



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA

Uso de Métodos Multivariados en el Estudio de Riesgo de Portafolios de Markowitz y Sharpe.

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela por el **Br. Juan Malca** para optar al título de Licenciado en Matemática.

Tutor: Dr. José Benito Hernández.

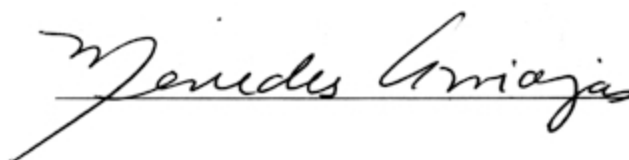
Caracas, Venezuela
Noviembre 2016

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado "Uso de Métodos Multivariados en el Estudio de Riesgo de Portafolios de Markowitz y Sharpe.", presentado por el Br. Juan Malca, titular de la Cédula de Identidad V 24.461.664, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de Licenciado en Matemática.



Dr. José Benito Hernández.

Tutor



Dra. Mercedes Arriojas.

Jurado



Dra. Mairene Colina.

Jurado

Dedicatoria

Dedico este Trabajo Especial de Grado a mis padres, Juan Malca y Luisa Atencio de Malca, ejemplos en mi vida. Su amor y lucha han hecho de ellos el gran ejemplo a para mí. Los amo con todo mi corazón.

Agradecimiento

Mi sincero agradecimiento está dirigido a todas aquellas personas que creyeron en mí y contribuyeron en la realización de este trabajo y especialmente a:

La ilustre Universidad Central de Venezuela, por ser más que un recinto universitario, mi segundo hogar en estos años de carrera, donde pude alcanzar esta meta y crecer como persona.

Mi tutor Dr. José Benito, por su disponibilidad, su dedicación, sus sugerencias oportunas y por toda su colaboración para que este trabajo fuese culminado.

Los integrantes de mi terna examinadora, por ser partícipes en la realización de este trabajo. En especial a la Dra. Mercedes Arriojas, por su disponibilidad, su exigencia en mi persona y su puntualidad.

Mis padres, Juan Malca y Luisa Atencio de Malca, quienes a lo largo de mi vida han velado por mi bienestar y educación siendo mi apoyo en todo momento, depositando su entera confianza en cada reto que se me presentaba sin dudar ni un solo instante en mi capacidad de superación y perseverancia.

Índice general

1. Preliminares	3
1.1. Nociones Básicas de Probabilidades.	3
1.2. Nociones Básicas de Estadística.	7
1.2.1. Inferencias sobre la población a partir de la información de la muestra.	9
2. Análisis Multivariante	12
2.0.2. Matrices y vectores de datos	12
2.0.3. Distribución Normal Multivariada	16
2.0.4. Independencia de las unidades experimentales.	16
2.0.5. Vectores de medias y matrices de varianzas-covarianzas.	17
2.0.6. Correlaciones y matrices de correlación	18
2.0.7. Pruebas estadísticas e intervalos de confianza para correlaciones.	21
2.0.8. Agrupamiento de datos basados en correlaciones.	23
2.0.9. La función de densidad de probabilidad normal multivariada.	24
2.0.10. <i>Estimaciones insesgadas.</i>	24
2.0.11. Datos estandarizados y/o valores Z	26
2.1. Análisis Univariado de Varianza.	27
2.1.1. Áreas de aplicación.	27
2.2. ANOVA de un solo factor	28
3. Portafolios de inversión de Markowitz.	31
3.1. Conceptos Financieros.	31
3.2. Portafolios Financieros.	33

3.2.1. Portafolio de Markowitz.	33
3.2.2. Portafolio CAPM	37
3.2.3. Índice de Sharpe.	39
3.2.4. Índice del Precio al Consumidor(IPC)	40
3.3. Práctica de portafolio de Markowitz.	41
3.3.1. Creación de Portafolios de Markowitz y frontera eficiente.	41
3.4. Práctica de ANOVA a los portafolios Min Var y Sharpe.	46
4. Conclusión	50
Bibliografía	51

Introducción

Este trabajo fue realizado con la finalidad de analizar portafolios de Mínima Varianza y Sharpe a partir de los precios de ventas de varias acciones tomadas de la Bolsa de Valores de Caracas. Los portafolios son elaborados usando el método tradicional de Markowitz y el propuesto por Sharpe, método CAMP(Capital Asset Pricing Model o Método de Fijación de Precios de Activos de Capital). Haremos comparaciones entre estos tipos de portafolios mediante el uso del ANOVA como método de análisis multivariado, para determinar el portafolio que mejor se adapta a nuestras necesidades, tener un criterio de cuan riesgoso puede llegar a ser nuestro portafolio.

En la práctica se presenta un modelo sencillo que está desarrollado en Excel, el cual permite obtener los portafolios que se encuentren en la llamada Frontera Eficiente cuyo fundamento teórico es el modelo de Markowitz el cual incluye desde la manera de medir el riesgo y rendimiento que tiene el conjunto de acciones de inversión.

Capítulo 1

Preliminares

En este capítulo presentaremos algunos conceptos básicos de Probabilidad como la definición de una variable aleatoria, su esperanza y varianza, y algunos tipos de función de densidad; los cuales van a tener un rol importante en este trabajo.

1.1. Nociones Básicas de Probabilidades.

Definición 1.1 Variable Aleatoria.

Sea $(\Omega, \mathfrak{F}, \mathbb{P})$ un espacio de probabilidad. Una variable aleatoria ó variable aleatoria real es una función:

$$X : \Omega \mapsto \mathbb{R}$$

tal que para todo intervalo $I \subseteq \mathbb{R}$ el conjunto

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathfrak{F}.$$

Definición 1.2 Variable Aleatoria Discreta.

Una variable aleatoria X se denomina discreta, si su rango es un conjunto discreto (finito ó numerable). En este caso existe un conjunto $\{x_n\}_{n \geq 1}$ (conjunto de valores de X) tal que:

$$\sum_{n \geq 1} \mathbb{P}(X = x_n) = 1.$$

Definición 1.3 Función de Probabilidad.

Dada una variable aleatoria discreta X cuyo conjunto de valores es un conjunto $\{x_n\}_{n \geq 1}$, la función:

$$p : \{x_n\}_{n \geq 1} \mapsto [0, 1]$$

$$x_n \rightarrow p(n) = p_n = \mathbb{P}\{X = x_n\}$$

se denomina función de probabilidad de la variable aleatoria X ó función de masa de probabilidad de la variable aleatoria X .

Definición 1.4 Función de Distribución.

Si X es una variable aleatoria, se define la función de distribución de X como la función $F_X : \mathbb{R} \rightarrow [0, 1]$ tal que para cada $y \in \mathbb{R}$.

$$F_X(y) = \mathbb{P}(\{w \in \Omega : X(w) \leq y\}) = \mathbb{P}(X \leq y).$$

Definición 1.5 Función de Densidad

Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función de densidad sobre \mathbb{R} si y sólo si f satisface las siguientes condiciones:

1. Para todo $x \in \mathbb{R}$, $f(x) \geq 0$.
2. $\int_{-\infty}^{\infty} f(u) du = 1$.

Definición 1.6 Variable Aleatoria Continua

Una variable aleatoria X es continua ó absolutamente continua si existe una función de densidad f tal que para todo $a \in \mathbb{R}$:

$$F(a) = \mathbb{P}(X \leq a) = \int_{-\infty}^a f(u) du.$$

Observación 1.7 Acerca de la función de densidad.

- La probabilidad de que X tome un valor en el intervalo (a, b) es el área bajo la curva de la función de densidad sobre dicho intervalo como se muestra en la figura 1.1.

Entre los estadísticos más importantes asociados a las variables aleatorias, y que serán de mucha utilidad en este trabajo, se encuentran la media y la varianza cuyas definiciones daremos a continuación.

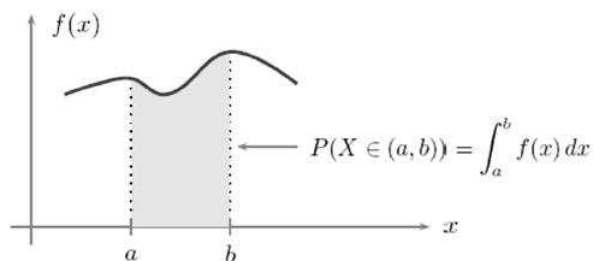


Figura 1.1: La probabilidad como área bajo la curva de la función de densidad.

Definición 1.8 Esperanza de una Variable Aleatoria Discreta.

Sea X una variable aleatoria discreta tal que su conjunto de valores es $\{x_n\}_{n \geq 1}$ y su función de probabilidad $\{p_n\}_{n \geq 1}$, donde $p_n = \mathbb{P}(X = x_n)$. La esperanza matemática, media ó valor esperado de X es el número denotado por $\mathbb{E}[X]$ y definido como:

$$\mu = \mathbb{E}[X] = \sum_{n=1}^{\infty} x_n \mathbb{P}(X = x_n) = \sum_{n=1}^{\infty} x_n p_n;$$

siempre y cuando la serie anterior sea absolutamente convergente. En este caso se dice que existe la esperanza matemática de X .

Definición 1.9 Esperanza de una Variable Aleatoria Continua

Sea X una variable aleatoria continua con función de densidad f . Se define la esperanza ó media de X como:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx;$$

siempre que la integral anterior sea finita. En este caso, $\mathbb{E}[X] < \infty$, se dice que X tiene esperanza o también que es integrable con respecto a la medida de probabilidad dada.

Definición 1.10 Varianza de una Variable Aleatoria.

Si X es una variable aleatoria con media $E[X] = \mu$, entonces la varianza de X , es denotada $\sigma^2[X]$, y se define como:

$$\sigma^2[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

La varianza de X es una medida de la dispersión de los valores de X alrededor de la media.

Definición 1.11 Desviación Típica o Estándar de una Variable Aleatoria.

Dada una variable aleatoria X con varianza σ^2 , la desviación típica o estándar de X es igual a σ .

Observación 1.12 La desviación típica o estándar es una medida de la dispersión de los valores de la variable aleatoria con respecto a su media, cuanto menor es la desviación estándar mayor será la concentración de los datos alrededor de la media, por lo que este parámetro describe apropiadamente la variable. Pero si los valores están ampliamente dispersos, entonces esta variabilidad podría eliminar cualquier información que parezca relevante que se desprende del promedio.

A continuación se presentan como ejemplo la distribución Normal y la distribución Chi cuadrado.

Ejemplo 1.13 Variable Aleatoria Normal.

Una variable aleatoria X se dice que tiene distribución normal ó que está normalmente distribuida, con parámetros μ = media y σ^2 = varianza ($\sigma > 0$), si su función de densidad de probabilidad está dada por:

$$N(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}; \forall x \in \mathbb{R}.$$

La función de distribución de una variable aleatoria normal es

$$F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t \exp^{-\frac{(x-\mu)^2}{2\sigma^2}} dx; \forall x \in \mathbb{R}.$$

y utilizamos la notación $X \sim N(x; \mu, \sigma)$ para decir que X está normalmente distribuida con parámetros μ y σ .

Definición 1.14 La distribución de una variable aleatoria normal con media 0 y varianza 1 se llama **distribución normal estándar**.

Definición 1.15 La función gamma se define como:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \text{ para } \alpha > 0$$

Definición 1.16 La variables aleatoria continua X tiene una distribución gamma, con parámetros α y β , si su función de densidad está dada por:

$$f(x, \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

donde $\alpha > 0$ y $\beta > 0$.

Un caso especial de la distribución gamma se obtiene al hacer $\alpha = v/2$ y $\beta = 2$, donde v es un entero positivo que corresponde a los grados de libertad. Esta distribución recibe el nombre de distribución Chi cuadrado.

Definición 1.17 La variable aleatoria continua X tiene una distribución chi cuadrado, con v grados de libertad, si su función de densidad está dada por:

$$f(x, v) = \begin{cases} \frac{1}{2^{v/2}\Gamma(v/2)} x^{v/2-1} e^{-x/2}, & x > 0 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

donde v es un entero positivo.

