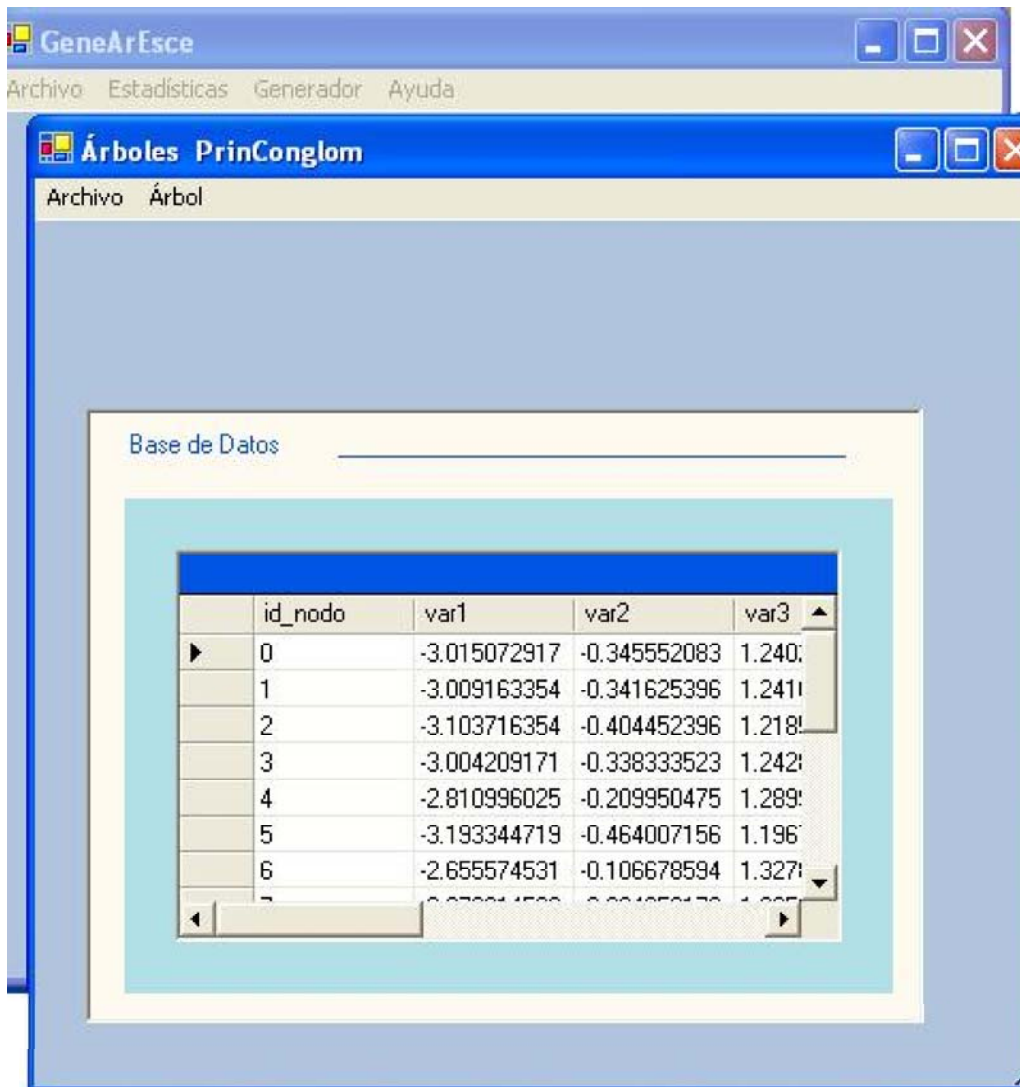
Figura 3.8: Conglomerados en GeneArEsce



The image shows a screenshot of the GeneArEsce software interface. The main window is titled "GeneArEsce" and has a menu bar with "Archivo", "Estadísticas", "Generador", and "Ayuda". A secondary window titled "Árboles Conglom" is open, showing a menu bar with "Archivo" and "Conglom". Inside this window, there is a "Base de Datos" section containing a data table with the following columns: "id_nodo", "var1", "var2", and "var3". The table contains 8 rows of data, with the last row partially cut off.

