



Munich Personal RePEc Archive

## **Inferential Statistics**

Alfaro, Rodrigo

Instituto de Estudios Bancarios

1 March 2009

Online at <https://mpra.ub.uni-muenchen.de/15618/>  
MPRA Paper No. 15618, posted 09 Jun 2009 07:04 UTC

# Inferencia Estadística

Rodrigo A. Alfaro

Este material es un extracto del curso de Inferencia Estadística (Estadística II) realizado durante los años 2001 y 2002 en las carreras de Ingeniería en Información y Contador Auditor. Versión 2008 incompleta.



# Índice general

<b>1. Teoría de Probabilidad</b>	<b>1</b>
1.1. Definiciones Básicas . . . . .	1
1.2. Factorial, Permutación y Combinatoria . . . . .	5
1.3. Variables Aleatorias . . . . .	6
1.4. Valor Esperado . . . . .	8
<b>2. Distribuciones de Probabilidad</b>	<b>12</b>
2.1. Distribuciones Univariadas . . . . .	12
2.2. Aplicación: Riqueza como Variable Aleatoria . . . . .	18
2.3. Distribuciones Conjuntas . . . . .	22
2.4. Aplicación: Mercados Financieros . . . . .	30
2.5. Modelo Binomial . . . . .	34
2.6. Aplicación: Factor de Descuento Estocástico . . . . .	36
<b>3. Análisis Descriptivo</b>	<b>39</b>
3.1. Los Datos . . . . .	39
3.1.1. Clasificación de Estacionariedad . . . . .	39
3.2. Primeras Definiciones . . . . .	40
3.3. Conceptos Técnicos . . . . .	41
3.4. Volatilidad del Retorno de un Activo . . . . .	42
<b>4. Estadísticas</b>	<b>44</b>
4.1. Definiciones . . . . .	44
4.2. Propiedades de las Estadísticas . . . . .	45
4.3. Muestreo Finito . . . . .	46
4.4. Modos de Convergencia . . . . .	46
4.5. Leyes de los Grandes Números . . . . .	47
4.6. Teoremas Centrales del Límite . . . . .	47
4.7. Distribución de $s^2$ . . . . .	48

<b>5. Métodos de Estimación: MV y MM</b>	<b>49</b>
5.1. Máxima Verosimilitud (MV)	49
5.1.1. Propiedades de la Optimización en MV	49
5.1.2. Propiedades del Estimador MV	50
5.1.3. Ejemplo: la distribución Poisson	50
5.1.4. Parámetros condicionados	51
5.2. Método de Momentos (MM)	51
<b>6. Tests e Intervalos de Confianza</b>	<b>52</b>
6.1. Dósimas	52
6.1.1. Una muestra	52
6.1.2. Dos muestras	53
6.1.3. Ejemplos	54
6.2. Ajuste de Distribución	56
6.2.1. Prueba Chi-cuadrado	56
6.3. Tablas de Doble Entrada	56
6.4. Remuestreo	57
6.4.1. Ejemplos	57
6.5. Simulación de Monte Carlo	57
6.5.1. Ejemplo: Simulación de Escenarios	58
<b>A. Apéndice Matemático</b>	<b>59</b>
A.1. Funciones	59
A.2. Diferenciación	60
A.3. Optimización	61
<b>B. Microsoft Excel</b>	<b>63</b>
B.1. Introducción	63
B.1.1. Funciones Útiles	63
B.1.2. Tablas Dinámicas	64
B.1.3. Herramientas para el análisis	64
B.1.4. Ejemplos	64

B.2. Gráficos . . . . .	65
B.3. Optimización . . . . .	66
B.4. Programando en Excel . . . . .	66
<b>C. Código en VBA</b>	<b>68</b>
C.1. Funciones . . . . .	68
C.1.1. Función de Densidad Beta . . . . .	68
C.1.2. Función Momento Muestral . . . . .	68
C.1.3. Correlación Bivariada Spearman . . . . .	69
C.1.4. Función Test sobre Medias . . . . .	69
C.1.5. Función Test sobre Varianzas . . . . .	70
C.1.6. Función de Autocorrelación . . . . .	70
C.1.7. Función de Autocorrelación Parcial . . . . .	71
C.2. Rutinas . . . . .	71
C.2.1. Histograma . . . . .	71
C.2.2. Monte Carlo y Remuestreo . . . . .	72
C.2.3. Variación Porcentual . . . . .	73
C.2.4. Interpolación . . . . .	74



## Teoría de Probabilidad

### 1. Definiciones Básicas

Para la mayoría, por no decir todos, de los sucesos de la vida no se puede establecer un resultado con certeza. Por ejemplo, en un lanzamiento de un dado no conocemos el resultado exacto, pero probablemente estaríamos más inclinados a creer que el número resultante es mayor que uno a que el resultado sea exactamente uno. Esa inclinación se ubica dentro de nuestro inconsciente y es fruto de la experiencia<sup>1</sup> Sin embargo, si la pregunta se refiriera a una apuesta sobre caballos, una conjetura en el bridge o en el backgammon de seguro se pediría mayor información.

Este mismo mecanismo de información fue utilizado por hombres a través de su historia. Con cierto grado de confianza el brujo de una tribu de indios podía saber si es que la próxima temporada sería lluviosa o no. Esta forma de hacer probabilidad se conoce como frecuentista.

Sin embargo, la formalidad de las actualidades probabilidades tienen su origen en el siglo XVIII, donde se destaca el aporte del matemático francés Pierre Simón de Laplace. Bajo esta línea de investigación toma lugar la probabilidad clásica. Por último, la probabilidad subjetiva ocurre cuando nuestro sentido nos lleva a inclinarnos más por una alternativa que por otra.

**Definición 1.1.** *Si un experimento que está sujeto al azar, resulta de  $n$  formas igualmente probables y mutuamente excluyentes, y si  $n_A$  de estos resultados tienen un atributo  $A$  en común. Entonces la probabilidad de  $A$  es la proporción de  $n_A$  sobre  $n$ , su notación será  $P(A) = n_A/n$ . En términos simples casos favorables sobre casos totales.*

**Definición 1.2.** *El conjunto de todos los posibles resultados de un experimento aleatorio se denomina Espacio Muestral ( $S$ ). Por ejemplo los alumnos de la clase.*

**Definición 1.3.** *Un Evento ( $E$ ) del  $S$  es un grupo de resultados contenidos en éste, cuyos miembros tienen*

---

<sup>1</sup>Casi todas las personas conocen y han lanzado un dado alguna vez en su vida.



una característica en común. En términos de la Teoría de Conjuntos será  $E \subseteq S$ . Por ejemplo los alumnos varones de la clase.

**Definición 1.4.** Se dirá que dos eventos,  $E_1$  y  $E_2$ , son mutuamente excluyentes si:  $E_1 \cap E_2 = \emptyset$ . En términos más intuitivos no pueden ocurrir simultáneamente. Por ejemplo ser hombre y estar embarazado.

Utilizando un diagrama de Venn (Figura 1.1), observamos que  $E_1$  y  $E_2$  son dos eventos disjuntos, mientras que  $E_3 \subset E_2$ . En este último caso notemos que  $E_2 \cap E_3 = E_3$ , del mismo modo se cumple para todo evento  $E \subset S$  que  $E \cap S = E$ .

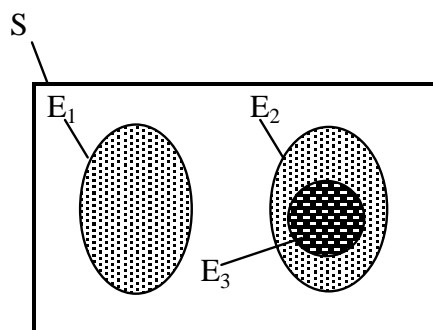


Figura 1.1: Espacio Muestral

**Definición 1.5.** Sea  $S$  y  $E$ . Se llama  $P(\cdot)$  a la función de probabilidad sobre  $S$ , que cumple con los siguientes axiomas:

1.  $P(E) \geq 0$ .
2.  $P(S) = 1$ .
3. Si  $E_1 \cap E_2 = \emptyset \Rightarrow P(E_1 \cup E_2) = P(E_1) + P(E_2)$ .

**Proposición 1.1.** Mostrar que:

1.  $P(\emptyset) = 0$
2.  $E \subset S \Rightarrow 0 \leq P(E) \leq 1$

**Ejemplo 1.1.** Suponga que se lanzan simultáneamente 2 dados. ¿Cuál es la probabilidad de adivinar el número resultante en cada uno de ellos?.

Debido a que cada dado tiene 6 posibilidades, lo cual es 6 para el dado 1 (azul) y 6 para el dado 2 (rojo), vale decir  $6 \times 6 = 36$  posibilidades en total. Si la regla es acertar exactamente el número en cada dado entonces existe 1 caso favorable, por lo que la probabilidad es  $1/36$ . En cambio si la regla no conlleva a un dado en particular (o ambos dados son iguales) entonces los casos favorables son 2, por tanto la probabilidad es  $1/18$ .

**Definición 1.6.** Se dirá Probabilidad conjunta a  $P(E_1 \cap E_2)$  cuando se aplica a la interacción de 2 eventos. Marginal cuando se refiere a sólo un evento ( $P(E_1)$ ). Y condicional en el caso de que el evento  $E_2$  depende del resultado del evento  $E_1$  ( $P(E_2|E_1)$ ).