1.2. Nociones Básicas de Estadística.

Ahora veremos las características de una población y una muestra, y como se puede llegar a inferir a partir de la muestra, haciendo uso de los estadísticos, las características más relevantes de una población sin llegar a conocer todo el contenido de una población.

Definición 1.18 Población.

Una población consiste en la totalidad de las observaciones de las que estamos interesados, puede ser finito o infinito.

Definición 1.19 Muestreo aleatorio simple.

El muestreo aleatorio simple es un procedimiento que consiste en seleccionar n elementos de una población de tamaño N de tal modo que cada una de las muestras posibles tenga la misma oportunidad de ser elegida.

Definición 1.20 Estadístico

Cualquier función medible de las variables aleatorias que forman una muestra aleatoria se llama **estadístico**.

Por ejemplo, supongamos que deseamos llegar a una conclusión con respecto a la proporción de personas bebedoras de café en Venezuela que prefieren cierta marca de café. Sería imposible preguntar a cada bebedor de café venezolano para calcular el valor del parámetro p que representa la proporción de la población. En cambio, se selecciona una muestra aleatoria grande y se calcula la proporción \hat{p} de personas en esta muestra que prefieren la marca de café en cuestión. El valor \hat{p} se utiliza ahora para hacer una inferencia con respecto a la proporción p verdadera.

Ahora, \hat{p} es una función de los valores observados en la muestra aleatoria; como es posible tomar muchas muestras aleatorias de una misma población, esperaríamos que \hat{p} variara algo de una muestra a otra. Es decir, \hat{p} es un valor de una variable aleatoria que representamos con P . Tal variable aleatoria se llama estadístico.

Un estadístico sirve para estimar determinados parámetros de la distribución de la que procede la muestra, por ejemplo la media de los valores de una muestra, la varianza muestral, etc.

Definición 1.21 Distribución muestral

La distribución de un estadístico se llama **distribución muestral**.

Si, por ejemplo, el estadístico de que se trata es la media muestral, la distribución se conoce como *distribución muestral de medias* o la *distribución muestral de la media*. Análogamente se obtendrían las distribuciones muestrales de las desviaciones típicas, varianzas, medianas, proporciones, etc.

Para cada distribución muestral se puede calcular, la media, desviación típica, etc. Así, pues, se puede hablar de la media y la desviación típica de la distribución muestral de medias, etc.

Ejemplo 1.22 Estadísticos comunes.

- La media muestral aleatoria:

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n) = \frac{1}{n} \sum_{i=1}^n X_i,$$

es un estadístico utilizado para medir el centro de un conjunto de datos.

- La varianza muestral centrada o cuasivarianza:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

es un estadístico utilizado para medir la variabilidad o la dispersión de las observaciones.

- La desviación estándar de la muestra aleatoria: S , es un estadístico.

Por otra parte, si S^2 es la varianza muestral de tamaño n , podemos escribir

$$\begin{aligned} S^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right]. \end{aligned}$$

Al multiplicar el numerador y el denominador por n , obtenemos la fórmula

$$S^2 = \frac{1}{n(n-1)} \left[n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right].$$

1.2.1. Inferencias sobre la población a partir de la información de la muestra.

Distribuciones muestrales de la media.

Uno de los teoremas más importantes de probabilidad en el Teorema Central del Límite que nos permite aproximar o estimar la distribución de una muestral de \bar{X} será aproximadamente normal con media μ y varianza finita σ^2 cuando el tamaño de la muestra es grande.

Teorema 1.23 Teorema Central del Límite.

Si \bar{X} es la media de una muestra aleatoria de tamaño "n" tomada de una población con media μ y varianza finita σ^2 , entonces la forma límite de la distribución de:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

conforme $n \rightarrow \infty$, es la distribución normal estándar $N(z; 0, 1)$.

La aproximación normal para \bar{X} , por lo general, será buena si $n \geq 30$. Si $n < 30$, la aproximación es buena sólo si la población no es muy diferente de una distribución normal y, como se estableció antes, si se sabe que la población es normal, la distribución muestral de \bar{X} seguirá una distribución normal exacta, no importa qué tan pequeño sea el tamaño de las muestras.

Una aplicación importante del **teorema central del límite** consiste en determinar valores razonables de la media de la población μ .

Ejemplo 1.24 Un importante proceso de fabricación produce partes de componentes cilíndricos para la industria automotriz. Es importante que el proceso produzca partes que tengan una media de 5 milímetros. El ingeniero implicado hace la conjetura de que la media de la población es de 5 milímetros. Se lleva a cabo un experimento donde se selecciona al azar 100 partes elaboradas por el proceso y se mide el diámetro de cada una de ellas. Se sabe que la desviación estándar de la población es $\sigma = 0,1$. El experimento indica un diámetro promedio de la muestra $\bar{X} = 5,027$ milímetros. ¿Esta información de la muestra parece apoyar o refutar la conjetura del ingeniero?

Si la media μ es 5, ¿cuál es la probabilidad de que \bar{X} se desvíe a lo más 0.027 milímetros?

$$\begin{aligned} P(|\bar{X} - 5| \geq 0,027) &= P(\bar{X} - 5 \geq 0,027) + P(\bar{X} - 5 \leq -0,027) \\ &= 2P\left(\frac{\bar{X} - 5}{0,1\sqrt{100}} \geq 2,7\right). \end{aligned}$$

Al estandarizar \bar{X} de acuerdo con el teorema del límite central, se tiene:

$$2P\left(\frac{\bar{X} - 5}{0,1\sqrt{100}} \geq 2,7\right) = 2P(Z \geq 2,7) = 2(0,0035) = 0,007.$$

De esta manera, se experimenta una \bar{X} que está al menos de 0.027 milímetros de la media en aproximadamente 7 de cada 1000 experimentos. Como resultado, este experimento con $\bar{X} = 5,027$ ciertamente no ofrece evidencia que apoye la conjetura de que $\mu = 5,0$.

Distribuciones muestrales de S^2 .

Como hemos visto anteriormente se puede inferir acerca de la distribución muestral de \bar{X} usando el teorema central del límite, por otra parte se puede inferir acerca de su varianza muestral usando los siguientes teoremas y la función Chi cuadrado.

Teorema 1.25 Si X_1, X_2, \dots, X_n son variables mutuamente independiente que tienen, respectivamente, distribuciones chi cuadrado con v_1, v_2, \dots, v_n grados de libertad, entonces la variable aleatoria

$$Y = X_1 + X_2 + \dots + X_n$$

tienen una distribución chi cuadrada con $v = v_1 + v_2 + \dots + v_n$ grados de libertad.

Teorema 1.26 Si S^2 es la varianza muestral centrada de una muestra de tamaño n que se toma de una población normal que tiene varianza σ^2 , entonces el estadístico

$$X^2 = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}$$

tiene una distribución chi cuadrada con $v = n - 1$ grados de libertad.

Demostración. Si se extrae una muestra aleatoria de tamaño n de una población normal con media μ y varianza σ^2 , y se calcula la varianza muestral, obtenemos un valor del estadístico $(n-1)S^2/\sigma^2$.

Mediante la suma y la resta de la media muestral \bar{X} , es fácil ver que:

$$\begin{aligned} \sum_{i=1}^n (X_i - \mu)^2 &= \sum_{i=1}^n [(X_i - \bar{X}) + (\bar{X} - \mu)]^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + 2(\bar{X} - \mu) \sum_{i=1}^n (X_i - \bar{X}) \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2. \end{aligned}$$

Al dividir cada término de la igualdad entre σ^2 y sustituir $(n-1)S^2$ por $\sum_{i=1}^n (X_i - \bar{X})^2$, obtenemos

$$\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 = \frac{(n-1)S^2}{\sigma^2} + \frac{(\bar{X} - \mu)^2}{\sigma^2/n}.$$

Sabemos también que:

$$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}$$

es una variable aleatoria chi cuadrado con n grados de libertad. Al usar el teorema 1.25 tenemos una variable aleatoria chi cuadrado con n grados de libertad dividida en dos componentes. El segundo término del lado derecho es Z^2 , que es una variable aleatoria chi cuadrado con 1 grado de libertad y resulta que $(n-1)S^2/\sigma^2$ es una variable aleatoria chi cuadrado con $n-1$ grados de libertad. \square

Capítulo 2

Análisis Multivariante

El análisis multivariante es un método estadístico utilizado para determinar la contribución de varios factores en un simple evento o resultado.

La información multivariante viene dada en una matriz de datos, donde la información de entrada consiste en distancias o similaridades que miden el grado de discrepancia entre los individuos.

Componentes estudiados por el análisis multivariante:

- **Unidad experimental** es cualquier objeto o concepto que se puede medir o evaluar de alguna manera.
- **Las variables** o también llamadas variables respuesta, proporcionan datos que se obtienen siempre que un investigador mide o evalúa más de un atributo o característica de cada unidad experimental.

Los métodos multivariados son de gran utilidad para ayudar al investigador a hacer que tengan sentido conjuntos grandes, complicados y complejos de datos que constan de una gran cantidad de variables. Muchas técnicas multivariadas tienden a ser de naturaleza exploratoria en lugar de confirmatoria, es decir, muchos métodos multivariados tienden a motivar la hipótesis en lugar de comprobarla.

2.0.2. Matrices y vectores de datos

En este trabajo se usará p para representar la cantidad de variables numéricas de respuesta que se están midiendo y N siempre representará el número de unidades experimentales sobre las cuales se están midiendo

las variables. Se usa la variable X_{rj} para identificar el valor de la j -ésima variable respuesta en la r -ésima unidad experimental, para $r = 1, 2, \dots, N$ y $j = 1, 2, \dots, p$.

Definición 2.1 Matrices de Datos.

Sea X una matriz de datos, de modo que X_{rj} es el elemento en el r -ésima fila y la j -ésima columna de la propia matriz. La matriz de datos es Np y se denota por X . Por tanto,

$$X = \begin{pmatrix} X_{11} & \dots & X_{1j} & \dots & X_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{r1} & \dots & X_{rj} & \dots & X_{rp} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{N1} & \dots & X_{Nj} & \dots & X_{Np} \end{pmatrix}.$$

es una matriz de datos.

Definición 2.2 Vectores de datos.

Las filas de una matriz de datos se denominan vectores fila. Los datos que se presentan en la r -ésima fila de X se denominan por X'_r . De este modo,

$$X'_r = [X_{r1} \ X_{r2} \ \dots \ X_{rp}].$$

Cuando los datos de la r -ésima fila de X se escriben en un vector columna, se denota por X_r . En este caso,

$$X_r = \begin{bmatrix} X_{r1} \\ X_{r2} \\ \vdots \\ X_{rp} \end{bmatrix}.$$

Asimismo, la matriz de datos se expresa por:

$$X = \begin{bmatrix} X'_1 \\ X'_2 \\ \vdots \\ X'_N \end{bmatrix}.$$

Observación 2.3 *Algunas observaciones más resaltantes que veremos a lo largo del capítulo:*

1. *Con bastante frecuencia encontraremos la utilidad de crear nuevas variables para cada unidad experimental, de modo que se pueden comparar entre sí con más facilidad. Algunas de las técnicas multivariadas que crean nuevas variables son el análisis de componentes principales, el análisis por factores, el análisis de correlación canónica, el análisis discriminante canónico y el análisis de variables canónicas.*
2. *Datos Outliers: Son aquellas unidades experimentales cuyas medidas parecen incoherentes con las mediciones realizadas en las otras unidades experimentales. Por medio de procesos gráficos se pueden identificar estos datos outliers, entre ellas tenemos las gráficas de Andrews, las caras de Chernoff y las gráficas de rayos o de estrellas. Al ser localizados los datos outliers el investigador debe determinar si se cometió un error de registro, si así es, el error debe ser corregido inmediatamente; y si no es un error de registro, estos tipos de datos deben eliminarse de todo análisis multivariado con la finalidad de no afectar las conclusiones más veraces, es decir que la conclusión más acertada puede ser manipulada por estos datos outliers dando una idea errada o de poca relevancia en los datos.*
3. *Valores faltantes, para estos casos cuando faltan datos en la matriz de Datos nos encontramos con tres posibles acciones favorables: Reemplazo de los valores faltantes por ceros, este procedimiento trata de reemplazar los valores faltantes por ceros. De modo que, por ejemplo, en un caso hipotético donde se este midiendo la altura de un hombre que pesa 70 kilos, y su altura es reemplazada por cero, en este caso ya se ha cometido un grave error suponer tal aberración. Por lo que este tipo de reemplazo no es recomendado para ningún caso.*

Reemplazo de valores por promedios, a primera vista esta acción parece ser la mejor, pero no es muy recomendable, por que puede llegar a crear datos outliers. Pongamos un ejemplo: un investigador está observando una muestra aleatoria de personas provenientes de alguna población dada. Se consideran ahora dos variables de interés 1) el peso de cada persona, y 2) la estatura de cada persona. Suponga que también el peso medio de las personas de esta muestra sea de 58 kilos y que la estatura media sea 1.60 metros. Supongamos ahora que hay un integrante en la muestra cuya estatura es de 2.13 metros, pero cuyo peso no se conozca. Esta técnica para reemplazar los valores faltantes asignará un peso de 58 kilos como el peso reglamentario de este hombre. Una asignación de este tipo generaría un dato outlier en los datos que se van a analizar.