	id_nodo	var1	var2	var3
▶	0	0.747088021	0.68375575	0.7690260:
	1	0.855826415	0.703031785	1.1036211:
	2	0.654660386	0.66737112	0.4846201:
	3	0.896603312	0.710260298	1.2290942:
	4	0.774272619	0.688574759	0.8526747:
	5	0.914443205	0.713422772	1.2839888:
	6	0.964394905	0.722277701	1.4376934:
	7	0.687112	0.673123812	0.5844759:

Figura 3.9: *Arboles con Conglomerados GeneArEsce*



Arboles con Conglomerados y Componentes Principales en GeneArEsce

3.4 Algoritmos de GeneArEsce

Una vez que se diseñan las clases se procedió a la implementación de los algoritmos fundamentales de la Aplicación GeneArEsce y se presentan a continuación:

- **Encadenamiento Simple** : Procedimiento que genera el criterio de distancia de Encadenamiento Simple entre grupos

- **Conglomerate** : Procedimiento que genera Dendrogramas aplicando Conglomeración Jerárquica Aglomerativa
- **PrinC**: Procedimiento que genera las Componentes Principales
- **GenMuLog**: Procedimiento que genera muestras aleatorias lognormales multivariadas
- **Conglom** : Procedimiento que genera Árboles de Escenario mediante Conglomerados
- **PrinConglom**: Procedimiento que genera Árboles de Escenario mediante Conglomerados y Componentes Principales
- **PrinCEs**: Procedimiento que genera Escenarios mediante Componentes Principales

Algorithm 1 Encadenamiento Simple

Ordenar las distancias en orden ascendente
 Buscar los candidatos con la distancia mínima
Si $C_i \in \text{ConglomCompuesto}$ **Entonces**
 Si $C_j \in \text{ConglomCompuesto}$ **Entonces**
 Fusionar ambos Conglomerados Compuestos
 Sino
 Fusionar C_j en el conglomerado C_i
 Fin Si
Sino $\{C_i \text{ es } \text{ConglomSimple}\}$
 Si $C_j \in \text{ConglomCompuesto}$ **Entonces**
 Fusionar C_i en el conglomerado C_j
 Sino
 Fusionar ambos candidatos en un nuevo Conglomerado Compuesto
 Fin Si
Fin Si
 Redimensionar $M_x GN$

Algorithm 2 Conglomerate

Importar la Matriz de Datos
 Obtener la Matriz de Distancias {Euclidianas}
 Generar el nivel $N-1$ con N grupos
 Actualizar los registros en el *Dendrograma*
 Construir el primer *Conglomerado Compuesto*
 Considerar la menor distancia de los grupos
 Actualizar los registros en el *Dendrograma*
 $Nivel = Nivel - 1$
Para $Nivel = N - 3$ a 0 **Hacer**
 Formar un nuevo Conglomerado aplicando criterio {Encadenamiento}
 Actualizar los registros en el *Dendrograma*
Fin Para

Algorithm 3 PrinC

Importar la Matriz de Datos
 Transformar la Matriz de Datos
 Restar las medias de cada variable a la Matriz Original
 Obtener:
 Matriz de Varianzas y Covarianzas
 Matriz de Correlaciones
 Valores Propios {Householder y QR}
Para cada valor propio **Hacer**
 Generar el vector propio correspondiente {Gauss con sustitución hacia atrás}
Fin Para
 Obtener datos resultantes
 Matriz de las *Componentes Principales*
 Matriz de los *Scores*
 Desviaciones Estándar indicando la importancia de las *Componentes Principales*

Algorithm 4 GenMuLog

Importar la Matriz de Datos
 Transformar la Matriz de Varianzas y Covarianzas a parametros normales
 Obtener el Vector de Medias μ
 Obtener la Matriz de Varianzas y Covarianzas Σ
 Aplicar la Factorización de Cholesky a Σ
 Guardar el resultado en L^T
 Generar $Z = (z_1, z_2, z_3, \dots, z_p)^T$ donde $z_i < i < p$ son variables aleatorias $N(0, 1)$
 Matriz de Correlaciones
 Aplicar las transformaciones de Box-Muller
 $x = \sqrt{-2\ln U_2} \cos(2\pi U_1)$ $y = \sqrt{-2\ln U_2} \sin(2\pi U_1)$
 Obtener $X = L^T * Z + \mu$
 Aplicar la función exponencial a X