**Teorema 1.1.** Si  $E_1 \subset E_2$ , entonces  $P(E_1) \leq P(E_2)$ .

Prueba: Si se define  $E^c$  como el complemento de  $E$ . Entonces  $E \cap E^c = \emptyset$  y  $E \cup E^c = S$ . De este modo  $E_2 = S \cap E_2 = (E_1 \cup E_1^c) \cap E_2 = (E_1 \cap E_2) \cup (E_1^c \cap E_2) = E_1 \cup (E_1^c \cap E_2)$ , la última igualdad se debe a que  $E_1 \subset E_2$ . Pero también se cumple que  $\emptyset \cap E_2 = E_1 \cap (E_1^c \cap E_2) = \emptyset$ . Concentremos la atención en los conjuntos  $E_1$  y  $(E_1^c \cap E_2)$  que son disjuntos por lo que la Definición 1.5 establece  $P(E_2) = P(E_1) + P(E_1^c \cap E_2)$ , además de la misma definición se obtiene que  $P(E_1^c \cap E_2) \geq 0$ , por lo que se demuestra que  $P(E_2) \geq P(E_1)$ .

**Teorema 1.2.**  $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$ .

Prueba: Consideremos que  $E_1 \cup E_2 = S \cap (E_1 \cup E_2) = (E_1 \cup E_1^c) \cap (E_1 \cup E_2) = E_1 \cup (E_1^c \cap E_2)$ , como los eventos  $E_1$  y  $(E_1^c \cap E_2)$  son disjuntos entonces por la Definición 1.5  $P(E_1 \cup E_2) = P(E_1) + P(E_1^c \cap E_2)$ . Por otro lado  $E_2 = S \cap E_2 = (E_1 \cup E_1^c) \cap E_2 = (E_1 \cap E_2) \cup (E_1^c \cap E_2)$ , nuevamente los eventos son disjuntos y haciendo uso de la misma definición  $P(E_2) = P(E_1 \cap E_2) + P(E_1^c \cap E_2)$ . Reordenando esta última ecuación tenemos que  $P(E_1^c \cap E_2) = P(E_2) - P(E_1 \cap E_2)$ , que al ser reemplazada en la probabilidad anterior se obtiene que  $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$ .

**Teorema 1.3.** Probabilidad Total: Sean  $\{E_i\}_{i=1}^k$  particiones mutuamente excluyentes de  $S$  y  $F$  un evento cualquiera dentro de  $S$ . Entonces se cumple que:

$$P(F) = \sum_{i=1}^k P(F \cap E_i)$$

Prueba: Notemos que  $F = F \cap S = F \cap (E_1 \cup E_2 \cup \dots \cup E_k) = (F \cap E_1) \cup (F \cap E_2) \cup \dots \cup (F \cap E_k)$ . Dado que los eventos son disjuntos podemos utilizar la Definición 1.5 por la cual se demuestra el teorema.

**Teorema 1.4.** Bayes: nos dice que para 2 eventos  $E_2$  y  $E_1$ , su probabilidad condicional viene dada por la siguiente proporción:

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

Cuadro 1.1: Ejemplo Probabilidad Conjunta

	Fuma( $B_1$ )	NoFuma( $B_2$ )
Hombres( $A_1$ )	$a$	$b$
Mujeres( $A_2$ )	$c$	$d$

**Definición 1.7.** Se dirá que 2 eventos son estadísticamente independientes si  $P(E_2|E_1) = P(E_2)$ . Utilizando Bayes, vemos que la independencia implica:  $P(E_1 \cap E_2) = P(E_1)P(E_2)$ .

**Ejemplo 1.2.** Suponga que cuenta con números enteros entre 1 y 50. Si es que se elige aleatoriamente uno de ellos, ¿cuál es la probabilidad de elegir un múltiplo de 3 o 5 o de ambos?.

Definimos  $A$  como el evento múltiplo de 3 y  $B$  como el evento múltiplo de 5. De este modo el problema se trata de buscar  $P(A \cup B)$ , que corresponde a  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ . Adicionalmente  $P(A) = 16/50 = 0,32$ , debido a que existe los siguientes casos favorables: (3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48). Por otro lado,  $P(B) = 10/50 = 0,20$ , dado: (5, 10, 15, 20, 25, 30, 35, 40, 45, 50). Finalmente  $P(A \cap B) = 3/50 = 0,06$ , por (15, 30, 45). Con todo,  $P(A \cup B) = 0,32 + 0,20 - 0,06 = 0,46$ .

**Ejemplo 1.3.** Suponga que un hombre necesita para mañana US\$1000 para pagar a la Mafia y sólo dispone de US\$2. Por tanto, compra dos billetes de lotería que cuestan un dólar cada uno en dos loterías independientes. El ganador de cada una de estas loterías obtiene un premio de US\$1000 (y no hay segundos premios). Si la probabilidad de ganar en cada lotería es  $q = 0,0001$  cuál es la probabilidad de que el hombre gane el dinero suficiente para pagar a la Mafia mañana?.

Definiendo a  $G_i$  como el evento ganar la lotería  $i$  y  $P_i$  perder aquella lotería. Tendríamos la siguiente solución:  $P(G_1 \cup G_2) = P(G_1) + P(G_2) - P(G_1 \cap G_2) = q + q - q^2$ . O alternativamente tomando el recíproco, es decir 1 menos la probabilidad de no ganar ninguna de las loterías:  $1 - P(P_1 \cap P_2) = 1 - (1 - q)^2 = 2q - q^2$ . Con el valor de  $q$ , el resultado sería 0,000199.

**Ejemplo 1.4.** El Cuadro 1.1 contiene las tabulaciones (cantidad de elementos) sobre 2 características de una muestra:  $A$  es su sexo y  $B$  si fuma o no. Asumiendo que es una muestra representativa, podemos expresar las siguientes probabilidades como casos favorables sobre casos totales: (1) Fumar y ser mujer, (2) Sólo fumar y (3) Dado que fuma, cuál es la probabilidad de ser mujer.

El primer punto es la conjunta, definida como  $P(A_2 \cap B_1) = c/(a + b + c + d)$ . El segundo es la marginal  $P(B_1) = (a + c)/(a + b + c + d)$ . Notemos que se recorre al evento  $A$  (como en la ley de probabilidad total), porque no interesa el sexo del individuo. El último caso es la condicional  $P(A_2|B_1) = c/(a + c)$ . Nos

centramos sólo en los elementos que fuman. Pero también corresponde al cociente de los puntos anteriores  $P(A_2|B_1) = P(A_2 \cap B_1)/P(B_1)$ , como lo establece Bayes.

## 2. Factorial, Permutación y Combinatoria

En la gran mayoría de los juegos de azar se mezclan una cantidad de elementos que pueden ser dispuestos de diversas formas. Para considerar esos eventos utilizamos algunos arreglos.

**Definición 1.8.** *Factorial corresponde a la multiplicación de todos los naturales desde 1 al número correspondiente, como se muestra en (1.1). Por convención  $0! = 1$ . En Excel la función es `fact()`.*

$$k! = 1 \times 2 \times \dots \times (n-1) \times n \quad (1.1)$$

**Ejemplo 1.5.** *Consideremos el número de palabras (con y sin sentido) que se pueden formar con las letras de la palabra AVIÓN (sin repetirlas).*

Dado que tiene 5 letras distintas tenemos: 5 posibilidades para la primera posición, 4 para la segunda y así sucesivamente. De este modo el número total de palabras será  $5 \times 4 \times \dots \times 1 = 5! = 120$  o `fact(5)=120`.

**Definición 1.9.** *Una permutación indica de cuántos grupos de elementos de tamaño  $k$  se pueden obtener de un tamaño  $n$ . Los elementos son tomados sin reemplazo y el orden interno es importante. Su formulación es como (1.2) y en Excel se puede obtener mediante `permutaciones(n;k)`.*

$$P_{k,n} = \frac{n!}{(n-k)!} \quad (1.2)$$

**Ejemplo 1.6.** *De cuántas formas se puede organizar una comisión de 3 personas si se cuenta con un total de 5 candidatos.*

Si el orden importa, por ejemplo si se eligen cargos directivos, entonces tratamos con una permutación  $P_{3,5} = 5!/2! = 5 \times 4 \times 3 = 60$  o `permutaciones(5;3)=60`. Ahora si se trata de una comisión que no tiene cargos trabajamos con una combinatoria de 5 sobre 3, que es  $C_{3,5} = P_{3,5}/3! = 60/3! = 60/6 = 10$  o `combinat(5;3)=10`.

**Ejemplo 1.7.** *Suponiendo que en una bolsa existe  $n$  bolitas numeradas del 1 al  $n$  y se desean extraer  $k$  bolitas, de uno en uno sin reemplazo. ¿Cuál es la probabilidad de acertar ordenadamente los números de*

cada una de las bolitas?

Sólo existe un caso favorable por lo que, la probabilidad será el inverso del número de combinaciones que se pueden formar en este problema. Para la primera bolita existen  $n$  posibilidades, para la segunda  $(n - 1)$  y así hasta  $(n - k + 1)$ , ya que se extraen  $k$  bolitas. Esto se asemeja a un factorial, pero necesitamos los  $(n - k)$  elementos multiplicativos, por ello multiplicamos y dividimos por  $(n - k)!$  formando (1.2).