El otro método es la eliminación de filas de la matriz de datos, a pesar de que la idea de quitar información hace

de esta técnica poco satisfactoria, es hasta los momentos el más seguro de los procedimientos, esto se debe a que aún no se tiene un software estadístico que trate adecuadamente los valores faltantes.

4. *El objetivo principal de los métodos de análisis multivariado es resumir grandes cantidades de datos por medio de relativamente pocos parámetros. Otro interés de las técnicas de análisis multivariante es encontrar relaciones entre las variables, las unidades experimentales y tanto las variables como las unidades experimentales.*
5. *Una distinción fundamental entre los métodos multivariados es que algunos se clasifican como “técnicas dirigidas por las variables” y “técnicas dirigidas por los individuos”.*

Las técnicas dirigidas por las variable son aquellas que se enfocan primordialmente en las relaciones que podrían existir entre las variables que se están midiendo. Algunos ejemplos de este tipo de técnica se encuentran en los análisis realizados a las matrices de correlación: el análisis de componentes principales, el análisis por factores, el análisis de regresión y el análisis de correlación canónica.

Las técnicas dirigidas por los individuos son aquellas que se interesan principalmente en las relaciones que podrían existir entre las unidades experimentales o individuos que se están midiendo. Algunos ejemplos de este tipo de técnica: análisis discriminante, el análisis por agrupación y el análisis multivariado de varianza (MANOVA: Multivariate Analysis Of Variance).

6. *Los métodos basados en la distribución normal multivariada por ejemplo análisis de componentes principales, análisis discriminante y ANOVA, funcionan bien incluso cuando los datos no provienen de poblaciones normales multivariadas, siempre que estos vectores tengan distribuciones de probabilidad independiente. También cuando los valores de cada una de la variables respuesta se pueden ordenar cuantitativamente de alguna manera, este es el caso en donde la variable no es ni continua ni discreta, se podrá obtener una respuesta satisfactoria por parte del método que se use.*

2.0.3. Distribución Normal Multivariada

Es importante comprender qué se requiere para que un vector de variables aleatorias, como

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

sea multivariado normalmente distribuido.

Definición 2.4 Se dice que un vector X p -variado tiene distribución normal multivariada si

$$a'X = [a_1 \ a_2 \ \dots \ a_p] \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = a_1X_1 + a_2X_2 + \dots + a_pX_p = \sum_{i=1}^p a_iX_i$$

tiene una distribución normal multivariada para todos los conjuntos posibles de valores para los elementos en el vector a .

Una consecuencia de la definición es que todos los elementos del vector X deben tener una distribución normal univariada, es decir, la distribución de frecuencias de los valores tanto en cada una como todas las variables respuesta, X_i , debe seguir una curva acampanada. Se podría esperar que si cada X_i estuviera normalmente distribuida, entonces el vector X tendría una distribución normal multivariada, es decir, la distribución de cada variable X_i es normalmente univariante $X_i \sim N(\mu_i, \sigma_{ii})$, $i = 1, \dots, p$.

Como hemos ya mencionado, las técnicas multivariadas que se basan en la distribución normal son robustas y funcionan bien cuando los vectores de datos no son multivariados normalmente distribuidos, siempre que estos vectores todavía tengan distribuciones de probabilidad independientes.

2.0.4. Independencia de las unidades experimentales.

Una condición importante que deben satisfacer los métodos multivariados es que las unidades experimentales deben ser independientes entre sí. En otras palabras, los datos observados de las variables medidas en una unidad experimental no deben influir sobre los valores observados de las variables medidas de otra unidad experimental. Este hecho es corroborado cuando se ve en detalle *el teorema central del límite multivariado*:

Teorema 2.5 Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias p -dimensionales independientes e idénticamente distribuidas. Sea \bar{X}_n el vector p -dimensional media aritmética de las n primeras variables:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Se tiene lo siguiente:

1. Si existe $E[X_i] = \mu$, entonces \bar{X}_n converge a μ casi seguramente (Ley de los grandes números p -dimensional).
2. Si, además, X_i tiene una matriz de varianzas y covarianzas finitas Σ , entonces

$$\sqrt{n}(\bar{X}_n - \mu) \rightarrow_D N_p(0, \Sigma)$$

donde la expresión \rightarrow_D indica convergencia en distribución (teorema central del límite p -dimensional).

Si no se logra la independencia entre las unidades, las técnicas multivariadas que se van a desarrollar no podrían aplicarse correctamente a los datos. La demostración del teorema anterior se verá al usar el concepto de matriz de correlación que se desarrollará más adelante.

2.0.5. Vectores de medias y matrices de varianzas-covarianzas.

Definición 2.6 La media de un vector de variables aleatorias X se denota por μ y la matriz de covarianza de X se denota por Σ . Éstos se definen por:

$$\mu = \mathbb{E}[X] = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_p] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} y$$

$$\Sigma = \text{Cov}(X) = \mathbb{E}[(X - \mu)(X - \mu)']$$

$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}$$

en donde

$$\begin{aligned}\sigma_{ii} &= \text{Var}(X_i) = \mathbb{E}[(X_i - \mu)^2], \text{ para } i = 1, 2, \dots, p, \text{ y} \\ \sigma_{ij} &= \text{Cov}(X_i, X_j) \\ &= \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)], \text{ para } i \neq j = 1, 2, \dots, p.\end{aligned}$$

2.0.6. Correlaciones y matrices de correlación

Definición 2.7 Coeficiente de correlación.

El coeficiente de correlación entre X_i y X_j se denotará por ρ_{ij} y se define por:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

el valor ρ_{ij} se encuentra dentro del rango $[-1, 1]$.

El coeficiente de correlación proporciona una medida de la relación lineal entre dos variables, se puede representar como la recta imaginaria que pasa sobre los puntos ó a una distancia muy cercana a ellos, tal que al graficar la recta dan una idea de asociación entre ambas variables. Más la asociación lineal no implica necesariamente causalidad, es decir, el hecho que halla correlación no significa que una variable sea la causante de que exista la otra, pero tampoco podemos rechazar esta idea de causalidad dado que también podría ocurrir. Este hecho es conocido como "**Cum hoc propter hoc**", es una frase en latín que significa "después de esto, por lo tanto, a consecuencia de esto", es una falacia que se comete al inferir entre dos eventos que están conectados causalmente por haberse observado una correlación estadística entre ellos. Si tenemos las variables X_i y X_j , con una correlación cercana a los extremos de $[-1, 1]$, existen cuatro posibilidades que podemos inferir y que son correctas:

1. Que X_i sea causante de la existencia de X_j .
2. La existencia de un factor desconocido que sea realmente la causa de la relación entre X_i y X_j .
3. Que la relación sea tan compleja y numerosa que los hechos sean simples coincidencias.
4. Que la relación sea bilateral causal, o una asociación simbiótica; X_i causa la existencia de X_j y X_j causa la existencia de X_i .

Sin un estudio profundo de la situación podemos hasta estar infiriendo una causalidad falsa, por ejemplo al asumir que X_i sea causante de la existencia de X_j sea correcto, podría ser que X_j sea causante de la existencia de X_i .

Ejemplo 2.8 Algunos tipos de correlación. Para la correlación tenemos dos clases: positiva cuando $\rho_{ij} \approx 1$, es decir, que cuando aumenta una variable lo hace también la otra, y la negativa cuando $\rho_{ij} \approx -1$, es decir, que cuando aumenta una variable la otra disminuye.

La idea de no estar correlacionadas, $\rho_{ij} = 0$, es equivalente a que las variables son independientes cuando las variables X_i y X_j tienen una distribución normal bivariada, la demostración de este hecho se verá más adelante.

Nunca debe calcularse un coeficiente de correlación entre dos variables sin también situar las dos variables una contra otra en una gráfica. Si es importante determinar la correlación entre dos variables, entonces es de igual importancia trazar una gráfica de dispersión de ellas. Por lo común, la gráfica de dispersión revelará si, en realidad, las dos variables están relacionadas entre sí, la propia gráfica revelará como podrían estar relacionadas. Como ya se ha mencionado con una correlación cercana a los extremos de $[-1, 1]$ existe una relación lineal, pero al tener una correlación igual a cero no debemos descartar el hecho que las variables esten relacionadas, es importante decir que no lo hará de forma lineal, pero por ejemplo lo que puede haber es una relación cuadrática; este ejemplo nos muestra que no debemos descartar la idea de que pueda existir una relación entre las dos variables.

Proposición 2.9 Cuando X y Y son estadísticamente independientes, se puede mostrar que la covarianza es cero $\sigma_{ij} = 0$. Lo opuesto, por lo general, no es cierto a menos que las variables tengan distribución normal bivariada.

Demostración. (\Rightarrow) Sean X y Y variables independientes. La covarianza de X y Y , sabiendo que μ_x y μ_y son sus medias respectivamente, viene dada por:

$$\begin{aligned}\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY] - \mu_X\mu_Y \\ &= E[X]E[Y] - \mu_X\mu_Y \\ &= \mu_X\mu_Y - \mu_X\mu_Y \\ &= 0.\end{aligned}$$

Usando el coeficiente de correlación

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{0}{\sigma_X\sigma_Y} = 0.$$

(\Leftarrow) Sea $\sigma_{XY} = 0$ y $\rho_{XY} = 0$, además X e Y tienen distribución normal multivariada conjunta. Veamos que X e Y son independientes, para eso introduciremos la teoría correspondiente para esta demostración:

Definición 2.10 La función de densidad conjunta de X e Y con distribución normal es:

$$f(X, Y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho_{XY}^2)}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right) + \left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}.$$

También hay que tener claro la independencia dentro de una función de densidad.

Definición 2.11 Dada la distribución de probabilidad conjunta $f(x, y)$ de las variables aleatorias X e Y , la distribución de probabilidad $g(x)$ de X sola se obtiene al sumar (caso discreto) o integrar (caso continuo) $f(x, y)$ sobre los valores de Y . De manera analoga, la distribución de probabilidad $h(y)$ de Y se obtiene al sumar (caso discreto) o integrar (caso continuo) $f(x, y)$ sobre los valores de X . Definimos a $g(x)$ y $h(y)$ como distribuciones marginales de X e Y respectivamente.

Definición 2.12 Sean X e Y dos variables aleatorias continuas, con distribución de probabilidad conjunta $f(x, y)$ y distribuciones marginales $g(x)$ y $h(y)$. Se dice que las variables aleatorias X e Y son estadísticamente independientes si y solo si:

$$f(X, Y) = g(X).h(Y)$$

donde $g(X) = \int_{-\infty}^{\infty} f(X, Y)dY$ y $h(Y) = \int_{-\infty}^{\infty} f(X, Y)dX$.