Algorithm 5 Conglom

Importar la Matriz de Datos
 Construir el nivel 0 cuyo Nodo Principal contiene los valores de la Base de Datos inicial
 Registro=($id_nodo = 0, var_1, \dots, var_n, prob = 1, id_padre = -1, nivel = 0$)
 Generar el Dendrograma a partir de la Matriz de Datos
 Elegir el número de grupos para determinar el número de ramas por nivel que conformarán el Árbol de Escenario
 Construir una Base de Datos para cada grupo seleccionado
Para $Nivel = 1$ a $L - 1$ **Hacer**
 Generar una Base de Datos Lognormal para cada grupo
 Calcular la probabilidad de ocurrencia de cada rama
 Para $grupo = 1$ a $TotalG$ **Hacer**
 Actualizar los registros en la base T.Conglom
 ($id_nodo, var_1(grupo), \dots, var_n(grupo), prob(grupo), id_padre(grupo), nivel$)
 Fin Para
Fin Para

Algorithm 6 PrinConglom

Importar la Matriz de Datos
 Generar la Base de Datos CP con las Componentes Principales de la Base de Datos original
 Construir el nivel 0 cuyo Nodo Principal contiene los valores de la Base de Datos CP
 $(id_nodo = 0, varCP_1, \dots, varCP_n, prob = 1, id_padre = -1, nivel = 0)$
 Generar el Dendrograma a partir de CP
 Elegir el número de grupos para determinar el número de ramas por nivel que conformarán el Árbol de Escenario
 Construir una Base de Datos para cada grupo seleccionado
Para $Nivel = 1$ a $L - 1$ **Hacer**
 Generar una Base de Datos Lognormal para cada grupo
 Calcular la probabilidad de ocurrencia de cada rama
 Para $grupo = 1$ a $TotalG$ **Hacer**
 Actualizar los registros en la base T.PrinConglomCP
 $(id_nodo, varCP_1(grupo), \dots, varCP_n(grupo), prob(grupo), id_padre(grupo), nivel)$
 Transformar las variables VarCP a las variables originales
 Actualizar los registros en la base T.PrinConglom
 $(id_nodo, var_1(grupo), \dots, var_n(grupo), prob(grupo), id_padre(grupo), nivel)$
 Fin Para
Fin Para

Algorithm 7 PrinCEs

Obtener las Componentes Principales
Si la Primera Componente posee la mayor proporción acumulada **Entonces**
 Ordenar los scores resultantes
 Obtener los percentiles
 Escoger los percentiles de los eventos cota de la distribución empírica de la Componente Principal
Sino {Si se requiere más de una Componente Principal}
 Se realiza el proceso anterior para cada una de las Componentes significativas
 Se genera una variable como combinación lineal de las Componentes
 Ordenar los scores resultantes
 Obtener los percentiles
 Escoger los percentiles de los eventos cota de la distribución empírica de la Componente Principal
Fin Si

3.4.1 Householder, QR y Cholesky

Estos algoritmos se encuentran sustentados en [Bur03], por lo que sólo se mencionará sus funcionalidades esenciales y el motivo de su implementación en GeneArEsce:

3.4.1.1 Householder

El objetivo de este método es reducir una Matriz Simétrica arbitraria a una Matriz Tridiagonal Simétrica. Esto se realiza con las Transformaciones de Householder que suprimen selectivamente bloques de columnas de matrices logrando estabilidad respecto al error de redondeo.

Este método es indispensable en la obtención de valores propios para la elaboración del algoritmo PrinC que genera las Componentes Principales.

3.4.1.2 QR

El objetivo de este método es determinar simultáneamente todos los valores característicos de una Matriz Tridiagonal Simétrica. El éxito del procedimiento radica en la conversión de la matriz original a una matriz con los valores propios en la diagonal, empleando matrices de rotación.

Este método es indispensable en la obtención de valores propios para la elaboración del algoritmo PrinC que genera las Componentes Principales.

3.4.1.3 Cholesky

El objetivo de este método es factorizar una matriz definida positiva cuadrada en una matriz de la forma LL^T donde L es una Matriz Triangular Inferior.

Este método se utiliza en la descomposición de la Matriz de Varianzas y Covarianzas, para el algoritmo GenMuLog.

Por lo que la Aplicación GeneArEsce queda conformada con estas funcionalidades, sólo resta probar su funcionalidad en un campo de Investigación específico como son las Carteras de Inversión.

Capítulo 4

GeneArEsce en Carteras de Inversión

Con las funcionalidades proporcionadas por GeneArEsce se procedió a la generación de Árboles de Escenario con la Base de Datos histórica obtenida en Economatica.