**Definición 1.10.** Una combinatoria corresponde a una permutación, pero donde el orden interno de los grupos no importa. Su formulación puede ser escrita como (1.3). En Excel es `combinat(n;k)`.

$$C_{k,n} = \binom{n}{k} = \frac{P_{k,n}}{k!} = \frac{n!}{(n - k)!k!} \quad (1.3)$$

**Ejemplo 1.8.** Siguiendo el Ejemplo 1.7, podemos levantar el supuesto de que el orden importa. Esto es lo típico de los juegos de azar, donde lo que cuenta son los números ganadores y no el orden con que fueron extraídos.

Notemos que si el orden no importa nuestra probabilidad debiera aumentar, debido a que el número de casos totales cae. Cómo se extraen  $k$  bolitas, entonces existen  $k!$  combinaciones de números que son igualmente válidas, por lo que el número de casos totales debe dividirse por  $k!$  y esto es exactamente lo que se aprecia en (1.3).

### 3. Variables Aleatorias

**Definición 1.11.** Sea  $S$  con  $P(\cdot)$ . Sea  $X$  una función de valor real definida sobre  $S$  de manera que transforme los resultados de  $S$  en puntos sobre  $\mathbb{R}$ . Se dice que  $X$  es una variable aleatoria (va).

**Definición 1.12.** Una función de probabilidad sobre  $X$  (fdp), corresponderá a  $p(x)$  en el caso discreto y  $f(x)$  en el caso continuo<sup>2</sup>. Asimismo, se denominará función de probabilidad acumulada a  $F(x)$ , de tal modo que en el caso discreto será:  $F(x) = \sum_{x_i \leq x} p(x_i) = P(X \leq x)$ . Para el caso continuo:  $F(x) = \int_{-\infty}^x f(t)dt$ . Los axiomas que corresponden a esta definición son:

1.  $F(-\infty) = 0$ .
2.  $F(+\infty) = 1$ . Lo que implica que  $\int_{-\infty}^{+\infty} f(x)dx = 1$ .

<sup>2</sup>Notar que de la Teoría de Probabilidades:  $p(x) \in [0, 1]$ . Sin embargo nada asegura que  $f(x) \in [0, 1]$ .

$$3. P(a < X < b) = F(b) - F(a).$$

$$4. f(x) = \frac{dF(x)}{dx}.$$

Notemos por ejemplo el caso de una distribución Beta, que tiene como dominio  $[0, 1]$ . En la Figura 1.2 se representa la masa que acumula la probabilidad mientras se agregan más unidades del set.

Un punto importante que se puede extraer de  $F(x)$  es que debido que cada elemento aporta con una mayor probabilidad  $F(x)$  debe ser no decreciente. Por otro lado, existe una relación uno a uno entre  $x$  y  $F(x)$ , lo que quiere decir que a través de  $F^{-1}(x)$  podemos encontrar los respectivos valores de  $x$ . Este último hecho nos permitirá encontrar cuáles son los valores de  $x$  que acumulan una cierta probabilidad acumulada. Además con las funciones  $F^{-1}(x)$  podemos generar número aleatorios provenientes de la distribución que deseemos.

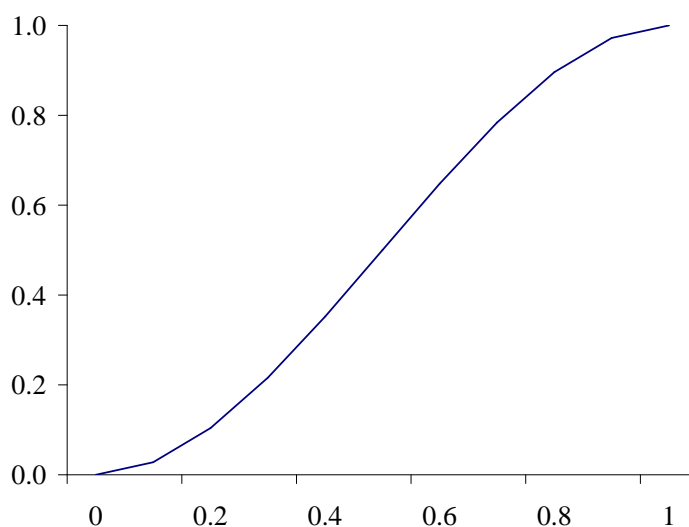


Figura 1.2: Distribución Acumulada para una Beta(2,2)

**Ejemplo 1.9.** *Suponga que el show de Pepito TV lo premia con una casa si abre la puerta correcta de las 3 que se presentan (A, B y C). Construya un mecanismo que elija a cada puerta con las siguientes probabilidades:  $P(A) = 0,5$ ,  $P(B) = 0,3$  y  $P(C) = 0,2$ .*

*Notar que la puerta A será elegida si la distribución acumulada es inferior a 0.5 (la probabilidad de un punto es cero), la puerta B será elegida si la distribución acumulada es inferior a 0.8 y C será elegido en los otros casos. Con esto hemos definido los eventos en función de su  $F(x)$ . El paso siguiente será entregar una probabilidad, que la colocaremos en `a1=aleatorio()`. Basados en la  $F^{-1}(x)$  la escribimos (notar que*

$x$	1	2	3	4	otros casos
$f(x)$	0,4	0,1	0,3	0,2	0

Cuadro 1.2: Ejemplo Variable Aleatoria Discreta

los sucesos son discretos) en  $a_2 = si(a_1 < 0,5; "A"; si(a_1 < 0,8; "B"; "C"))$ , luego presionando F9 las puertas se eligen aleatoriamente.

**Definición 1.13.** *Parámetro:* se denomina a un número que le da una característica especial a una distribución dentro de una familia.

#### 4. Valor Esperado

Permite encontrar el valor medio de una función de una va. Se encuentra definido como (1.4) para el caso discreto y como (1.5) para el caso continuo:

$$E[g(X)] = \sum_x g(x) p(x) \quad (1.4)$$

$$E[g(X)] = \int_X g(x) f(x) dx \quad (1.5)$$

Definiendo  $E()$  como operador esperanza, notaremos que este es lineal como se muestra en (1.6) para todo par de funciones  $g_1(X)$  y  $g_2(X)$ . En particular, si consideramos  $g_1(X) = aX$  y  $g_2(X) = b$ , con  $a$  y  $b$  constantes, tenemos que  $E(aX + b) = aE(X) + b$ .

$$E[g_1(X) + g_2(X)] = \int_X [g_1(x) + g_2(x)] dx = \int_X g_1(x) dx + \int_X g_2(x) dx = E[g_1(X)] + E[g_2(X)] \quad (1.6)$$

**Ejemplo 1.10.** Sea  $X$  una va discreta con la fdp como se muestra en la Tabla 1.2. Computar  $E(X)$ .

Como se trata de un caso discreto podemos considerar los valores ponderados por sus respectivas probabilidades como sigue<sup>3</sup>:

$$E(X) = \sum_{i=1}^4 i f(x_i) = 1 \cdot \frac{4}{10} + 2 \cdot \frac{1}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{2}{10} = \frac{23}{10} = 2,3$$

<sup>3</sup>También es posible usar la función `sumaproducto()` en caso de contar con los datos en una hoja de cálculo.

**Ejemplo 1.11.** Para la va  $X$  encuentre el valor  $E(X)$  y  $E(X^2)$

$$\begin{aligned} f(x) &= 2(1-x) && 0 < x < 1 \\ &= 0 && \text{en otros casos} \end{aligned}$$

Considerando los momentos centrados en cero tenemos para  $E(X)$  y  $E(X^2)$  como (1.7) y (1.8), respectivamente.

$$E(X) = 2 \int_0^1 x(1-x) dx = 2 \left[ \frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 = \frac{1}{3} \quad (1.7)$$

$$E(X^2) = 2 \int_0^1 x^2(1-x) dx = 2 \left[ \frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 = \frac{1}{6} \quad (1.8)$$

**Ejemplo 1.12.** Para la va  $X$  encuentre  $E[X^2]$ .

$$\begin{aligned} f(x) &= (1/5) \exp(-x/5) && x > 0 \\ &= 0 && \text{en otros casos} \end{aligned}$$

Consideremos la Función Gamma (explicada en A.5), donde para un entero se tiene que  $\Gamma(x) = (x-1)!$ . Por otro lado, tomaremos la variable  $y = x/5$  de modo que  $dx = 5dy$ . Reemplazando tenemos  $E[x^2] = 25\Gamma(3) = 50$ , cuyo desarrollo es:

$$E[x^2] = \frac{1}{5} \int_0^\infty x^2 \exp\left(-\frac{x}{5}\right) dx = \frac{1}{5} \int_0^\infty (5y)^2 \exp(-y) (5dy) = 25 \int_0^\infty y^2 \exp(-y) dy$$

**Definición 1.14.** La Función Generadora de Momentos (FGM), nos permite caracterizar completamente a una distribución<sup>4</sup>. Esto porque a partir de ella es posible extraer todos los momentos. Para  $X$  va, corresponde a (1.9)

$$M_X(t) = E[\exp(tX)] \quad (1.9)$$

Derivando  $r$  veces la expresión anterior y considerando  $t = 0$  podemos encontrar los momentos centrados en cero (1.10):

$$\frac{d^r M_X(0)}{dt^r} = E(X^r) \quad (1.10)$$

Si se considera  $\dot{x} = dx/dt$  y  $\ddot{x} = d^2x/dt^2$ , entonces:

<sup>4</sup>No todas las distribuciones poseen FGM, por lo que se utiliza la función característica que es un concepto más general. Siendo  $i = \sqrt{-1}$ , se define como:  $\varphi_X(t) = E[\exp(itX)]$ .



$$1. \dot{M}_X = E[X \exp(tX)]|_{t=0} = E(X).$$

$$2. \ddot{M}_X(0) = E(X^2).$$

**Teorema 1.5.** Considerando el logaritmo de la FGM como  $m(t) = \log M(t)$ , se tiene que:  $m(0) = 0$ ,  $\dot{m}(0) = E(X)$  y  $\ddot{m}(0) = V(X)$ .