Luego usando la definición de densidad conjunta de una distribución normal

$$f(X, Y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho_{XY}^2)}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right) + \left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}.$$

y usando el hecho $\rho_{XY} = 0$,

$$\begin{aligned} f(X, Y) &= \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left\{-\frac{1}{2}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2 + \left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right]\right\} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\sigma_X\sigma_Y} \exp\left\{-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2\right\} \exp\left\{-\frac{1}{2}\left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right\} \\ &= \left\{\frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2\right\}\right\} \left\{\frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left\{-\frac{1}{2}\left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right\}\right\} \\ &= g(X)h(Y). \end{aligned}$$

Entonces X e Y son independientes con función de densidad normal. □

Definición 2.13 Matriz de correlación.

La matriz de correlación para un vector aleatorio X se denota por P y se define por:

$$P = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \dots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{bmatrix}.$$

2.0.7. Pruebas estadísticas e intervalos de confianza para correlaciones.

Lo que en realidad se necesita saber es cuáles correlaciones son suficientemente grandes como para tener alguna importancia práctica. Por ejemplo cuando se reúnen datos dentro de ambientes controlados, como los laboratorios, las correlaciones mayores que 0,9 no son desacostumbradas. Sin embargo, cuando se reúnen datos de poblaciones en las que el investigador tiene muy poco control, puede ser que correlaciones mayores que 0,7 sean difíciles de obtener.

Una guía aproximada, pero no tanto confiable, para datos que se reúnen de poblaciones poco o no controlados: para muchas *situaciones experimentales* es que su correlación cuyo valor absoluto sea mayor que 0,6 puede considerarse bastante grande como para tener una importancia práctica. En esos casos, es probable que la relación lineal entre las dos variables sea suficientemente fuerte como para que sea posible usar una de ellas como sustituto de las otra. Para datos que se reúnen provenientes de personas, en ocasiones se van a considerar las correlaciones cuyos valores absolutos sean mayores que 0,5 tan grandes como para indicar relaciones importantes de variables.

En general, no se debe calcular correlaciones de la muestra si el tamaño de ésta es menor que 12, debido a que la cantidad de observaciones es demasiado pequeña como para dar resultados confiables y, con mucho, es fácil que los investigadores interpreten de manera equivocada las correlaciones medidas a partir de muestras pequeñas.

Intervalos de confianza por la aproximación de Fisher.

Fisher, según informan Kendall y Stuart (1961), demostró que, cuando se toman muestras de tamaño mayor que 25 de una distribución normal bivariada con correlación ρ , la tangente hiperbólica inversa de la correlación de la muestra tiene poco más o menos una distribución normal con media igual a la tangente hiperbólica inversa de ρ y varianza $1/(N - 3)$. Es decir,

$$U = \operatorname{tanhinv}(r)$$

tiene una distribución aproximadamente normal con media

$$\delta = \operatorname{tanhinv}(\rho)$$

y varianza $1/(N - 3)$, cuando $N > 25$. Dado que en muchas calculadoras se cuenta con la inversa hiperbólica ($\operatorname{tanhinv}$), se puede usar este resultado para construir intervalos de confianza en torno a ρ para tamaños grandes de muestra. En particular, un intervalo de confianza del $(1 - \alpha)100\%$ para ρ se expresa por

$$\operatorname{tanh} \left\{ \operatorname{tanhinv}(r) - \frac{z_{\alpha/2}}{\sqrt{N-3}} \right\} < \rho < \operatorname{tanh} \left\{ \operatorname{tanhinv}(r) + \frac{z_{\alpha/2}}{\sqrt{N-3}} \right\}.$$

Para poner un ejemplo, supongamos que $N = 25$ y $r = 0,7$. Entonces $\operatorname{tanhinv}(0,7) = 0,86673$. Por consiguiente, el punto extremo izquierdo de un intervalo de confianza de 95 % es más o menos

$$\operatorname{tanh} \left\{ 0,8673 - \frac{1,96}{\sqrt{22}} \right\} = \operatorname{tanh}[0,8673 - 0,4179] = \operatorname{tanh}[0,4494] = 0,421$$

y el punto extremo derecho está alrededor de

$$\operatorname{tanh} \left\{ 0,8673 + \frac{1,96}{\sqrt{22}} \right\} = \operatorname{tanh}[0,8673 + 0,4179] = \operatorname{tanh}[1,2852] = 0,858.$$

Como consecuencia un intervalo de confianza de 95 % para ρ es

$$0,42 < \rho < 0,86.$$

Como conclusión podemos decir que el coeficiente de correlación calculado es significativamente distinto de cero.

Si no disponemos de calculadora, con la función tangente hiperbólica que viene dada con la función logaritmo natural, se tendrá:

$$\operatorname{tanh}^{-1}(r) = (1/2)\log[(1+r)(1-r)]$$

en donde \log es la función logaritmo natural y

$$\operatorname{tanh}(w) = (\exp^w - \exp^{-w}) / (\exp^w + \exp^{-w}) = (\exp^{2w} - 1) / (\exp^{2w} + 1).$$

Intervalos de confianza por aproximación de Ruben.

La aproximación de Ruben es más exacta que la de Fisher.

Sea N el tamaño de la muestra y r la correlación observada de la misma, u el punto crítico superior $\alpha/2$ de la distribución normal estandar. Con los siguientes elementos que deben ser hallados:

$$r^* = \frac{r}{(1-r^2)^{1/2}}, \quad a = 2N - 3 - u^2,$$

$$b = r^* [(2N - 3)(2N - 5)]^{1/2} \text{ y } c = (2N - 5 - u^2)r^{*2} - 2u^2.$$

Supongamos que y_1 e y_2 son las raíces de la ecuación cuadrática $ay^2 - 2by + c = 0$, entonces los límites de confianza, superior e inferior, del $(1 - \alpha)100\%$ para ρ son $y_1 / (1 + y_1^2)^{1/2}$ e $y_2 / (1 + y_2^2)^{1/2}$, respectivamente.

2.0.8. Agrupamiento de datos basados en correlaciones.

Cuando se miden varias variables en cada una de un gran número de unidades experimentales, a menudo resulta interesante ver si esas variables están interrelacionadas y de que manera lo están. Para examinar estas interrelaciones, cuando varias parejas de variables están intensamente correlacionadas entre sí, se podría intentar la partición de las variables respuesta en grupos, de modo que las variables dentro de un grupo tengan elevadas correlaciones entre sí y las que se encuentren en grupos diferentes tengan bajas correlaciones.

2.0.9. La función de densidad de probabilidad normal multivariada.

Definición 2.14 Rango de una matriz.

Sea A una matriz, el rango de la matriz A es el número de líneas de esa matriz (filas) que son linealmente independientes. La denotaremos como $\text{rang}(A)$ ó $r(A)$.

Una fila es linealmente independiente de otras cuando no se puede establecer una combinación lineal entre ellas.

Existen muchos métodos para calcular el rango de una matriz, usaremos el método de Gauss que propone el descarte de (1) la fila donde todos sus coeficientes son ceros, (2) hay dos filas iguales, (3) una fila es proporcional a otra y (4) una fila es combinación lineal de otras. Después $\text{rang}(A)$ va a ser igual al número de filas.

Supongamos que X es un vector aleatorio p -variado que tiene distribución normal multivariada con vector de medias μ y matriz de varianzas-covarianzas Σ . Expresado también como:

$$X \sim N_p(\mu, \Sigma).$$

Además de ser matrices simétricas, tanto Σ como P son matrices no negativas y tienen el mismo rango. Si el rango de Σ es igual a p , existe una función de densidad de probabilidad para X . En este caso, la función de densidad de probabilidad de X se debe expresar como:

$$f_x(x; \mu, \Sigma) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp[-(1/2)(x - \mu)' \Sigma^{-1} (x - \mu)] \text{ para } X \in E_p.$$

En donde E_p presenta el espacio vectorial p dimensional de números reales. Es decir:

$$E_p = x : -\infty < x_i < \infty, \text{ para } i = 1, 2, \dots, p.$$

2.0.10. Estimaciones insesgadas.

Sean X_1, X_2, \dots, X_N una muestra aleatoria de una distribución multivariada que tiene un vector de medidas, μ , una matriz de covarianzas, Σ , y una matriz de correlación, P . Las estimaciones insesgadas de μ y Σ se expresan por.

$$\hat{\mu} = \frac{1}{N} \left(\sum_{r=1}^N X_r \right) = \frac{X_1 + X_2 + \dots + X_N}{N}$$

y

$$\hat{\Sigma} = \frac{1}{N-1} \left[\sum_{r=1}^N X_r (X_r - \hat{\mu})(X_r - \hat{\mu})' \right]$$

respectivamente. Notemos que:

$$\hat{\mu} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix} = \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \vdots \\ \hat{\mu}_p \end{bmatrix} \text{ y}$$

$$\hat{\Sigma} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \dots & \hat{\sigma}_{1p} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \dots & \hat{\sigma}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{p1} & \hat{\sigma}_{p2} & \dots & \hat{\sigma}_{pp} \end{bmatrix}$$

en donde

$$\hat{\sigma}_{ii} = \widehat{Var}(X_i) = \frac{1}{N-1} \sum_{r=1}^N (X_{ri} - \bar{X}_i)^2$$

y

$$\hat{\sigma}_{ij} = \widehat{Cov}(X_i, X_j) = \frac{1}{N-1} \sum_{r=1}^N (X_{ri} - \bar{X}_i)(X_{rj} - \bar{X}_j).$$

Los estimadores de los coeficientes de correlación ρ_{ij} suelen tomarse como $r_{ij} = \hat{\sigma}_{ij} / \sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}$, aún cuando r_{ij} no sea una estimación insesgada de ρ_{ij} , para cualquier i y j . Aunque r_{ij} no sea una estimación insesgada ρ_{ij} , sí tiene la propiedad de que $-1 \leq r_{ij} \leq 1$, para toda $i \neq j$. Es posible multiplicar r_{ij} por una constante para hacerla insesgada, pero después de seguir este procedimiento se podría encontrar que la correlación estimada pudiera ser mayor que 1 o menor que -1, lo cual no tendría sentido. No usaremos las estimaciones insesgadas de los coeficientes de correlación.

La matriz de correlación de la muestra se denota por R y se define por:

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1p} \\ r_{21} & 1 & r_{23} & \dots & r_{2p} \\ r_{31} & r_{32} & 1 & \dots & r_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & r_{p3} & \dots & 1 \end{bmatrix}.$$

2.0.11. Datos estandarizados y/o valores Z

A veces, los datos son más fáciles de comprender y de comparar cuando las variables respuesta se estandarizan, de modo que se midan en unidades comparables. Esto suele hacerse mediante la eliminación de las unidades de medición en conjunto. Defínase

$$Z_{rj} = \frac{x_{rj} - \hat{\mu}_j}{\sqrt{\hat{\sigma}_{jj}}}, \text{ para } r = 1, 2, \dots, N, j = 1, 2, \dots, p$$

La variable Z_{ij} se llama valor Z para la j -ésima unidad experimental y

$$Z = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1p} \\ Z_{21} & Z_{22} & \dots & Z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \dots & Z_{Np} \end{bmatrix}$$

Se conoce como la matriz de valores Z.

Se recomienda la estandarización cuando las variables medidas están en unidades completamente diferentes. Por ejemplo, suponga que un investigador tiene medidas de la estatura y el peso de personas. Necesariamente, estas medidas están en unidades totalmente diferentes. Con frecuencia, es mucho más fácil comparar individuos con respecto a estas dos variables si se estandariza cada una de las variables.

Observación 2.15 Significancia de la correlación de una muestra.