4.1 Validación de Resultados

El proceso de validación de resultados quedo conformado con las siguientes etapas:

- **Comprobación de la correspondencia de la simulación con la realidad**
- **Variabilidad de los Resultados de los periodos del Árbol de Escenario**
- **Resultados del comportamiento del Árbol de Escenario con los métodos de generación de Árboles de Escenario**

4.1.1 Comprobación de la correspondencia de la simulación con la realidad

Una vez obtenidas las Bases de Datos simuladas con el Generador Multivariado Lognormal de GeneArEsce, es necesario verificar que la información simulada presente una relación de correspondencia significativa con la Base de Datos recopilada mediante Economatica.

Los resultados aparecen en la Tabla 4.1, donde se muestra la aplicación de tres criterios estadísticos mediante el procedimiento MANOVA por la plataforma Minitab. Cabe destacar que uno de los activos fue eliminado durante este análisis debido a la alta correlación que presentaba con otro activo.

Como se observa en la segunda columna de la Tabla 4.1 aparecen los valores de los estadísticos Wilks, LawleyHotelling y Pillai, mientras que en la tercera columna aparece el valor de la Prueba F que aproxima a los estadísticos mencionados en la tabla.

Los valores P que aparecen en la sexta columna, indican que la hipótesis nula (existe correspondencia de la base simulada con la base histórica mediante el análisis de sus vectores de medias), no es rechazada, por lo cuál se concluye que la simulación corresponde a la realidad.

Criterio	Estadístico	F	Numerador	Denominador	P
Wilks	0.69995	0.402	49	46	0.999
LawleyHotelling	0.42867	0.402	49	46	0.999
Pillai	0.30005	0.402	49	46	0.999

Tabla 4.1: Resultados de la correspondencia de los datos simulados con los datos de Economatica

4.1.2 Variabilidad de los Resultados de los periodos del Árbol de Escenario

En este caso el estudio se basó en estudiar la correspondencia de la variabilidad de cada uno de los activos entre los periodos y dentro de cada uno de ellos. Para este estudio se aplicó la Prueba de igualdad de Varianzas iguales proporcionado por Minitab. Se realizaron pruebas para los métodos de Conglomerados y Conglomerados con Componentes Principales.

4.1.3 Resultados de los periodos con el método de Conglomerados

4.1.3.1 Variabilidad entre los periodos

La figura 4.1 muestra el comportamiento de los activos simulados del segundo periodo respecto a los activos históricos del primer periodo. Como puede observarse en la primera gráfica se observa la correspondencia de las varianzas entre las muestras, este resultado es sustentado tanto por la Prueba F así como la Prueba de Levene. Por otro lado los intervalos de confianza de Bonferroni también muestran correspondencia de los datos debido a que en la primera gráfica las dos líneas se traslapan.

La tabla 4.2 muestra el concentrado de los datos de cada uno de los nodos de los periodos 1-3 respecto al periodo 0 donde se encuentran los datos históricos. Observé que en todos los casos los valores de p son mayores que el nivel de significancia $\alpha = 0.05$, por lo que no se rechaza la hipótesis nula de varianzas iguales.