Prueba: Notemos que  $M(0) = 1$ , entonces  $m(0) = 0$ . Por otro lado:  $\dot{m} = \dot{M}(0)/M(0) = \dot{M}(0) = E(X)$  y finalmente  $\ddot{m} = M(0)^{-2}[\ddot{M}(0)M(0) - \dot{M}(0)^2] = \{E(X^2) - [E(X)]^2\} = V(X)$ .

**Definición 1.15.** Los Momentos Teóricos, pueden ser clasificados en 2 grupos: (1) centrados en cero como en (1.11) y (2) centrados en la media como (1.12).

$$\theta_r = E[X^r] \tag{1.11}$$

$$\mu_r = E[(X - \mu)^r] \tag{1.12}$$

El primer momento centrado en cero se conoce como esperanza, mientras que al segundo momento centrado en la media se le denomina varianza. Adicionalmente se usan momentos escalados como (1.13) llamado *skewness* que representa la asimetría de la distribución. Visualmente una distribución con asimetría positiva (negativa) concentrará su masa de probabilidad a la izquierda (derecha), mientras que las distribuciones simétricas (en torno a su media) mostrarán asimetría cero.

$$\frac{\mu_3}{\sqrt{\mu_2^3}} \tag{1.13}$$

**Ejemplo 1.13.** Para el caso de una distribución simétrica  $X$  que tiene media cero podemos decir que un valor  $x > 0$  acumula la misma masa de probabilidad a su derecha que el que acumula  $-x$  a la izquierda.

El cuarto momento escaldado se conoce como *curtosis* y es computado como (1.14). Este indicador nos señala cuánto pesan las colas de la distribución en términos de la masa de probabilidad. Vale decir una distribución platocurtica que es aquella que posee colas altas o pesadas, acumula bastante probabilidad en las colas por lo que los valores extremos de la va son más probables que en el caso de una leptocúrtica, donde encontraremos la mayoría de la masa de probabilidad concentrada en la media y por tanto, las colas serán menos pesada o más bajas.

$$\frac{\mu_4}{\mu_2^2} \tag{1.14}$$

**Definición 1.16.** *Transformación de va.* Se puede realizar de 2 formar. La primera de ellas es a través de la FGM y la segunda es, a través de la transformación Jacobiana.

**Teorema 1.6.** *Sea  $X$  una va continua con fdp  $f_X(x)$  y defínase  $Y = g(X)$ . Si  $y = g(x)$  y  $x = g^{-1}(y)$  son funciones univaluadas, continuas y diferenciables, la fdp de  $Y$  está determinada por:  $f_Y(y) = f_X[g^{-1}(y)] \cdot J$ , donde  $J = |dx/dy|$  recibe el nombre de Jacobiano de la transformación.*

*En caso que la relación entre  $X$  e  $Y$  no sea directa, se debe ajustar la fdp acumulada para que se mantenga en rango  $[0, 1]$ .*

**Teorema 1.7.** *Desigualdad de Chebyshev:* Sea  $u(X)$  una no negativa función de la va  $X$ . SI  $E[u(X)]$  existe entonces para toda constante  $c > 0$  se cumple que:

$$P[u(X) \geq c] \leq \frac{E[u(X)]}{c} \quad (1.15)$$

Prueba: Consideremos  $X$  una va continua con fdp  $f(x)$ . Además se define el conjunto  $A = \{x : u(x) \geq c\}$  y  $A^c$  su complemento. Entonces:

$$E[u(X)] = \int_X u(x)f(x)dx = \int_A u(x)f(x)dx + \int_{A^c} u(x)f(x)dx$$

Notar que el segundo término de la derecha corresponde a una probabilidad por tanto es no negativo. Dado que  $x \in A$ , se establece que  $u(x) \geq c$  como se muestra a continuación:

$$E[u(X)] \geq \int_A u(x)f(x)dx \geq c \int_A f(x)dx = cP(A)$$

Adicionalmente  $P(A) = P(X \in A) = P[u(X) \geq c]$ . Dividiendo a ambos lado de la desigualdad por  $c$  queda demostrado el teorema.

## Distribuciones de Probabilidad

### 1. Distribuciones Univariadas

#### Distribuciones Discretas

Estas distribuciones están pensadas para sucesos que son va definidas en los naturales (incluido el cero). Por un lado la *Binomial* cuando  $n = 1$  se conoce como *Bernoulli* y es ampliamente utilizada como un modelo binario de elección. Por ejemplo el individuo consume el bien o no. Adicionalmente la binomial tiene la facultad de tener como distribución límite (cuando  $n \rightarrow \infty$ ) un normal, que corresponde a una distribución continúa.

Por otro lado, la *Poisson* despierta un interés especial a los investigadores que trabajan en el área de datos contables como por ejemplo la ocurrencia de accidentes la elección del tamaño de la familia (número de hijos). En Finanzas, la distribución Poisson permite agregar salto no esperados en el precio de un activo, cuando este último es modelado en tiempo continuo.

**Definición 2.1.** *Binomial, sea  $p$  la probabilidad de éxito. Para un experimento que se realiza  $n$  veces se requieren  $x$  éxitos. Así su fdp estará dada por (2.1) y la función en Excel corresponde a `distr.binom(x; n; p; t)`, donde  $t$  es el argumento que indica si se trata de la función de densidad (0) o la acumulada (1).*

$$p(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x} \quad (2.1)$$

Si consideramos  $p = 0,5$ , para 5 ensayos, entonces el histograma (o gráfico que representa las probabilidades para cada evento)<sup>1</sup> será como el de la Figura 2.1, donde el valor de la segunda barra de derecha a izquierda corresponde a `distr.binom(4;5;0,5;0)=0,15625`.

**Proposición 2.1.** *Para  $X \sim B(p)$  demuestre que:*

<sup>1</sup>El código para su construcción en Excel se encuentra en el Anexo C.

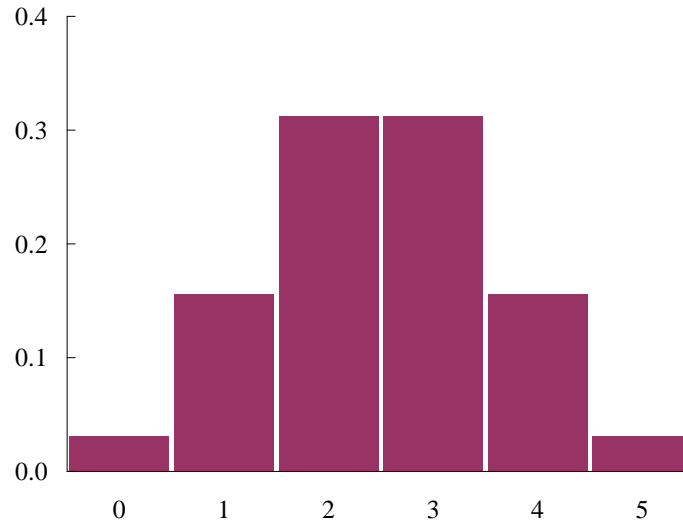


Figura 2.1: Histograma para una Binomial

1.  $m_X(t) = n \log[(1 - p) + \exp(t)p]$ .
2.  $E(X) = np$  y que  $V(X) = np(1 - p)$ .

**Ejemplo 2.1.** Si la probabilidad de un perno defectuoso es 0.1, hallar para la distribución de pernos defectuosos en un total de 400 la media y la desviación estándar.

La media será  $E(X) = np = 400 \cdot 0,1 = 40$  y la varianza  $V(X) = np(1 - p) = 400 \cdot 0,1 \cdot 0,9 = 36$ , luego la desviación estándar es 6.

**Definición 2.2.** Poisson, cuando la va representa el número de eventos independientes que ocurren a una velocidad constante. Su fdp es (2.2) y  $\text{poisson}(x; \lambda; t)$  corresponde a la fórmula de Excel que devuelve la probabilidad de obtener  $x$  éxitos con una media ( $\lambda$ ) conocida. Si  $t$  es 0 devuelve la marginal, con  $t = 1$  se obtiene la probabilidad acumulada. Por ejemplo para  $\lambda = 3,8$  y  $x = 8$  (éxitos) acumula 0,984 o  $\text{poisson}(8; 3,8; 1) = 0,984$ .

$$p(x; \lambda) = \frac{\exp(-\lambda)\lambda^x}{x!} \quad (2.2)$$

Si consideramos  $\lambda = 2$ , entonces el histograma será como el de la Figura 2.2.

**Proposición 2.2.** Para  $X \sim \text{Poisson}(\lambda)$  demuestre que:

1.  $m_X(t) = \lambda[\exp(t) - 1]$ .