Saber que grado de significancia tiene nuestra correlación muestral es de suma importancia, porque así se concluirá si dos variables disponen de la propiedad de correlación o no. Por poner un ejemplo supongamos que la correlación de la muestra de dos variables respuesta es igual a 0,90. ¿Encontraríamos interesante ese resultado? ¿Que pasaría si la correlación de la muestra fuera igual a 0,30? ¿Cómo cambiaría nuestra interpretación si la correlación de 0,90 no fuera significativamente diferente de cero y el de 0,30 si lo fuera?. Hablaremos acerca de algunas fórmulas para el cálculo de la significancia, así como la manera en que se podrían usar esas correlaciones para hacer inferencias acerca de la población de la que se ha extraído la muestra.

Es importante construir intervalos de confianza para los valores verdaderos de los coeficientes de correlación de la población. Las gráficas de dispersión son importantes porque dan información visual referente a si podrían existir otras relaciones entre las parejas de variables, que no sean lineales. Los intervalos de confianza son importantes porque dan información confiable en cuanto al tamaño numérico real de un coeficiente de correlación de la población.

La técnica que se va a usar para el estudio de los datos es el análisis ANOVA.

2.1. Análisis Univariado de Varianza.

Ronald Fisher fue innovador del uso de los métodos estadísticos en el diseño de experimentos. Fue él quien desarrollo y usó por primera vez el análisis de varianza como herramienta primaria para el análisis estadístico en el diseño experimental. Muchas de las primeras aplicaciones de los métodos del diseño experimental se dieron en el área de agricultura y ciencias biológicas; sin embargo, sus primeras aplicaciones industriales se hicieron en la década de 1930 en las industrias textil y de lana británica.

2.1.1. Áreas de aplicación.

El análisis de varianza tiene una amplia aplicación en muchas disciplinas; su uso es muy importante para la mejora de procesos de manufactura, el desarrollo de nuevos procesos, el desarrollo de nuevos productos y la mejora de otros ya existentes.

El análisis de varianza descompone la variabilidad del resultado de un experimento en componentes independientes. Por ejemplo, se puede considerar la productividad de tres máquinas compradas a un mismo proveedor que, aunque provienen de la misma empresa, pueden producir cantidades distintas de piezas. Esta variabilidad puede producirse por muchos factores controlables (ajuste de máquinas, presión ejercida, etc), donde cada uno puede presentar diferentes niveles (ajuste de primer nivel, de segundo o cantidades distintas de presión), o bien por factores no controlables (clima, cansancio del operador, etc).

También es útil para conocer qué campaña publicitaria tiene más impacto en las ventas de algún producto, qué tipo de aprendizaje repercute en las calificaciones de los alumnos de ciertas carreras o cuál dieta ayuda a bajar de peso. Se pueden modelar los fenómenos que surjan en los procesos productivos, procedimientos de manufactura, métodos de aprendizaje y sobre todo cuando a los individuos, servicios o productos se les aplican diferentes tratamientos como en medicina y psicología, entre otros usos.

Definición 2.16 *El análisis de Varianza es una técnica que permite comparar las medias de dos o más poblaciones.*

Etapas del análisis de varianza.

El análisis de varianza está conformado por cinco etapas que son las siguientes:

1. Formulación del problema.
2. Selección de las variables.
3. Selección del tipo de análisis de varianza.
4. Análisis de la información.
5. Interpretación de los resultados.

Supuestos del análisis de varianza.

Antes de introducir algunos de los tipos de análisis de varianza es necesario saber que, al igual que en cada técnica, aquí también se tienen supuestos:

1. Las fuentes de variación en el experimento deberán permanecer constantes o ser iguales, a esto se le conoce como **homoscedasticidad u homogeneidad de varianza**.
2. Los datos deben distribuirse normalmente, para esto se sugiere hacer un histograma de los residuos. Dicha gráfica debe parecerse a una extraída de una distribución normal. Otro procedimiento consiste en realizar una gráfica de probabilidad de los residuos, para su construcción los residuos deben colocarse en orden ascendente y graficar el k -ésimo de estos residuos ordenados contra su punto de probabilidad acumulada sobre la probabilidad normal. Si la distribución de los errores es normal, la gráfica será una línea recta.
3. Se supone que un **modelo aditivo** es el que mejor explica el comportamiento de la variable dependiente.

2.2. ANOVA de un solo factor

El modelo unifactorial se presenta cuando se tiene un solo factor con diferentes niveles, que influye sobre una variable dependientes que mide el resultado del experimento, es decir:

$$Y = f(X_1).$$

La hipótesis que se prueba en el análisis de varianza de un factor es que las medias poblacionales, esto es, las medias de la variable dependiente en cada nivel de las variable son iguales. Si las medias poblacionales son iguales significa que el comportamiento de la variable dependiente es el mismo en las diferentes alternativas que se manejan en el experimento y que, por lo tanto, el factor es independiente de la variable dependiente, con lo cual se concluirá que el factor no afecta a la dependiente.

Para poner a prueba la hipótesis de igualdad de medias se debe obtener un estadístico llamado F que refleja el grado de diferencia entre las medias que se comparan.

Al hablar de variación se refiere a lo distinto o diferente que son los datos comparados con el promedio de los mismos. En ANOVA de un factor hay tres tipos de variaciones a considerar:

Casos	Variable 1	Variable 2	Variable 3	...	Variable p
Caso 1	X_{11}	X_{12}	X_{13}	...	X_{1p}
Caso 2	X_{21}	X_{22}	X_{23}	...	X_{2p}
Caso 3	X_{31}	X_{32}	X_{33}	...	X_{3p}
⋮	⋮	⋮	⋮	⋮	⋮
Caso N	X_{N1}	X_{N2}	X_{N3}	...	X_{Np}
Suma	$X_{.1}$	$X_{.2}$	$X_{.3}$...	$X_{.p}$

1. **Variación total (SS_T).** Se refiere a lo distinto que son los datos, sin considerar que pertenecen a diferentes niveles del factor, es decir sin clasificarlos. Esta variación resulta de comparar cada dato con respecto al promedio general de datos.

$$SS_T = \sum_{i=1}^N \sum_{j=1}^p (X_{ij} - \bar{X})^2. \tag{2.1}$$

Donde:

X_{ij} : el valor del dato en la posición ij .

\bar{X} : el promedio general de los datos.

Esta variación total se forma por la variación explicada y la variación no explicada que veremos más adelante.

2. **Variación Explicada (SS_B).** Se encuentra al comparar los datos de los diferentes niveles del factor. Esta

variación resulta de comparar cada promedio de cada nivel con el promedio general de datos.

$$SS_B = N \sum_{j=1}^p (X_{ij} - \bar{X})^2. \quad (2.2)$$

Donde:

X_{ij} : el valor del dato en la posición ij .

\bar{X} : el promedio general de los datos.

3. **Variación no explicada** (SS_W). Es la que existe entre los datos dentro del mismo factor y no se considera en el experimento. Puede verse en lo distinto que son los datos dentro del mismo nivel de factor y es el resultado de la comparación de cada dato de cada nivel de factor con respecto a su correspondiente promedio de factor.

Así entonces:

$$SS_T = SS_B + SS_W.$$

Generalmente, y como veremos más adelante, la mayoría de los software presenta la información en una tabla denominada tabla ANOVA, la cual se presenta a continuación.

Fuentes de Variación	Suma de Cuadrados	Grados de Libertad	Cuadrados Medios	Prueba F
Variación Explicada	SS_B	$a - 1$	$CM_B = \frac{SS_B}{a-1}$	$F_C = \frac{CM_B}{CM_W}$
Variación No Explicada	SS_W	$N - a$	$CM_W = \frac{SS_W}{N-a}$	
Variación Total	SS_T	$N - 1$		

A continuación se realizará la prueba de hipótesis:

$H_0 : \mu_1 = \mu_2 = \mu_3$: con $\alpha = 0,05$ acepto si $F_C > 0,05$. Al aceptar la hipótesis nula concluiríamos que las medias de los tres grupos estudiados son iguales, dicho de otro modo, los grupos proceden de poblaciones con medias iguales.

$H_a : \mu_1 \neq \mu_2 \neq \mu_3$: con $\alpha = 0,05$ acepto si $F_C < 0,05$. En el caso de aceptar la hipótesis alternativa (H_a) se concluye que al menos uno de los grupos tiene una media distinta al resto.

Capítulo 3

Portafolios de inversión de Markowitz.

En este capítulo presentaremos algunos conceptos financieros que nos ayudará a comprender los pasos para crear un portafolio de inversión.

3.1. Conceptos Financieros.

Definición 3.1 *Los instrumentos de inversión son aquellos que canalizan el ahorro hacia la inversión, facilitan el acceso de la empresa a recursos financieros necesarios para el desarrollo de proyectos de inversión. Entre algunos instrumentos de inversión tenemos: los bonos, oro, acciones, entre otros.*

Definición 3.2 *Un portafolio es un conjunto de instrumentos financieros o inversiones seleccionados por una misma persona natural o jurídica.*

Definición 3.3 *Un Portafolio eficiente es un conjunto específico de activos de inversión que crean el retorno más alto posible para un determinado nivel de riesgo deseado.*

Definición 3.4 *La Bolsa de Valores es una organización privada que brinda las facilidades necesarias para que sus miembros, atendiendo los mandatos de sus clientes, introduzcan órdenes y realicen negociaciones de compra y venta de valores, tales como acciones de sociedades o compañías anónimas, bonos públicos y privados, certificados, títulos de participación y una amplia variedad de instrumentos de inversión.*

Definición 3.5 *Mercado, en economía, es un conjunto de transacciones de procesos o intercambio de bienes o servicios entre comprador y vendedor.*

Definición 3.6 El **mercado bursátil**, es un tipo de mercado el cual está relacionado con las operaciones o transacciones que se realizan en las diferentes bolsas alrededor del mundo.

El mercado bursátil se considera como un mercado centralizado y regulado. Este mercado permite a las empresas financiar sus proyectos (conseguir el dinero necesario) y actividades a través de la venta de diferentes productos, activos o títulos. Igualmente, da a los inversionistas posibilidades de inversión a través de la compra de éstos.

Definición 3.7 Un **Corredor de Bolsa**, es la persona que compra y vende activos financieros tales como divisas, acciones, productos derivados, etc. Hoy en día encontramos instituciones e independientes que cumplen con esta función.

Ya sea que una institución quiera hacer un portafolio o una persona de manera independiente, debe poseer conocimientos mínimos para ejercer este tipo de actividad, por ejemplo:

- El análisis técnico, dentro del análisis bursátil, es el estudio de la acción del mercado, principalmente a través del uso de gráficas, con el propósito de predecir futuras tendencias en el precio.
- Es importante que el corredor de bolsa comprenda que para lograr un beneficio debe siempre asumir un riesgo, y de optimizar la relación Rentabilidad / Riesgo.

Dentro de la teoría financiera moderna, se han identificado dos componentes principales del riesgo: el riesgo sistemático y el no sistemático.

Definición 3.8 Un **crack bursátil** es una caída vertiginosa de las cotizaciones en la mayoría de las acciones en una bolsa, durante un corto período.

Definición 3.9 El **Riesgo Sistemático** es el riesgo inherente a un mercado. En otras palabras, no afecta a una acción o un sector en particular, sino al mercado en su totalidad. Por ejemplo, en una gran crisis financiera o en un “crack bursátil” todas las acciones tienden a bajar de manera simultánea. Es un riesgo impredecible pero también imposible de evitar completamente.

Se dice que el riesgo sistemático es un riesgo no diversificable dentro de un mismo mercado. No obstante, uno puede invertir en mercados distintos y esto también es diversificación. De esta forma, el riesgo sistemático se puede mitigar a través de una estrategia de asignación de activos (es decir invirtiendo en mercados distintos, como bonos acciones). Denotaremos el Riesgo Sistemático como (β), más adelante daremos su definición y uso práctico.

Definición 3.10 El **Riesgo no Sistemático** es el riesgo particular de cada activo, es decir, es aquél que resulta de factores propios y específicos de cada instrumento. Es decir, situaciones que afectan de manera particular a un activo y no al resto de activos dentro del mismo mercado.