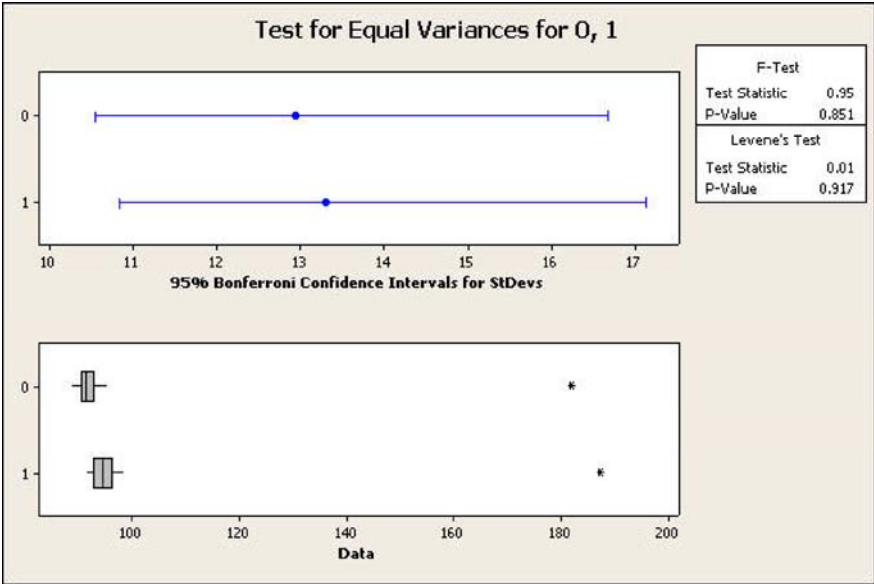


Figura 4.1: Variabilidad entre los periodos

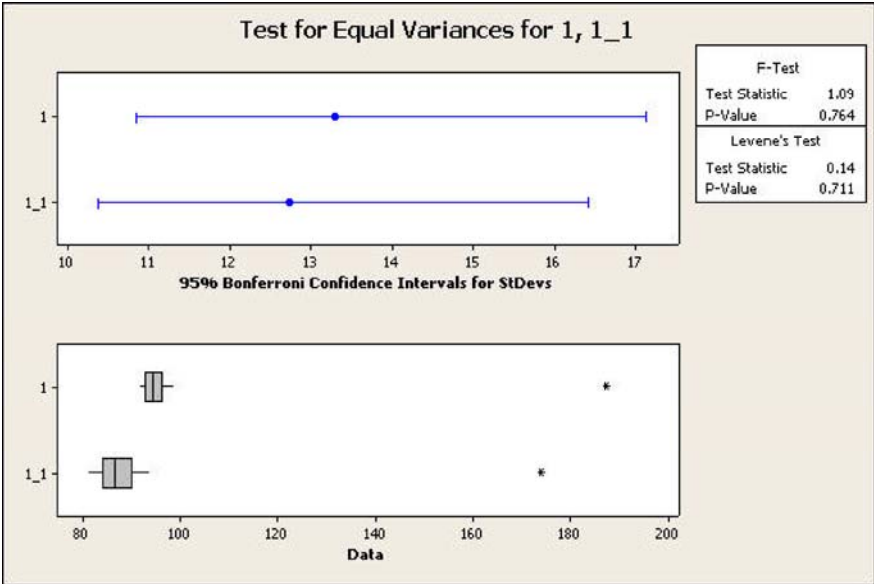


Figura 4.2: Variabilidad dentro de los periodos

periodo	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
1	0.95	0.851	0.01	0.917
1	1.03	0.910	0.23	0.630
2	0.93	0.808	0.02	0.900
2	0.95	0.852	0.03	0.852
2	1.09	0.771	0.26	0.611
2	0.96	0.899	0.59	0.443
3	0.94	0.826	0.10	0.749
3	0.93	0.796	0.28	0.595
3	0.88	0.648	0.18	0.675
3	0.91	0.756	0.07	0.785
3	1.13	0.675	0.18	0.670
3	0.97	0.911	0.71	0.403
3	0.95	0.864	0.61	0.438
3	0.88	0.664	0.79	0.376

Tabla 4.2: Resultados de la varianza del periodo 0 con los periodos sucesivos

4.1.3.2 Variabilidad dentro de los periodos

En este caso se analizó la varianza entre los nodos hijos en cada uno de los periodos generados. En la figura 4.2 el análisis es similar al de la figura 4.1, sólo que ahora se valora el comportamiento de los activos simulados dentro del primer periodo, es decir los nodos 1 y 2 del Árbol de Escenario. Como en el caso anterior la correspondencia de las varianzas entre ellos fue aceptada por las Pruebas F y Levene y por los intervalos de confianza de Bonferroni.

En las tablas 4.3, 4.4 y 4.5 aparecen los resultados concentrados de la comparación de varianzas tanto por la Prueba F como la de Levene que no rechazaron la hipótesis nula: existe correspondencia de varianzas entre los nodos hijos de cada periodo.

nodos	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
1-2	1.09	0.764	0.14	0.711

Tabla 4.3: Resultados de la varianza entre los nodos del periodo 1

nodos	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
3-4	1.02	0.955	0.00	0.953
5-6	0.89	0.676	0.08	0.781

Tabla 4.4: Resultados de la varianza entre los nodos del periodo 2

nodos	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
7-8	0.99	0.970	0.04	0.834
9-10	1.04	0.884	0.02	0.884
11-12	0.86	0.595	0.20	0.660
13-14	0.93	0.793	0.02	0.901