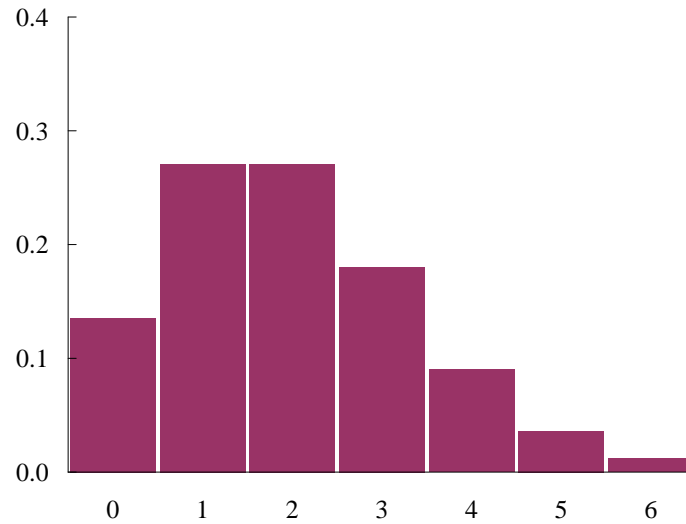


Figura 2.2: Histograma para una Poisson

2.  $E(X) = V(X) = \lambda$ .

El hecho que en esta distribución el primer y segundo momento de esta distribución sean iguales resulta ser un supuesto poco realista en el caso de los datos que podrían ser de nuestro interés (precios de activos, tasas de interés, etc.). Por ejemplo, cuando la varianza supera a la esperanza en escala estamos con un problema de sobredispersión para el cuál se propone como alternativa utilizar un modelo más general como el binomial negativo.

### Distribuciones Continuas

Un punto importante para el análisis posterior es el hecho que la naturaleza de una variable puede ser un tema confuso. Por ejemplo se podría considerar que las tasas de interés que se reciben por un depósito no son variables, pero son numerables. Esto último quiere decir que no existen un evento en que la tasa de interés tome un valor como 0.0399907131532813, sino que sólo se observa un 4%. La conclusión es correcta, pero por simplicidad y por la densidad de estos conjuntos numerables los consideraremos continuos.

**Definición 2.3.** *Normal, proviene del límite de la función de probabilidad de una binomial; fue descubierta por DeMoivre en 1733. Su principal característica es que ajusta muy bien diversos fenómenos estocásticos y posee sólo 2 parámetros, lo que la hace muy útil en estudios empíricos pues agota sólo 2 grados de libertad*

para poder caracterizarla. Su fdp es como sigue:

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right]$$

Toda  $X$  que se distribuye  $N(\mu, \sigma^2)$  puede ser escrita como una  $N(0, 1)$  generando  $Z = (X - \mu)/\sigma$ . Notamos que como se trata de una función lineal de una variable aleatoria, podemos asegurar que es univaluada. Siguiendo la regla del jacobiano y con la función inversa caracterizada como:  $x = \sigma z + \mu$  tenemos que  $J = dx/dz = \sigma$  y así:

$$f_Z = f_X[g^{-1}(z)] \frac{dx}{dz} = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\sigma z + \mu - \mu}{\sigma} \right)^2 \right] \sigma = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right)$$

$Z$  se conoce como la *Normal Estándar* y es la que aparece en las tablas de distribución y su dominio son todos los reales. La función de Excel `distr.norm.inv(p; μ; σ)` devuelve el valor  $x$  de una distribución  $N(\mu, \sigma^2)$  que acumula una probabilidad de  $p$ . Por ejemplo para una  $N(0, 1)$  que acumula una masa de 0.488 el resultado es -0.03 o `distr.norm.inv(0.488, 0, 1)=-0.03`.<sup>2</sup>

**Proposición 2.3.** Probar para una va  $X$  que se distribuye  $N(\mu, \sigma^2)$ :

1.  $m_X(t) = \mu t + 0,5(\sigma t)^2$ .
2.  $E(X) = \mu$  y  $V(X) = \sigma^2$ .

**Definición 2.4.** Beta, se dice que la va  $X$  se distribuye beta si su fdp es como:

$$f(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad \text{donde } \Gamma(\beta) = \int_0^\infty t^{\beta-1} \exp(-t) dt$$

Notemos que los parámetros  $\alpha$  y  $\beta$  modifican la forma de la distribución (Figura 2.3).<sup>3</sup>

**Proposición 2.4.** Probar para una va  $X$  que se distribuye  $Beta(\alpha, \beta)$ :  $E(X) = \alpha/(\alpha + \beta)$  y  $V(X) = \alpha\beta/[(\alpha + \beta)^2(\alpha + \beta + 1)]$ .

**Definición 2.5.** Uniforme, se dice que una va  $X$  está distribuida uniformemente sobre el intervalo  $I = [a, b]$  si su fdp está dada por (2.3) y  $f(x; a, b) = 0$  si  $x \notin I$ .

$$f(x; a, b) = \frac{1}{b - a} \tag{2.3}$$

<sup>2</sup>También está disponible la función `distr.norm.estand.inv()`, pero su aproximación es menos precisa.

<sup>3</sup>La sección C contiene el código para generar la función de densidad de la Beta.

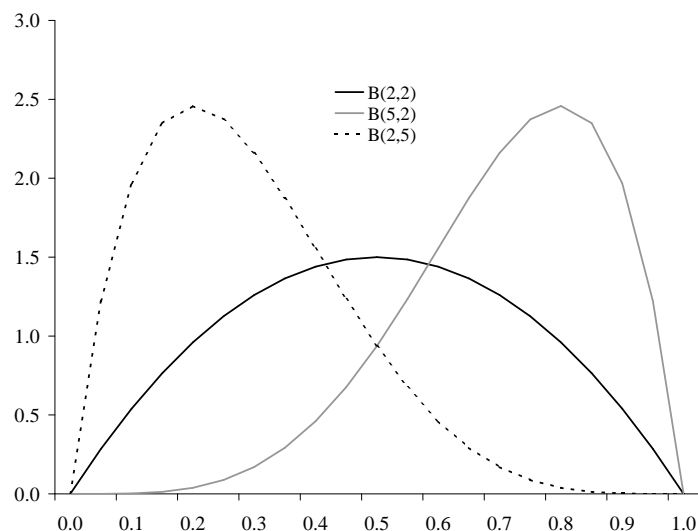


Figura 2.3: Formas de la Beta

**Proposición 2.5.** Probar para una va  $X$  que se distribuye  $U(a, b)$ :  $E(X) = (a+b)/2$  y  $V(X) = (b-a)^2/12$ .

De gran utilidad es la distribución  $U(0, 1)$ , que se encuentra en la función `aleatorio()` de Excel. Debido a que con un valor de esta distribución podemos encontrar el valor de una distribución cualquiera a través de  $F^{-1}(x)$ , de manera absolutamente aleatoria. Por ejemplo para obtener números aleatorios de una  $Beta(2,2)$  basta incorporar en el argumento de probabilidad una  $U(0, 1)$ : `distr.beta.inv(aleatorio();2;2)`.

**Definición 2.6.** Gamma, se dice que la va  $X$  se distribuye gamma si su fdp es como:

$$f(x; \alpha, \theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\theta}\right)$$

**Proposición 2.6.** Probar para una va  $X$  que se distribuye  $Gamma(\alpha, \theta)$ :

1.  $m_X(t) = -\alpha \log(1 - \theta t)$ .
2.  $E(X) = \alpha\theta$  y  $V(X) = \alpha\theta^2$ .

Sus casos particulares son Chi-cuadrado con  $2\alpha$  grados de libertad cuando  $\theta = 2$  como se muestra en la Figura 2.4 y Exponencial Negativa cuando  $\alpha = 1$ . La función asociada en Excel es `distr.gamma.inv(p; \alpha; \theta)` que podemos utilizar para calcular una  $\chi_v^2$  (donde  $v$  son los grados de libertad). Esto corresponde a una  $Gamma(0,5v, 2)$ . Por ejemplo para saber cuál es el valor chi con 1 grado de libertad que acumule 95% a la izquierda encontraremos un valor de 3.841 o `distr.gamma.inv(0,95;0,5;2)=3,841`.

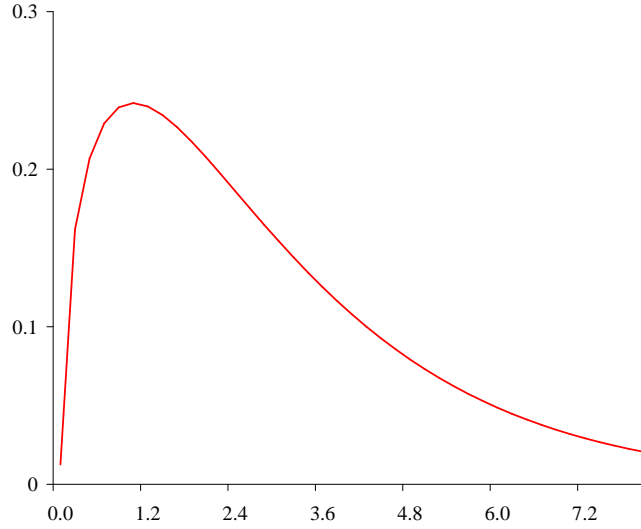


Figura 2.4: Chi-cuadrado con 3 grados de libertad

**Definición 2.7.** Si  $X$  es una normal estándar, entonces  $Y = X^2$  corresponde a una  $\chi_1^2$ .

Considerando la transformación jacobiana para  $y = x^2$ , entonces  $g(x) = x^2$ , luego  $g^{-1}(y) = \sqrt{y}$  y así  $J = dx/dy = 0,5/\sqrt{y}$ . Sin embargo se debe considerar que  $P(Y \leq b) = P(X^2 \leq b) = P(-\sqrt{b} \leq X \leq \sqrt{b})$ , vale decir no tiene una única solución. Pero como  $X$  es simétrica en cero la solución sólo pasa por considerar el doble de la acumulada, entonces:

$$F_Y(b) = 2 \int_0^{\sqrt{b}} f_X[g^{-1}(x)] \left| \frac{dx}{dy} \right| dy = 2 \int_0^{\sqrt{b}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \frac{1}{2\sqrt{y}} dy$$