Como veremos más adelante el riesgo no sistemático se dice que es diversificable, porque se puede reducir o controlar con una diversificación adecuada. En resumidas cuentas, al combinar acciones de distinto tipo se puede maximizar el rendimiento esperado y reducir el riesgo, diversificar de manera inteligente significa lograr un portafolio óptimo.

3.2. Portafolios Financieros.

Como ya hemos dicho, el portafolio es una combinación de activos realizado por una institución o un individuo. La creación de un portafolio es parte de la estrategia, además se tienen que tomar en cuenta una serie de pasos que van a permitir que halla una diversificación apropiada dentro de la cartera. Los tipos de activos para el portafolio pueden ser:

1. Acciones,
2. Bonos,
3. Monedas Extranjeras,
4. Fondos de Inversión,
5. Bienes Raíces,
6. Cuentas Bancarias,
7. Certificados de oro y plata,
8. Materia Prima,
9. O cualquier otro elemento que se espera que conserve su valor.

3.2.1. Portafolio de Markowitz.

El uso de la diversificación dentro de un portafolio no es mera coincidencia, se basa en un efecto básico muy importante que presentan los portafolios eficientes, ya que va en contra de la frase de “colocar todos los huevos en una misma canasta”, esto se debe a que al ser diversificable se está reduciendo un riesgo de carácter no sistemático como ya lo hemos explicado. Pero existe una gran diferencia entre una diversificación ingenua o principiante y la diversificación eficiente de Markowitz.

En la diversificación ingenua los inversionistas realizan el ejercicio de elegir diferentes activos, de diferente naturaleza o de diferente emisor, pero no tienen en cuenta la correlación que se presenta entre los mismos. A diferencia de la diversificación eficiente que se fundamenta básicamente en la observación de la correlación de los diferentes activos que conforman el portafolio.

Por otro lado la diversificación eficiente propuesta por Markowitz en su artículo “Modern Portfolio Theory” publicado en 1952, lleva consigo la combinación de inversiones con una correlación cercana a cero con el fin de reducir el riesgo de la cartera o portafolio sin clasificar sus rendimientos, en general cuanto menor sea la correlación de los activos en una cartera, menor será su riesgo.

El mismo Markowitz lo expresa de la siguiente manera:

“No sólo supone diversificación (el análisis de carteras) sino la clase apropiada para el motivo apropiado. Los inversores no consideran que la suficiencia de una diversificación dependa del número de acciones que compongan una cartera. Por ejemplo, una cartera con sesenta clases de acciones de ferrocarriles no estará tan bien diversificada como otra que esté compuesta de unas cuantas acciones de ferrocarriles, y otras de servicios públicos, minería, diversas empresas industriales. La razón está en que es más probable que empresas pertenecientes a una misma industria tengan resultados peores al mismo tiempo que cuando se trata de empresas pertenecientes a actividades diferentes”.

La teoría de Markowitz permite determinar lo que se denomina la frontera eficiente, la cual se define como el conjunto de portafolios conformados por todas las combinaciones de riesgo-rendimiento que se pueden obtener entre los diversos activos que hacen parte y que ofrecen el rendimiento esperado más alto para cualquier nivel de riesgo dado. La curva eficiente, que veremos en la figura 3.1, nos muestra los portafolios eficientes y los que no lo son.

Markowitz define una cartera eficiente si cumple con tres condiciones:

- Debe representar la máxima rentabilidad dentro de su clase de riesgo.
- Debe tener el riesgo mínimo dentro de su clase de riesgo.
- Debe ser legítima.

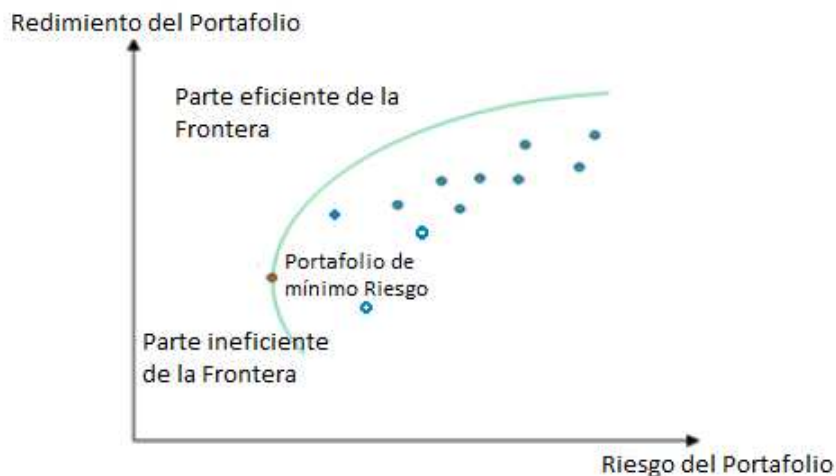


Figura 3.1: Frontera de portafolios.

Definición 3.11 *El apalancamiento financiero consiste en utilizar deuda para aumentar la cantidad de dinero que podemos destinar a una inversión. Es decir, en lugar de utilizar fondos propios, se hará con fondos propios y un crédito.*

Una cartera o portafolio es legítimo cuando no contenga títulos con proporciones negativas. Desde el punto de vista financiero quiere decir que no admiten los apalancamientos o ventas en el corto plazo.

Durante el estudio se va a tomar las decisiones observando dos parámetros de la distribución de probabilidad de la rentabilidad: la rentabilidad esperada y la varianza o la desviación estándar de los rendimientos.

Construcción de un portafolio de Markowitz.

El modelo no aconseja en cuantos ni en cuales activos invertir. En cambio resuelve, una vez escogido el número y los activos que conforman el portafolio, que cantidades invertir en cada uno de ellos de acuerdo al máximo de la media y la mínima varianza.

Definición 3.12 *El Riesgo financiero se refiere a la probabilidad de ocurrencia de un evento que tenga consecuencias financieras negativas para una organización o persona.*

Por ejemplo, Markowitz establece el objetivo de fijar el menú de las posibles combinaciones de rentabilidad de portafolio (\mathbf{R}) con un riesgo que se puede elegir (σ_p), siendo el peso asignado a los activos (\mathbf{W}) la variable sobre la cual va a tener capacidad de decisión el inversor. Si sólo tuviésemos que decidir entre dos activos (Activo 1 y Activo 2) la formulación estadística del portafolio sería:

$$R_p = W_1 \cdot R_1 + W_2 \cdot R_2$$

$$\sigma_p^2 = \sum_{i=1}^2 W_i^2 \sigma_i^2 + \sum_{i,j=1; i \neq j}^2 W_i W_j \sigma_{ij}.$$

Donde:

R_1 : es el rendimiento del Activo 1,

R_2 : es el rendimiento del Activo 2,

R_p : es el rendimiento del portafolio en general,

σ_1^2 : es la varianza del Activo 1,

σ_2^2 : es la varianza del Activo 2,

σ_p^2 : es la varianza del portafolio en general y,

W_i : es el peso (dado en %) del capital en la acción i .

Si extrapolamos a portafolios compuestos por hasta n activos la fórmula sería:

$$R_p = \sum_{i=1}^n W_i R_i, \text{ como el rendimiento del portafolio, y}$$

$$\sigma_p^2 = \sum_{i=1}^n W_i^2 \sigma_i^2 + \sum_{i,j=1; i \neq j}^n W_i W_j \sigma_{ij}, \text{ como la varianza del portafolio.}$$

Con frecuencia, el desempeño de la cartera se evalúa por lo menos durante cuatro años consecutivos, con rendimientos medidos durante periodos dentro del intervalo, por lo general, mensual o trimestralmente. Este método mantiene un tamaño de la muestra bastante adecuado para la evaluación estadística (por ejemplo, si los rendimientos se miden trimestralmente durante cuatro años, habrá 16 observaciones).

En nuestro caso supondremos la situación más sencilla, el cliente no deposita ni retira dinero de la cartera durante un lapso de tiempo. Como resultado, el cálculo del rendimiento periódico de la cartera es sencillo. Todo lo que se necesita es saber el valor de mercado de la cartera en dos puntos de tiempo, uno inicial y uno

final dentro de un periodo de tiempo, que en nuestro caso será diario.

Después de la teoría de Markowitz, y basados en ésta, se realizaron diversos estudios, como los modelos de estimación de retornos o valoración de activos “CAPM”, utilizado por William Sharpe en 1963, para explicar las ventajas de la diversificación, e introducir los conceptos de riesgo sistemático y riesgo no sistemático.

3.2.2. Portafolio CAPM

El **Capital Asset Pricing Model** o Modelo de Fijación de precios de activos de capital (CAPM). Este modelo propuesto por Sharpe (1964) establece que el rendimiento de un activo o portafolios es igual a la tasa libre de riesgo, más un premio por el riesgo que tiene ese instrumento o portafolios medido por el coeficiente beta.

Definición 3.13 *Un concepto básico en valoración es la **tasa libre de riesgo**. Como el mismo nombre la define, es aquella tasa de rendimiento que se obtiene al invertir en un activo financiero que no tiene riesgo de incumplir su pago.*

El **coeficiente beta** (β) es una medida de la volatilidad de un activo relativa a la variabilidad del mercado, de modo que al tener valores altos de beta denotan más volatilidad, y un beta 1,0 es equivalente al mercado.

Un beta de 1,75 es 75 % más volátil que el mercado. En otro caso un beta 0,7 sería 30 % menos volátil que el mercado.

Fórmula para hallar el beta(β) dentro de un activo.

Una primera forma de estimar el Beta es calculando mediante una regresión lineal entre el rendimiento del mercado como variable independiente y el rendimiento del activo como variable dependiente. Esta regresión se expresa mediante la siguiente fórmula:

$$R_i = \alpha + \beta \times R_m$$

donde:

R_i : es la tasa de rendimiento esperado del capital sobre el activo i ,

β : es la beta como riesgo,

R_m : es el rendimiento del mercado.

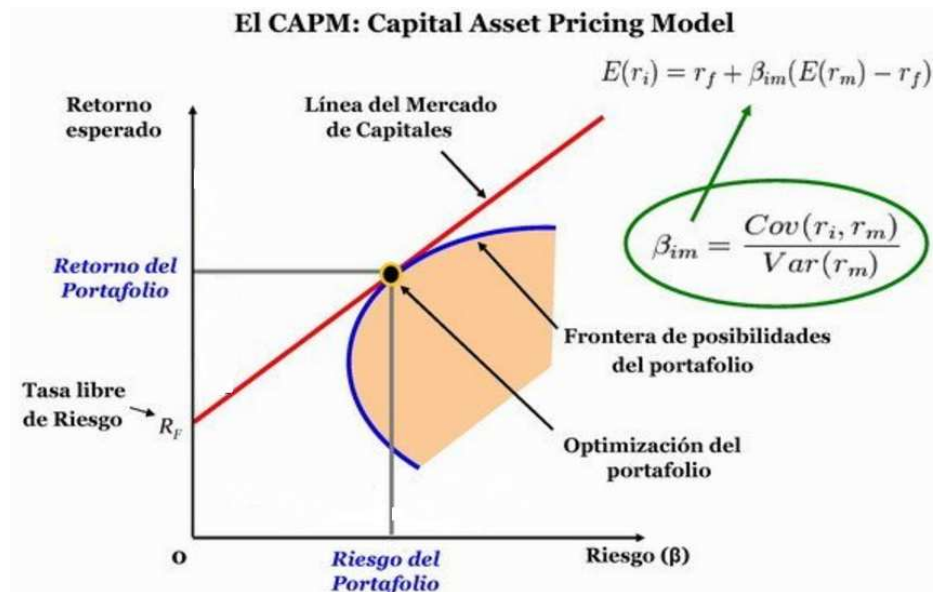


Figura 3.2: Modelo CAPM.

Fórmula para hallar la beta (β) de un portafolio.