Tabla 4.5: Resultados de la varianza entre los nodos del periodo 3

4.1.4 Resultados de los periodos con el método de Conglomerados y Componentes Principales

4.1.4.1 Variabilidad entre los periodos

El procedimiento fue el mismo que se aplicó al método de Conglomerados, la tabla 4.6 muestra el concentrado de los datos de cada uno de los nodos de los periodos 1-3 respecto al periodo 0 donde se encuentran los datos históricos.

periodo	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
1	0.93	0.804	0.01	0.941
1	1.04	0.899	0.06	0.806
2	0.89	0.676	0.01	0.933
2	0.96	0.886	0.01	0.940
2	1.03	0.913	0.33	0.568
2	0.98	0.940	0.23	0.631
3	0.88	0.667	0.07	0.788
3	0.87	0.618	0.35	0.553
3	0.91	0.742	0.05	0.816
3	0.88	0.666	0.06	0.812
3	1.04	0.881	0.23	0.633
3	0.96	0.892	0.92	0.340
3	0.94	0.818	0.39	0.534
3	0.88	0.666	0.66	0.419

Tabla 4.6: Resultados de la varianza del periodo 0 con los periodos sucesivos

4.1.4.2 Variabilidad dentro de los periodos

Como en el caso de variabilidad dentro de los periodos con el método de conglomerados, los resultados de la comparación de varianzas tanto por la Prueba F como la de Levene no rechazaron la hipótesis nula: existe correspondencia de varianzas entre los nodos hijos de cada periodo, los concentrados por cada periodo se encuentran en las tablas 4.7, 4.8 y 4.9.

nodos	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
1-2	1.11	0.708	0.03	0.867

Tabla 4.7: Resultados de la varianza entre los nodos del periodo 1

nodos	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
3-4	1.08	0.784	0.00	0.991
5-6	0.95	0.854	0.01	0.934

Tabla 4.8: Resultados de la varianza entre los nodos del periodo 2

nodos	Estadístico F	p empírico F	Estadístico Levene	p empírico Levene
7-8	0.98	0.945	0.10	0.753
9-10	0.97	0.919	0.00	0.995
11-12	0.92	0.775	0.25	0.620
13-14	0.94	0.840	0.04	0.847

Tabla 4.9: Resultados de la varianza entre los nodos del periodo 3

4.1.5 Resultados del comportamiento del Árbol de Escenario con los métodos de generación de Árboles de Escenario

Una vez establecida la correspondencia de los datos simulados con la realidad, así como los resultados de los métodos de generación de Árboles de Escenario, se procedió a la realización de un análisis descriptivo que contemplará el comportamiento de los datos resultantes con ambos métodos propuestos.

El criterio estudiado en este caso, fue el comportamiento de la probabilidad en cada uno de los nodos de los Árboles resultantes de ambos métodos, puesto que es factible comparar la probabilidad de cada uno de los nodos de los árboles.

En la figura 4.3 aparece el comportamiento de la probabilidad de cada uno de los nodos de Árboles sometidos a Conglomerados y Conglomerados con Componentes Principales.

Como se observa en la mayoría de los nodos (Eje X), la probabilidad (Eje Y) arrojada por cada método mantiene una proporcionalidad similar, excepto en los nodos 4 y 12 donde los picos son demasiado pronunciados.

en la figura 4.4 se aprecian cada una de las probabilidades de los nodos en cada gráfica, en este caso se observa que el método de Conglomerados con Componentes Principales arroja probabilidades que tienden a los extremos del intervalo (0,1), en cambio el método de Conglomerados mantiene a las probabilidades en el rango medio del intervalo (0,1). Cabe destacar que los nodos que se encuentran inmersos en el intervalo (5-10), presentan un comportamiento distinto al grupo de nodos, puesto que mientras que en la gráfica probC(conglomerados)