Retomando el hecho que una  $\chi^2$  corresponde a una gamma y que  $\Gamma(0,5) = \sqrt{\pi}$ , se tiene que (2.4) es una gamma con  $\alpha = 0,5$  y  $\theta = 2$ , es decir una  $\chi_1^2$ .

$$f_Y = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y}{2}\right) \frac{1}{\sqrt{y}} = \frac{1}{\Gamma(0,5)2^{0,5}} y^{0,5-1} \exp\left(-\frac{y}{2}\right) \quad (2.4)$$

**Definición 2.8.** Lognormal: Considere la va  $Y = \exp(X)$ , donde  $X$  se distribuye  $N(\mu, \sigma^2)$ . Notemos que las funciones  $y = \exp(x)$  y  $x = \log y$  son univalueadas, dado que para el cómputo de la probabilidad  $P(Y \leq b) = P[\exp(X) \leq b]$ , sólo existe una solución que corresponde a  $P(X \leq \log b)$ . El Jacobiano de la transformación es  $J = |dx/dy| = 1/y$ , por tanto siguiendo la transformación la fdp de  $Y$  estará determinada por:  $f_Y(y) = f_X[g^{-1}(y)] \cdot J$  como se presenta en (2.5) y corresponde a la fdp de una lognormal.



$$f(y; \mu, \sigma^2) = \frac{1}{\sigma y \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\log y - \mu}{\sigma} \right)^2 \right] \quad (2.5)$$

**Proposición 2.7.** Probar para una va  $X$  que se distribuye  $\text{LogN}(\mu, \sigma^2)$ :

1.  $\log E(X) = E(\log X) + 0,5V(\log X)$
2.  $\log E(X^2) = 2[E(\log X) + V(\log X)]$

## 2. Aplicación: Riqueza como Variable Aleatoria

### Funciones de Utilidad

La existencia de una función de utilidad cardinal representa un mapeo de preferencias ( $\succ$ ) de los consumidores que cumple con los siguientes axiomas :

1. Completitud: Para dos resultados  $x$  e  $y$  se establece que  $x \succ y$  ( $x$  es preferido a  $y$ ) o que  $y \succ x$  ( $y$  es preferido a  $x$ ) o  $x \sim y$  (el individuo es indiferente entre  $x$  e  $y$ ).
2. Transitividad: Si  $x \succ y$  y además  $y \succ z$ , entonces  $x \succ z$ . También es válido para cuando el individuo es indiferente ( $\sim$ ).
3. Independencia Fuerte: Sea  $G(x, z; \pi)$  un juego donde el individuo recibe  $x$  con probabilidad  $\pi$  y  $z$  con probabilidad  $(1 - \pi)$ . Entonces si  $x \sim y$  se cumple que  $G(x, z; \pi) \sim G(y, z; \pi)$ .
4. Medibilidad: Si  $x \succ y \succeq z$  o  $x \succeq y \succ z$ , entonces existe un único  $\pi$  tal que  $y \sim G(x, z; \pi)$ .
5. Ranking: Si  $x \succeq y \succeq z$  y  $x \succeq u \succ z$ , luego si  $y \sim G(x, z; \pi_1)$  y  $u \sim G(x, z; \pi_2)$ . Si ocurre que  $\pi_1 > \pi_2$  entonces  $y \succ u$ , mientras que si  $\pi_1 = \pi_2$  entonces  $y \sim u$ .

El objetivo del individuo es maximizar su utilidad. Pero si la riqueza<sup>4</sup> es una va, entonces debe maximizar la utilidad esperada ( $E[u(w)]$ ), que puede ser calculada a partir de (1.4). La segunda derivada de la función de utilidad nos indica cuál es la posición de los individuos frente al riesgo, denominándose: (1) averso cuando  $u''(w) < 0$ , (2) neutro si es que  $u''(w) = 0$  y (3) amante en el caso de que  $u''(w) > 0$ .

<sup>4</sup>Utilizando el teorema del bien compuesto podemos ampliar las posibilidades de bienes que consume el individuo. Para ellos sólo debemos mantener los precios relativos constantes. Vale decir, con precios relativos fijos podemos considerar el valor de la canasta de bienes del individuo (riqueza).

El análisis económico considera sólo los individuos tipo (1) y (2), notemos que para estos individuos siempre se cumplirá para todo  $w$  que  $E[u(w)] \leq u(E[w])$ <sup>5</sup>. Las funciones de utilidad de la riqueza mayormente usadas para caracterizar individuos que sean aversos al riesgos son: (1) logarítmica ( $u = \log w$ ), (2) raíz cuadrada ( $u = \sqrt{w}$ ) y (3) cuadrática ( $u = aw + bw^2$ , con  $b < 0$ ).

### Medidas de Aversión al Riesgo

Suponiendo que el individuo (no amante) posee una riqueza inicial ( $w_0$ ), con la cual podría contratar un seguro que le permita recibir un flujo sin incertidumbre, el valor de este seguro ( $s$ ) se obtiene de resolver  $u(w_0 - s) = E(u)$ . Consideremos un juego definido como una va  $X$ , que tiene  $E(X) = 0$  y  $V(X) = \sigma^2$  al cual el consumidor está expuesto.

Utilizando una Aproximación de Taylor (AT) de segundo orden<sup>6</sup> tenemos que  $u(w_0 + x) \approx u(w_0) + u'(w_0)x + 0,5u''(w_0)x^2$ . De este modo,  $E[u(w_0 + x)] \approx u(w_0) + 0,5\sigma^2u''(w)$ . Por otro lado, considerando una AT de primer orden para  $u(w_0 - s) \approx u(w_0) + u'(w_0)(-s)$ . Igualando ambas aproximaciones se obtiene en (2.6) el valor del seguro<sup>7</sup>. Notemos que el factor que está ponderando el paréntesis es siempre positivo, lo que da lugar a las definiciones más generales de aversión al riesgo.

$$s = \frac{\sigma^2}{2} \left( -\frac{u''(w_0)}{u'(w_0)} \right) \quad (2.6)$$

**Definición 2.9.** *Medida de Aversión al Riesgo Absoluta:*  $ARA = -u''(w)/u'(w)$ .

**Definición 2.10.** *Medida de Aversión al Riesgo Relativa:*  $ARR = wARA = -wu''(w)/u'(w)$ .

**Ejemplo 2.2.** *Considere un agente que tiene una riqueza ( $w$ ) inicial de 10 y una función de utilidad:  $u = \sqrt{w}$ , existen 2 estados que son: Pérdida [ $P(w = 5) = 0,7$ ] y Ganancia [ $P(w = 15) = 0,3$ ]. Calcule  $E(w)$ ,  $u(E(w))$  y  $E(u)$ . ¿Cuál es el valor del seguro ( $s$ ) que el individuo está dispuesto a pagar para que le den su utilidad esperada?*

*Como se trata de una va discreta, su valor esperado es la probabilidad de cada evento por el valor que toma la variable bajo cada evento, es decir:  $E(w) = p_1w_1 + p_2w_2 = 0,7 \cdot 5 + 0,3 \cdot 15 = 8$ . Este valor genera una utilidad de:  $u(E(w)) = \sqrt{8} = 2\sqrt{2} \approx 2,828$ . De igual modo, la utilidad esperada será:  $E(u) = p_1u(w_1) + p_2u(w_2) = 0,7 \cdot u(5) + 0,3 \cdot u(15) \approx 2,727$ . Por último el seguro puede ser resuelto de:  $10 - s = u^{-1}(E(u)) = (2,727)^2$ . Lo*

<sup>5</sup>Esto nos asegura que la función de utilidad es estrictamente cóncava y se conoce en estadística como la desigualdad de Jensen.

<sup>6</sup>Ver Definición A.8.

<sup>7</sup>Para el caso del agente neutro el valor de este seguro es cero.

que significa que  $s \approx 2,563$ . Intuitivamente el individuo está dispuesto a sacrificar ese monto de su riqueza inicial para evitar la incertidumbre y quedarse (con probabilidad uno) con el valor esperado de su utilidad.

En la Figura 2.5 se encuentran graficadas la función de utilidad (curva), la función de probabilidad binomial (diagonal) y el valor esperado de la utilidad (horizontal). Como vimos este corresponde a  $E(u) = 2,727$  e interseca tanto a la función de utilidad como a la función de probabilidad. De este modo la diferencia medida en el eje de abscisas corresponde al seguro que el individuo toma para eliminar la incertidumbre.