Por otra parte en el caso del cálculo de la rentabilidad y riesgo de un portafolio p es:

$$R_p = \alpha + \beta \times R_m$$

donde:

R_p : es la tasa de rendimiento esperado del capital sobre el portafolio p ,

β : es la beta,

R_m : es el rendimiento del mercado.

Como se muestra en la figura (3.2) podemos observar una recta que refleja el comportamiento de los portafolios juntamente con el riesgo sistemático.

La importancia del factor beta (β)

Como ya hemos dicho la beta es el riesgo no diversificable y que depende del riesgo de mercado. Si el beta es cero, nuestro retorno esperado será solamente R_f , el valor del activo libre de riesgo, que sería su mínimo valor, a medida que el beta comienza a aumentar (desplazamiento hacia la derecha) por la curva horizontal de nuestra figura 3.2 aumenta también el retorno esperado. Cuando la beta es igual a 1, nuestro retorno esperado será igual al retorno del mercado. Esta es la razón por la cual un beta muy alto tiende a amplificar la respuesta del sistema. Por ejemplo si el beta es 2, el retorno del portafolio aumentará mucho más rápidamente si el mercado sube, por ejemplo, un 10%; pero también caerá más rápido si el mercado sufre una baja. Un beta elevado amplifica la tendencia, mientras que un beta menor a 1 la amortigua.

3.2.3. Índice de Sharpe.

EL índice de Sharpe mide el rendimiento en exceso (o prima de riesgo) por unidad de desviación típica en los activos de inversión.

Así, el índice de Sharpe representa como la rentabilidad de un activo, cartera o portafolio, compensa el riesgo que se asume al invertir en él. Al comparar dos activos frente a un punto de referencia común (índice del mercado), el que tiene mayor Ratio de Sharpe proporciona una mejor rentabilidad para el mismo riesgo.

Índice de Sharpe para un Activo

Fórmula de la Ratio de Sharpe para un activo i :

$$Sharpe = \frac{R_i - R_f}{\sigma_i}.$$

Donde:

R_i : Rendimiento del activo i ,

R_f : Tasa libre de riesgo(4%),

σ_i : Volatilidad observada del rendimiento del activo i .

Índice de Sharpe para un Portafolio

Fórmula de la Ratio de Sharpe de un portafolio p :

$$Sharpe = \frac{R_p - R_f}{\sigma_p}.$$

Donde:

R_p : Rendimiento del portafolios,

R_f : Tasa libre de riesgo(4%),

σ_p : Volatilidad observada del rendimiento del portafolios en el mismo periodo.

Del valor numérico del índice de Sharpe podemos extraer las siguientes conclusiones. En términos de rentabilidad, mientras mayor sea el índice de Sharpe, mejor es la rentabilidad del fondo comparado directamente a la cantidad de riesgo que se ha asumido en la inversión, es decir, la mejor relación Riesgo-Rendimiento. Si el índice o Ratio de Sharpe es negativo, indica un rendimiento inferior a la rentabilidad libre riesgo. Todo índice de Sharpe inferior a uno significa que el rendimiento del activo es inferior al riesgo que estamos asumiendo al invertir en un activo determinado. Cuando la volatilidad del fondo de inversión es grande, asumimos más riesgo y por ende el índice de Sharpe será menor, a no ser que el rendimiento del fondo en concreto compense esa mayor rentabilidad.

Importancia del índice de Sharpe

El índice de Sharpe puede utilizarse para evaluar el desempeño del portafolio sobre una base histórica en un periodo específico, siempre y cuando se conozca el monto del capital originalmente invertido. Por ejemplo, se podría calcular el indicador de Sharpe de manera trimestral en los últimos cinco años. Este análisis permitiría contestar las siguientes preguntas: ¿el negocio está tomando sistemáticamente más riesgo? ¿los rendimientos del portafolios están alineados con los riesgos que se han estado asumiendo?.

3.2.4. Índice del Precio al Consumidor(IPC)

Es un indicador estadístico que tiene como objetivo medir el cambio promedio en un periodo determinado, en los precios a nivel del consumidor de una lista de bienes y servicios representativos del consumo familiar, con respecto al nivel de precios vigente para el año escogido como base (información tomada de “El índice de

Precios al Consumidor, Año base 1997", autor Banco Central de Venezuela).

Por lo que el portafolio de IPC está destinado a cubrir las necesidades del consumidor promedio por medio de una lista de bienes y servicios.

3.3. Práctica de portafolio de Markowitz.

3.3.1. Creación de Portafolios de Markowitz y frontera eficiente.

Para hacer el uso de las herramientas de análisis de datos multivariado se ha usado el precio de cierre de acciones financieras registradas en la bolsa de valores.

Primeramente se ha tomado como referencia todos los Boletines Mensuales impreso por la Bolsa de Valores de Caracas RIF. J-00004114-8 NIT. 006397306, de los años sucesivos 2010-2014, antes de tomar una decisión concreta sobre qué acciones se va a tomar para crear el portafolio.

De cada Boletín anual entre 2010 - 2014 se han tomado cuatro acciones que fueron las más negociadas, y que han suministrado información diaria de los precios de forma continua.

Nuestro portafolio contiene las siguientes acciones:

1. Dominguez & CIA., S.A. (**DOM**): Empresa manufacturera. La actividad inherente a esta empresa es la fabricación de envases de hojalata, aluminio y plástico destinados a la industria de alimentos, pinturas, lubricantes, etc.
2. MANUFACTURAS DE PAPEL (MANPA) S.A. C.A. (**MPA**): Empresa papelera. Su objeto es la producción de pulpa, papeles, cartulinas, recolección de fibras reciclables, bolsas de papel, sacos multipliegos, papel higiénico, toallas faciales y servilletas; además de la elaboración de productos escolares.
3. BCO. PROVINCIAL (**BPV**): El objetivo primordial de la institución es la actividad de intermediación financiera. Consistente en la captación de recursos con la finalidad de otorgar créditos o financiamientos, y además operaciones y servicios financieros que sean compatibles con su naturaleza.
4. MERCANTIL SERV. FINANCIEROS C.A. CLS. (B) (**MVZ.B**): La compañía tiene por objeto la realización de toda clase de inversiones en acciones, bonos, cuotas de participación y demás obligaciones prove-

nientes de entidades mercantiles; la promoción de compañías de comercio en todas sus alternativas y modalidades; actuar como agente mediador de comercio y como comisionista, y operaciones de promoción e inversión financiera.

Agrupados por el tipo de empresas tenemos:

1. Empresas del sector financiero:

- BCO. PROVINCIAL,
- MERCANTIL SERV. FINANCIEROS C.A. CLS. (B).

2. Empresas de productos:

- DOMINGUEZ & CIA., S.A.,
- MANUFACTURAS DE MAPEL (MANPA) S.A.C.A.

Ahora procedemos al cálculo de los rendimientos promedios (R) y la desviación estándar (DESVEST) de nuestros cuatro titulares en el segundo trimestre del año 2010 y el índice del Precio al Consumidor (IPC) obtenido de la página del Banco Central de Venezuela en ese mismo periodo. Por otro lado el coeficiente de variación (CV) mostrará la relación entre el tamaño de la media y la desviación estándar de la variable.

	IPC	DOM	MPA	BPV	MVZ.B
R 2º trimestre 2010	9,21 %	15,88 %	0,67 %	11,51 %	18,92 %
DESVEST 2º trimestre 2010	5,79 %	31,65 %	28,05 %	26,85 %	20,81 %
CV	0,63	1,99	41,94	2,33	1,10

De la anterior tabla podemos destacar el rendimiento y la desviación estandar del activo MVZ.B, nos muestra un alto rendimiento (8,92 %) con respecto a otros activos y un nivel de dispersión(20,81 %) tolerable con respecto a los demás.

Calculamos el índice Beta, usando regresión lineal de cada uno de los cuatro titulares ya escogidos y el IPC, tomando como variable dependiente los retornos diarios de los activos y la variable independiente los retornos diarios del índice de precio al consumidor (IPC).

	IPC	DOM	MPA	BPV	MVZ.B
Beta	1	0,1507	0,0809	1,0602	-0,0881

También nos dispusimos en calcular el índice de Sharpe tomando en cuenta la tasa libre de riesgo (4 %).

TLR: 4%	IPC	DOM	MPA	BPV	MVZ.B
Charpe	0,9001	0,3753	-0,1187	0,2796	0,7173

Tomando en cuenta la desviación estándar de nuestros titulares ya escogidos, realizamos una tabla de correlación (CORR) y una tabla de varianza-covarianza (COV-VAR).

CORR	DOM	MPA	BPV	MVZ.B
DOM	1	0,0436	-0,0374	0,1699
MPA	0,0436	1	-0,0638	-0,0981
BPV	-0,0374	-0,0638	1	0,0111
MVZ.B	0,1699	-0,0981	0,0111	1

COV-VAR	DOM	MPA	BPV	MVZ.B
DOM	0,1002	0,0039	-0,0032	0,0112
MPA	0,0039	0,0787	-0,0048	-0,0057
BPV	-0,0032	-0,0048	0,0721	0,0006
MVZ.B	0,0112	-0,0057	0,0006	0,0433

A partir de los datos de rendimiento trimestral que denotaremos por R , elaboraremos varios portafolios para poder encontrar uno que sea eficiente. Comenzaremos por considerar que nuestro capital, que tomaremos como 100 % de todo lo que poseemos, lo colocaremos de forma equitativa a lo largo de los cuatro titulares obteniendo así 25 % de capital para cada uno de los titulares de forma inicial. También obtendremos la beta del portafolio (Beta P), la varianza dentro del portafolio, el rendimiento del portafolio, su desviación estándar y su índice de Sharpe.

Modelo Inicial	W_i	R. Anual	Beta
DOM	25 %	15,88 %	0,1507
MPA	25 %	0,67 %	0,0809
BPV	25 %	11,51 %	1,0602
MVZ.B	25 %	18,92 %	-0,0881
Sums W_i	100,00 %	Beta P	0,3009

Obteniendo así la siguiente tabla:

R Portafolio	11,74 %
Varianza Portafolio	0,0186
DESVEST Portafolio	13,65 %
I Sharpe	0,57

Después de tener el portafolio inicial, nos hemos planteado tener un total de ocho portafolios con diferentes índices de rendimiento partiendo desde 4 % que es el valor de libre riesgo hasta un 18,92 % el rendimiento máximo alcanzado por acción MVZ.B, este valor de rendimiento anual es alcanzado si colocamos el 100 % de nuestra capital al titular MVZ.B. Después usando la aplicación RESOLVER de Excel obtenemos la siguiente tabla con los ocho tipos de portafolios y tres elementos que aún no vamos a llenar:

Potf Markowitz	Desv Estd	R Esperado	I Sharpe	Beta P
1	20,6 %	4,0 %	0,0	0,4
2	17,5 %	6,1 %	0,2	0,4
3	15,0 %	8,3 %	0,3	0,3
4	13,4 %	10,4 %	0,5	0,3
5	12,9 %	12,5 %	0,6	0,3
6	13,8 %	14,7 %	0,7	0,2
7	15,8 %	16,8 %	0,8	0,2
8	20,8 %	18,9 %	0,7	-0,1
Min VAR				
Sharpe				
IPC				

Antes de terminar de rellenar la tabla tenemos que reconocer que Min Var, es decir la mínima varianza, va a generar el más óptimo de todos los portafolios, en otras palabras va a asumir un mínimo riesgo; mientras que el índice de Sharpe buscará la cartera o portafolio que más pague, por último la cartera del índice de Precios al Consumidor (IPC), creará un portafolio con todas las acciones existentes en el mercado.