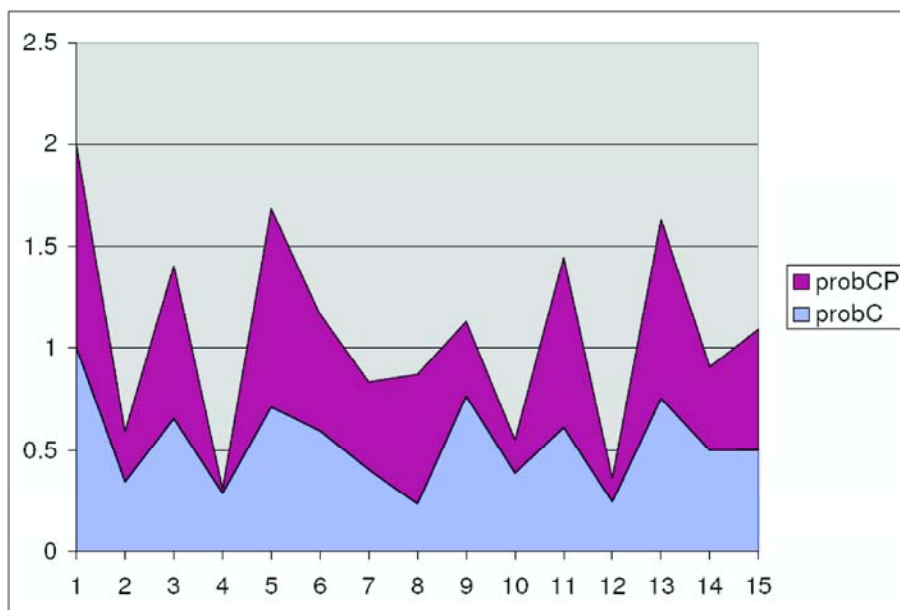


Figura 4.3: Probabilidades de los nodos en los mtodos de generación de Árboles con tendencias de comportamiento

la probabilidad disminuye, en la gráfica probCP(conglomerados-componentes) la probabilidad aumenta.

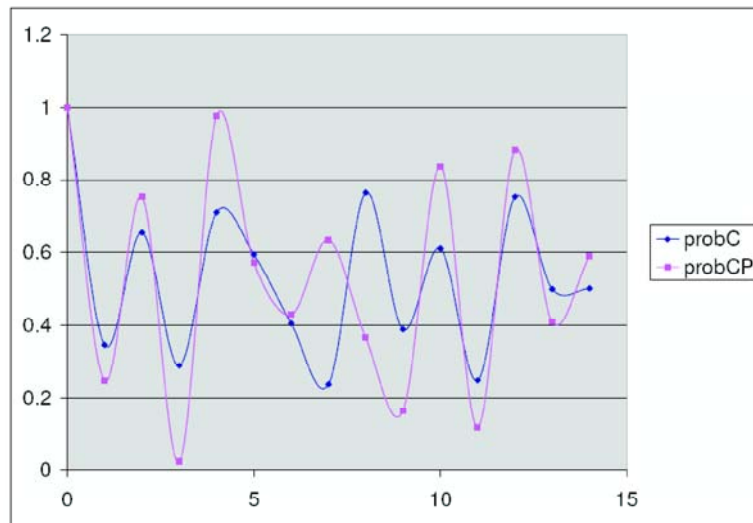


Figura 4.4: Probabilidades de los nodos en los métodos de generación de Árboles con valores específicos

Capítulo 5

Conclusiones

Los métodos de generación de Árboles de Escenario, son una herramienta potencial para satisfacer las necesidades de modelación de incertidumbre dado un horizonte de planeación específico. La puesta en marcha de estos métodos implica la comprensión y manejo de la Estadística, el Análisis Financiero, el tratamiento de la información, las Carteras de Inversión y los paradigmas de Ingeniería de Software, de forma que al fusionar estos elementos se obtiene una perspectiva clara de que pasos seguir en la generación de Árboles de Escenario.

Como aportaciones de esta investigación se cuentan con metodologías de generación de Árboles de Escenario, de implementación de software y estadísticas para el análisis de y procesamiento de datos financieros.

Las metodologías de generación de los Árboles Financieros son sustentados, por el análisis multivariante con sus procedimientos de conglomerados y componentes principales, que ocasionaron tres métodos, dos para la generación de Árboles, Conglomerados y Conglomerados-Componentes Principales mientras que el tercero es una alternativa para generar Escenarios sencillos sin considerar la Arborescencia.

Como metodología de implementación de software se empleó el modelo en Espiral, abarcando no sólo cuestiones de implementación sino también aplicándose al desarrollo por etapas de la investigación.

Las metodologías estadísticas facilitaron el análisis y la interpretación de la información, en etapas iniciales del proyecto, se aplicó el criterio de las componentes principales para deducir los activos más importantes, se arrojaron datos inconsistentes respecto a la distribución de los tres tipos de activos estudiados y se analizaron los datos extremos de comportamiento mediante el criterio de conglomeración. En la etapa final del proyecto estas metodologías facilitaron la interpretación de los resultados mediante la técnica MANOVA se verificó la correspondencia de la información simulado respecto a la real y con las pruebas estadísticas F, Levene, se sustentaron los criterios de correspondencia de las varianzas entre los periodos del Árbol y dentro de los nodos por periodo del mismo.