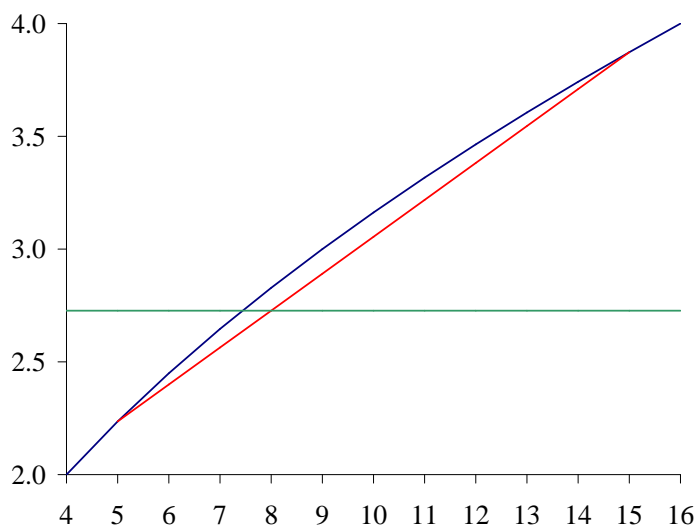


Figura 2.5: Utilidad Esperada y Seguro

### Dominancia estocástica

Consideremos dos procesos aleatorios ( $X$  e  $Y$ ), que tienen las fdp acumulada  $F$  y  $G$  y definen la distribución de la riqueza ( $w$ ). Vale decir, bajo cada proceso la riqueza se distribuye de un modo en particular. Ambos procesos se Dominan Estocásticamente (DE) mutuamente según las siguientes definiciones.

**Definición 2.11.** *DE de Primer Orden:* Se dice que el proceso  $X$  domina a  $Y$  si para todo  $w$  se cumple que  $F(w) \leq G(w)$  y para algún  $w_i$  ocurre que  $F(w_i) < G(w_i)$ .

La anterior definición es aplicable a toda función de utilidad creciente. Intuitivamente, si las fdp acumulada no se cruzan, debieramos observar que  $G$  se ubica por arriba de  $F$  lo que indica que para toda probabilidad el proceso  $Y$  entregaría menores valores de  $w$ .

Cuadro 2.1: Ejemplo Dominancia Estocástica de Primer Orden

$w$	$p(x)$	$p(y)$	$F_x$	$G_y$
-5	0.2	0.2	0.2	0.2
0	0.1	0.2	0.3	0.4
5	0.1	0.2	0.4	0.6
10	0.3	0.2	0.7	0.8
15	0.3	0.2	1	1

Cuadro 2.2: Ejemplo Dominancia Estocástica Segundo Orden

$w$	$p(x)$	$p(y)$	$F_x$	$G_y$	$G - F$	$\sum G - F$
-5	0.2	0.2	0.2	0.2	0	0
0	0.1	0.2	0.3	0.4	0.1	0.1
5	0.1	0.2	0.4	0.6	0.2	0.3
10	0.3	0.2	0.7	0.8	0.1	0.4
15	0.3	0.2	1	1	0	0.4

**Ejemplo 2.3.** Consideremos que la riqueza ( $w$ ) se puede distribuir bajo 2 procesos  $X$  e  $Y$ , como se muestra en la Tabla 2.1. Debido a que  $G_y \geq F_x$ , entonces  $X$  domina de primer orden a  $Y$ .<sup>8</sup> Esto quiere decir que el proceso  $X$  siempre será elegido cuando se coteje con el  $Y$ .

**Definición 2.12.** DE de Segundo Orden: Se dice que el proceso  $X$  domina a  $Y$  si para todo  $w$  se cumple que  $\int_{-\infty}^w [G(v) - F(v)]dv \geq 0$  y para algún  $w_i$  ocurre que  $F(w_i) \neq G(w_i)$ .

En este caso la definición es sólo aplicable para el caso de agentes que son averso, quienes preferirán el proceso  $X$  a  $Y$  por su menor riesgo. Si consideramos que ambos procesos ofrecen el mismo valor esperado para la riqueza, por lo que las fdp acumuladas se cruzan por lo menos en un punto, entonces el proceso  $Y$  presenta una mayor dispersión de resultados.

**Ejemplo 2.4.** Continuando con el Ejemplo 2.3, confirmamos que  $X$  DE también en segundo orden a  $Y$ , como se aprecia en el Cuadro 2.2.

---

<sup>8</sup>En este caso la definición coincide con el valor medio de la riqueza ya que  $E(w|X) = 7$  y  $E(w|Y) = 5$ .

### 3. Distribuciones Conjuntas

#### Definiciones

Establecemos algunas definiciones para cuando existen dos variables aleatorias. El análisis puede extenderse a más variables aleatorias sin mayores dificultades.

**Definición 2.13.** Sean  $X$  y  $Y$  dos variables continuas. Si existe una función  $f(x, y)$  tal que la probabilidad conjunta sea como (2.7) donde la fdp cumple con (2.8)

$$P(a < X < b, c < Y < d) = \int_a^b \int_c^d f(x, y) dy dx \quad (2.7)$$

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \quad (2.8)$$

**Definición 2.14.** Distribuciones marginales: se obtiene a partir de la distribución conjunta cuando se “barre” la otra variable. En el caso discreto sería como (2.9) y como (2.10) para cuando la variable es continua. Por tanto el cálculo de la probabilidad para la variable  $X$  sería:  $P[X \leq x] = F_X(x) = \int_{-\infty}^x f_X(t) dt$ .

$$p_X(x) = \sum_Y p(x, y) \quad (2.9)$$

$$f_X(x) = \int_Y f(x, y) dy \quad (2.10)$$

**Ejemplo 2.5.** Suponga que  $X$  e  $Y$  se distribuyen conjuntamente y tienen una fdp nula salvo cuando  $a < x < b$  y  $c < y < d$ , donde  $f(x, y) = mx + ny$ . Determine cuál es el valor de  $m$  en función de  $n$  y de los límites de los intervalos de cada variable. Considere  $I = F(b, d)$ , que al integrar con respecto a  $y$ :

$$I = \int_a^b \int_c^d (mx + ny) dy dx = \int_a^b \left[ mx(d - c) + \frac{n}{2} (d^2 - c^2) \right] dx = \left( \frac{d - c}{2} \right) \int_a^b [2mx + n(d + c)] dx$$

Ordenando e integrando con respecto a  $x$ :

$$I = \left( \frac{d - c}{2} \right) [m(b^2 - a^2) + n(b - a)(d + c)] = \left[ \frac{(a - b)(c - d)}{2} \right] [m(a + b) + n(c + d)]$$

Haciendo  $I = 1$  tenemos que la solución es:

$$m = \frac{1}{(a+b)} \left[ \frac{2}{(a-b)(c-d)} - n(c+d) \right]$$

**Definición 2.15.** *Valor esperado: corresponde a valorar una función de la va, considerando una función real  $g(x, y)$ , sería como (2.11)*

$$E[g(x, y)] = \int_X \int_Y g(x, y) f(x, y) dy dx \quad (2.11)$$

Sin embargo, una función interesante corresponde al cálculo de momentos, los más interesante son el  $r$ -ésimo momento centrado en cero (2.12) y el momento centrado en la media cruzado (2.13). En este último cuando  $r = s = 1$ , lo llamaremos  $Cov(X, Y)$ .

$$E[X^r] = \int_X x^r f_X(x) dx \quad (2.12)$$

$$E[(X - \mu_X)^r (Y - \mu_Y)^s] = \int_X \int_Y (x - \mu_X)^r (y - \mu_Y)^s f(x, y) dy dx \quad (2.13)$$

**Proposición 2.8.** *Probar que:*

1.  $Cov(X, Y) = E(XY) - E(X)E(Y)$

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

**Definición 2.16.** *Se llama distribución condicional a la fdp que sigue a Bayes como:  $f(y|x) = f(x, y)/f_X(x)$ . Si  $X$  y  $Y$  son independientes entonces  $f(y|x) = f_Y(y)$ . Esto implica que  $f(x, y) = f_X(x)f_Y(y)$ .*

**Proposición 2.9.** *Si  $X$  y  $Y$  son estadísticamente independientes entonces:*

1.  $P(a < X < b, c < Y < d) = P(a < X < b) \cdot P(c < Y < d)$ .

2.  $E[g(X)g(Y)] = E[g(X)]E[g(Y)]$ . Con  $g()$  una función univaluada.

**Definición 2.17.** *Sean  $\{X_i\}_{i=1}^n$  una ma con un fdp conjunta  $f(x_1, x_2, \dots, x_n)$ . Entonces para cualquier set  $\{a_i\}_{i=1}^n$ :*

$$E[\sum_{i=1}^n a_i X_i] = \sum_{i=1}^n [a_i E(X_i)]$$

$$V[\sum_{i=1}^n a_i X_i] = \sum_{i=1}^n a_i^2 V(X_i) + \sum_{j=1, j \neq i}^n \sum_{i=1}^n a_i a_j \cdot Cov(X_i, X_j).$$

**Ejemplo 2.6.** Considere las va  $X_1$  y  $X_2$  que se distribuyen conjuntamente y una nueva va  $Y = a_1X_1 + a_2X_2$ . Tomando por definición que  $\mu_1 = E(X_1)$  y  $\mu_2 = E(X_2)$ . Entonces: (1)  $E(Y) = a_1\mu_1 + a_2\mu_2$  y (2)  $V(Y) = a_1^2V(X_1) + a_2^2V(X_2) + 2a_1a_2Cov(X_1X_2)$ .