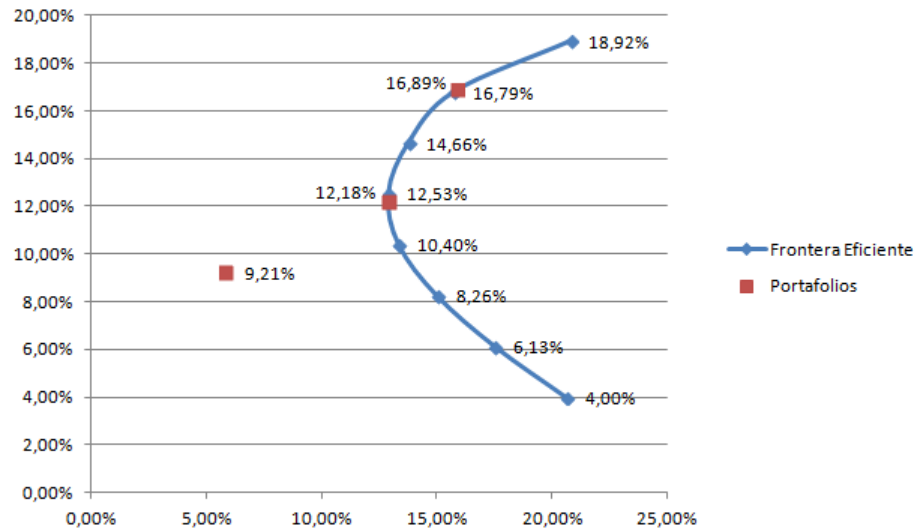


Figura 3.3: Frontera Eficiente

Potf Markowitz	Desv Estd	R Esperado	I Sharpe	Beta P
1	20,6 %	4,0 %	0,0	0,4
2	17,5 %	6,1 %	0,2	0,4
3	15,0 %	8,3 %	0,3	0,3
4	13,4 %	10,4 %	0,5	0,3
5	12,9 %	12,5 %	0,6	0,3
6	13,8 %	14,7 %	0,7	0,2
7	15,8 %	16,8 %	0,8	0,2
8	20,8 %	18,9 %	0,7	-0,1
Min VAR	12,9 %	12,2 %	0,6	0,3
Sharpe	15,9 %	16,9 %	0,8	0,2
IPC	5,6 %	9,2 %	0,9	1

Podemos ver en la figura 3.3 el portafolio de mínima varianza con rendimiento (12,2%), el portafolio de Sharpe (16,9%) y el portafolio del mercado (9,2%). Con este pequeño estudio podemos optar por el portafolio de Min Var, el portafolio 6 o 7, o el portafolio de Sharpe como buenos portafolios a invertir.

Ahora conozcamos un poco más de como se puede distribuir en porcentaje (W) el capital en el portafolio

de Min Var y el portafolio de Sharpe para hacer una comparación constructiva.

Min Var	W_i	R. Anual	Beta
DOM	12,15 %	15,88 %	0,1507
MPA	24,80 %	0,67 %	0,0809
BPV	24,91 %	11,51 %	1,0602
MVZ.B	38,14 %	18,92 %	-0,0881
Sums W_i	100,00 %	Beta P	0,2688

Sharpe	W_i	R. Anual	Beta
DOM	16,80 %	15,88 %	0,1507
MPA	0,00 %	0,67 %	0,0809
BPV	20,54 %	11,51 %	1,0602
MVZ.B	62,66 %	18,92 %	-0,0881
Sums W_i	100,00 %	Beta P	0,1878

Los portafolios, tanto Min Var como Sharpe, en su mayoría apoyan las empresas del Sector Financiero; finalizando con el orden de prioridad, dedica el resto del capital al titular Dominguez, productor de envases y otros utensilios domésticos, y MANPA productor de papel.

Como conclusión de las observaciones podemos ver que los portafolios más eficientes de los cuatro activos que escogimos dentro de los diez más negociados, además que producen información con regularidad, los más sobresalientes son en su mayoría los del sector financiero dejando en segundo lugar a los productores.

3.4. Práctica de ANOVA a los portafolios Min Var y Sharpe.

Se aplicará la herramienta ANOVA para hallar diferencias o semejanzas significativas entre los portafolios de Mínima varianza y el portafolio Sharpe.

Para dar respuesta a la tan controversial discusión se elaboraron portafolios trimestrales de ambos tipos, tanto MinVar como Sharpe, para los años 2010-2014, generando la siguiente tabla, en la que la información considerada, para cada tipo de portafolio, consiste en la desviación estándar y el rendimiento:

Usando el programa de Statistica se obtuvieron los siguientes resultados, pruebas a priori, que se realizan previo al uso del ANOVA:

Portafolios	Año	Devest	Rendimiento	Portafolios	Año	Devest	Rendimiento
Minvar	2010	14,29%	2,54%	Sharpe	2010	26,71%	10,30%
Minvar	2010	12,89%	12,18%	Sharpe	2010	15,86%	16,89%
Minvar	2010	3,39%	1,80%	Sharpe	2010	22,68%	13,69%
Minvar	2010			Sharpe	2010		
Minvar	2011	2,83%	2,33%	Sharpe	2011	18,94%	25,86%
Minvar	2011	13,55%	5,29%	Sharpe	2011	21,49%	19,06%
Minvar	2011	5,22%	5,33%	Sharpe	2011	27,70%	32,13%
Minvar	2011	7,57%	7,67%	Sharpe	2011	14,33%	17,15%
Minvar	2012	9,55%	28,76%	Sharpe	2012	10,82%	35,74%
Minvar	2012	23,19%	12,66%	Sharpe	2012	34,40%	33,63%
Minvar	2012	12,69%	13,54%	Sharpe	2012	15,99%	19,14%
Minvar	2012	27,60%	31,37%	Sharpe	2012	30,21%	38,46%
Minvar	2013	11,69%	15,04%	Sharpe	2013	13,47%	23,16%
Minvar	2013	12,25%	43,58%	Sharpe	2013	14,12%	61,42%
Minvar	2013	22,57%	44,04%	Sharpe	2013	22,96%	46,13%
Minvar	2013	18,92%	31,17%	Sharpe	2013	21,94%	40,55%
Minvar	2014			Sharpe	2014		
Minvar	2014	14,84%	-10,64%	Sharpe	2014	44,44%	28,77%
Minvar	2014	16,10%	25,64%	Sharpe	2014	18,69%	33,17%
Minvar	2014	16,49%	20,73%	Sharpe	2014	20,56%	29,98%

Cuadro 3.1: Portafolios trimestrales 2010-2014.

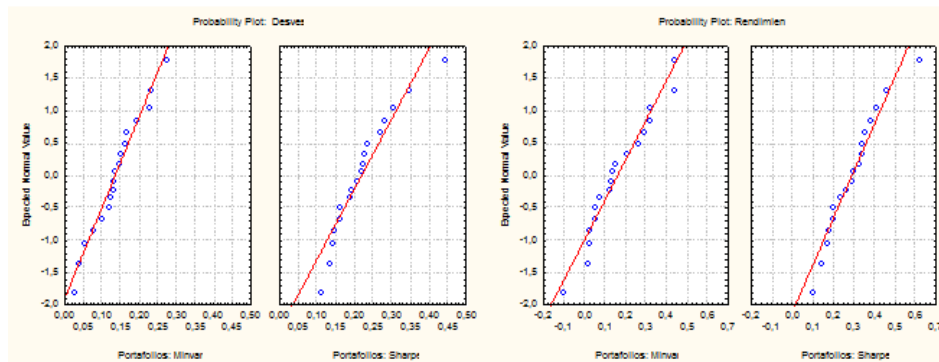


Figura 3.4: Normalidad de los datos.

1. Comprobación de la normalidad de los datos.

Según las gráficas de la figura 3.4 se puede observar que los datos son aproximadamente normales, esto se debe a que los valores representados en color azul se aproximan a la recta imaginaria de color rojo.

2. Estudio de la homogeneidad de la varianza.

Levene Test of Homogeneity of Variances (Spreadsheet19) Marked effects are significant at p < ,05000								
	SS - Effect	df - Effect	MS - Effect	SS - Error	df - Error	MS - Error	F	p
Desvest	0,001235	1	0,001235	0,083118	34	0,002445	0,505141	0,482097
Rendimiento	0,005743	1	0,005743	0,217424	34	0,006395	0,898071	0,349987

Cuadro 3.2: Homogeneidad de la Varianza.

Criterio de Homogeneidad de Varianza:

$$H_0 : \sigma_1^2 = \sigma_2^2, \text{ con } \alpha = 0,05. \text{ Acepto si la } p > 0,05.$$

$$H_1 : \text{al menos una es diferente con } \alpha = 0,05. \text{ Acepto si la } p < 0,05.$$

El test de homogeneidad de varianza visto en la tabla 3.2 nos muestra que existe dicha homogeneidad entre los datos de MinVar y Sharpe.

3. ANOVA

Analysis of Variance (Spreadsheet19) Marked effects are significant at p < ,05000								
	SS - Effect	df - Effect	MS - Effect	SS - Error	df - Error	MS - Error	F	p
Desvest	0,062239	1	0,062239	0,195929	34	0,005763	10,80057	0,002361
Rendimiento	0,149817	1	0,149817	0,667346	34	0,019628	7,63290	0,009181

Cuadro 3.3: Tablas de ANOVA.

Se aceptará o rechazará la siguiente hipótesis nula:

$$H_0 : \mu_1 = \mu_2 \text{ con } \alpha = 0,05. \text{ Se acepta si la } p > 0,05$$

H_1 : al menos una es diferente. Acepto si la $p < 0,05$

En respuesta a las reglas del ANOVA a través de los resultados de la tabla 3.3 podemos concluir que existe diferencia entre MinVar y Sharpe en sus variables Desviación estándar y Rendimiento.

Capítulo 4

Conclusión

En la elaboración de ambos portafolios, Min Var y Sharpe, nos encontramos como está distribuido el capital a través de cada activo que fue usado para tal propósito. Teniendo en gran peso el sector financiero y en una minoría el sector productivo. Podemós estar tentados en concentrarnos en el sector financiero, en tal caso estaríamos haciendo caso omiso a lo que había declarado Markowitz de tener un “portafolio diversificado”, generando un riesgo aún mayor que solo tener el riesgo perteneciente a este sector.

El uso del ANOVA ha sido de gran ayuda al discernir en la escogencia de ambos portafolios, con lo cual promovemos el uso de los métodos de análisis multivariados, respentado siempre los supuestos que corresponde a cada método. Hallar un diferencia significativa entre ambos portafolios siempre fue nuestro objeto principal, y lo hemos logrado gracias al uso del ANOVA.

En vista de lo anterior expuesto es de interés determinar el Valor en Riesgo (**VaR**) de cada portafolio, este estudio nos permitirá observar con más detenimiento las pérdidas incurridas a un nivel α deseado. Esto requerirá un estudio de estadísticos de orden, concepto de cuantil y concepto del VaR.

Bibliografía

- [1] ARRIOJAS M. (2004). Teoría de las Probabilidades. Caracas.
- [2] RONALD E. WALPOLE, RAYMOND H. MYERS, SHARON L. MYERS Y KEYING YE. Probabilidad y Estadística para Ingeniería y Ciencias. 8ª ed. Pearson, 2007.
- [3] ALFONSO DE LARA. Medición y control de riesgos financieros. 2ª ed. Limusa, 2002.
- [4] JORGE DE LA GARZA, BLANCA MORALES Y BEATRIZ GONZÁLES. Análisis Estadístico Multivariante. McGraw Hill, 2013.
- [5] DALLAS E. JHONSON. Métodos Multivariados Aplicados al Análisis de Datos. Thomson Editores, 2000.
- [6] JESÚS M TAPIA. Introducción al Análisis de Datos Multivariantes. Universidad Nacional Experimental De Los Llanos Ezequiel Zamora, 2006.
- [7] BOLSA DE VALORES DE CARACAS. Boletín Mensual. Diciembre 2010.
- [8] BOLSA DE VALORES DE CARACAS. Boletín Mensual. Diciembre 2011.
- [9] BOLSA DE VALORES DE CARACAS. Boletín Mensual. Diciembre 2012.
- [10] BOLSA DE VALORES DE CARACAS. Boletín Mensual. Diciembre 2013.
- [11] BOLSA DE VALORES DE CARACAS. Boletín Mensual. Diciembre 2014.