Los resultados de este proyecto incluyen la aplicación GeneArEsce que pro-

porciona herramientas estadísticas como la generación de matrices de covarianzas, distancias, Dendrogramas, Componentes Principales y con estos elementos constituye opciones más complejas como la generación de Árboles de Escenario con los métodos de Conglomerados y Conglomerados con Componentes Principales.

Los resultados sobre el análisis de los datos provenientes de Economatica muestran que es preferible conservar los activos como CETES y Acciones y restar importancia a los Bonos dado que el análisis de Componentes Principales mostró que son los activos menos significativos, aunque se incluyeron en la investigación con el fin de contar con activos de distintos tipos de riesgo.

Los resultados numéricos arrojaron Bases de Datos simuladas que corresponden a la realidad, correspondencia de las varianzas de cada uno de los nodos en cada periodo y entre cada uno de ellos, en el caso del método de Conglomerados, los Árboles de Escenario tiene un comportamiento más suave dada la probabilidad de ocurrencia de los valores de sus nodos, mientras que en el caso del método de Conglomerados con Componentes Principales la probabilidad de ocurrencia de los valores de los nodos tiende a buscar los extremos.

Como perspectivas futuras se sugiere utilizar Cadenas de Markov para la modelación de los Escenarios, utilizar las metodologías implementadas en otros campos de la ciencia como el análisis de suelos, educación y toda área donde se provea de grandes volúmenes de datos que necesiten procesarse para analizar su comportamiento en el tiempo.

Bibliografía

- [Bur03] R. Burden, D. Faires, Análisis Numérico, Thomson, 2003
- [Ale03] G. Alexander, W. Sharpe, J. Bailey. Fundamentos de Inversiones Teoría y Práctica, Prentice Hall, 2003.
- [Eco05] Manual de Referencia de Economática®, 2005, disponible en <http://www.economica.com>
- [Gul01] N. Gulpinar, B. Rustem y R. Settergren. Optimisation and simulation approaches to scenario tree generation en Journal of Economics Dynamics and Control, 2002.
- [Hoy01] K. Hoyland, S. W. Wallace. Generating scenario trees for multi stage decision problems en Management Science, volume 47, 2001, pags 295-307.
- [Kau03] K. M. Kaut, S. W. Wallace. Evaluation of scenario-generation methods for stochastic programming en Stochastic Programming, E-Print Series, 2003.
- [Kou01] R. Kouwenberg. Scenario generation and stochastic programming models for asset liability en management. European Journal of Operation Research, volume 134, 2001, pags 279-292.
- [Lin90] G. Linares. Análisis de Datos, ENPES, 1990.
- [Lor97] M. Loretan. Generating market risk scenarios using principal components analysis: Methodological and practical considerations. The Measurement of Aggregate Market Risk , CGFS Publications No. 7, Nov. 1997, pags 23-60.
- [Oso04] M. A. Osorio, R. Settergren, B. Rustem, N. Gulpinar. Post-Tax Optimization with Stochastic Programming en European Journal of Operational Research, volumen 157, 2004, pags 152-168.
- [Oso05] M. A. Osorio, A. Sánchez, Designing Decision Support Systems for Portfolio Management, Reporte Técnico, Facultad de Ciencias de la Computación, Universidad Autónoma de Puebla, 2005.

- [Pus98] J-F Puzraszeri. Portfolio Optimization under Uncertainty, Part 11, 1998
- [Rey04] A. Reynoso, Opening up a Securities Market: Mexico's New Push for Liberalization 2003-2004 en Mexican Stock Exchange, SCID Latin America Conference, 2004.
- [Tre04] A. Trejos, P. Carbajal, M. Barrera. Generador de Procesos Normales Multivariados, en Scientia et Técnica, No 25, 2004, pags 185-189.
- [Ven99] W. N. Venables, B.D. Ripley. Modern Applied Statistics with S-PLUS. tercera edición. Estados Unidos, Nueva York: Springer Verlag, 1999.
- [Wan98] S. Wang, Aggregation of Correlated Risk Portfolios: Models and Algorithms , Proceedings of the Casualty Actuarial Society, vol 85, 1998.