1.  $E()$  es un operador lineal, entonces:  $E(a_1X_1 + a_2X_2) = a_1E(X_1) + a_2E(X_2) = a_1\mu_1 + a_2\mu_2$ .
2. Por otro lado:  $V(Y) = E\{[Y - E(Y)]^2\} = E\{[a_1X_1 + a_2X_2 - a_1\mu_1 - a_2\mu_2]^2\}$ .
3. Conmutando y agrupando:  $V(Y) = E\{[(a_1X_1 - a_1\mu_1) + (a_2X_2 - a_2\mu_2)]^2\}$ .
4. Elevando:  $V(Y) = E\{(a_1X_1 - a_1\mu_1)^2 + (a_2X_2 - a_2\mu_2)^2 + 2(a_1X_1 - a_1\mu_1)(a_2X_2 - a_2\mu_2)\}$ .
5. Factorizando:  $V(Y) = E[a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + 2a_1a_2(X_1 - \mu_1)(X_2 - \mu_2)]$ .
6. Aplicando  $E()$ :  $V(Y) = a_1^2E[(X_1 - \mu_1)^2] + a_2^2E[(X_2 - \mu_2)^2] + 2a_1a_2E[(X_1 - \mu_1)(X_2 - \mu_2)]$ .
7. Usando la definición de  $V()$ :  $V(Y) = a_1^2V(X_1) + a_2^2V(X_2) + 2a_1a_2Cov(X_1X_2)$ .

**Teorema 2.1.** Si  $R$  y  $S$  son dos variables aleatorias con fdp conjunta  $f(r, s)$ . Si existe  $q = g(r, s)$  y  $\tilde{s} = h(q, r)$ , tal que puede ser computable  $J = dh/dq$  entonces:

$$f_q = \int_R f(r, \tilde{s})|J|dr$$

**Ejemplo 2.7.** Considerando  $X$  e  $Y$  con fdp conjunta  $f(x, y)$  tal que  $Z = X + Y$  computar  $f_z$ . Definiendo  $\tilde{y} = z - x$  y  $x = x$  se tiene  $J = 1$ . Así:

$$f_z = \int_X f(x, z - x)dx$$

**Ejemplo 2.8.** Suponga  $X_1$  y  $X_2$  son dos va con distribución conjunta como:  $f(x_1, x_2) = 6x_2$  si  $0 < x_2 < x_1 < 1$  y nula para los otros casos. Encontrar: (1)  $f_{2|1}(x_2|x_1)$  y (2)  $E(X_2|x_1)$ . Considere ahora, que  $Y = E(X_2|X_1)$  es una nueva va. Construya su fdp y compute  $E(Y)$  y  $V(Y)$

La fdp marginal de  $X_1$  es nula salvo en el intervalo  $0 < x_1 < 1$  que corresponde a:

$$f_1(x_1) = \int_0^{x_1} 6x_2 dx_2 = 3x_1^2$$

La fdp condicional de  $X_2$  dado  $X_1 = x_1$  es cero salvo en el intervalo  $0 < x_2 < x_1$ , donde corresponde a  $f_{2|1} = (6x_2)/(3x_1^2) = 2x_2/x_1^2$ . Finalmente la esperanza condicional puede ser obtenida como sigue:

$$E(X_2|x_1) = \int_0^{x_1} x_2 \left( \frac{2x_2}{x_1^2} \right) dx_2 = \frac{2}{3}x_1$$

Para encontrar los momentos de  $Y$  debemos considerar que  $Y = E(X_2|X_1) = (2/3)X_1$ , y  $0 < x_1 < 1$ . Esto implica que la función de distribución será  $G(y)$  y  $0 \leq y < 2/3$ .

$$G(y) = P(Y \leq y) = P\left(X_1 \leq \frac{3y}{2}\right) = \int_0^{3y/2} 3x_1^2 dx_1 = \frac{27y^3}{8}$$

Por supuesto,  $G(y) = 0$  si  $y < 0$  y  $G(y) = 1$  si  $2/3 < y$ . Finalmente considerando que la fdp de  $Y$  es  $g(y) = (81/8)y^2$ , entonces se tiene que:

$$E(Y) = \int_0^{2/3} y \left( \frac{81y^2}{8} \right) dy = \frac{1}{2}$$

$$V(Y) = \int_0^{2/3} y^2 \left( \frac{81y^2}{8} \right) dy - \frac{1}{4} = \frac{1}{60}$$

**Ejemplo 2.9.** Sean  $X$  y  $Y$  va que se distribuyen conjuntamente. Suponiendo que  $E(Y|X) = a + bX$ , encuentre los valores de  $a$  y  $b$ . Utilizando el resultado anterior y asumiendo que  $V(Y|X) = k(x)$ , determine  $E[k(x)]$ .

El supuesto planteado equivale a una esperanza cuya fdp es condicional que puede ser escrita como la conjunta sobre la marginal utilizando Bayes. Ponderando en ambos lados por  $f(x)$  llegamos a:

$$\int yf(x, y)dy = (a + bx)f(x) \tag{2.14}$$

$$E(Y|X) = a + bX = \int yf(y|x)dy = \int y \frac{f(x, y)}{f(x)} dy \Rightarrow \int yf(x, y)dy = (a + bx)f(x)$$

Algunos hechos a considerar son:

- Integrando (1) con respecto a  $x$ :  $E(Y) = a + bE(X)$ .
- Ponderando (1) por  $x$  e integrando con respecto a  $x$ :  $E(XY) = aE(X) + bE(X^2)$ .
- $Cov(X, Y) = E(XY) - E(X)E(Y) = \rho\sqrt{V(X)V(Y)}$ .

Sea  $\mu = E()$  y  $\sigma^2 = V()$ , el problema pasa por resolver el sistema de ecuaciones:  $\mu_y = a + b\mu_x$  y



$\rho\sigma_x\sigma_y + \mu_x\mu_y = a\mu_x + b(\sigma_x^2 + \mu_x^2)$ . Los resultados son:  $a = \mu_y - (\rho\sigma_y/\sigma_x)\mu_x$  y  $b = \rho\sigma_y/\sigma_x$ .

Consideramos la definición de  $V(Y|X)$ , ponderamos por  $f(x)$  e integramos con respecto a  $x$ :

$$V(Y|X) = \int \left[ (y - \mu_y) - \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x) \right]^2 \frac{f(x, y)}{f(x)} dy \Rightarrow E[V(Y|X)] = E[k(X)]\sigma_y^2(1 - \rho^2)$$

### Co-movimientos

La forma como conjuntamente se mueven dos variables nos permite tener una inspección rápida de la posible relación que exista entre dichas variables. Dos errores comúnmente interpretados son: (1) la causalidad de las variables (si  $X$  toma un valor entonces  $Y$  toma un valor definido) y (2) establecer que una correlación cero implica que las variables son independientes. En efecto, las distribuciones conjuntas no manifiestan necesariamente un orden de como los sucesos ocurren y por otro lado, los índices de correlación suelen ser aproximaciones sólo a dependencias lineales de las variables.

**Definición 2.18.** *Correlación Lineal o de Pearson ( $\rho$ ): permite medir el grado de relación lineal que existe entre 2 va. Este coeficiente se ubica en el intervalo  $[-1, 1]$  y se define como:*

$$\rho = \frac{Cov(X, Y)}{\sqrt{[V(X)V(Y)]}} = \frac{\int \int xyf(x, y)dxdy - \int xf_X(x)dx \int yf_Y(y)dy}{\sqrt{\int x^2f_X(x)dx - [\int xf_X(x)dx]^2} \sqrt{\int y^2f_Y(y)dy - [\int yf_Y(y)dy]^2}}$$

Notemos que cuando dos va son independientes entonces su coeficiente de correlación es cero, pero no al revés. Consideremos el caso de que la distribución conjunta de  $X$  e  $Y$  es discreta con el conjunto  $R = \{(0, 1); (0, -1); (1, 0); (-1, 0)\}$  de resultado posibles. Suponiendo que cada uno tiene igual probabilidad de ocurrencia (1/4) podemos calcular  $E(X) = (1/4)(0 + 0 + 1 - 1) = 0$  y  $E(Y) = (1/4)(1 - 1 + 0 + 0) = 0$ . Además se cumple que  $X_i Y_i = 0$ , por lo que  $E(XY) = 0$  y de este modo  $Cov(X, Y) = 0$  que implica que  $\rho = 0$ .

Del ejercicio anterior podemos notar que  $\rho = 0$  puede ocurrir cuando las va no son independientes. En particular cualquier distribución conjunta que proviene de una combinación no lineal de las va no será capturada por  $\rho$ . Su estimación numérica puede computada con la función `coef.de.correl()`.

**Teorema 2.2.** *Si  $X_1$  y  $X_2$  son independientes, entonces  $\rho = 0$ .*

Prueba: Si  $X_1$  y  $X_2$  son ind,  $\implies E(X_1 X_2) = E(X_1)E(X_2) \implies Cov(X_1, X_2) = 0 \implies \rho = 0$ .

**Definición 2.19.** Spearman<sup>9</sup>: es similar a  $\rho$  presentado en la definición anterior pero sustituye el valor de las variables por su probabilidad acumulada. También puede ser entendido como el orden que manifiestan distintos procesos aleatorios. Su definición por tanto es:

$$\zeta = \frac{\int \int F_X(x)F_Y(y)f(x,y)dxdy - \int F_X(x)f_X(x)dx \int F_Y(y)f_Y(y)dy}{\sqrt{\int F_X^2(x)f_X(x)dx - [\int F_X(x)f_X(x)dx]^2} \sqrt{\int F_Y^2(y)f_Y(y)dy - [\int F_Y(y)f_Y(y)dy]^2}}$$

Notemos que  $\int F_X(x)f_X(x)dx = \int_0^1 udu = 1/2$  y  $\int F_X^2(x)f_X(x)dx = \int_0^1 u^2du = 1/3$ . Simplificando:

$$\zeta = 12 \int \int F_X(x)F_Y(y)f(x,y)dxdy - 3$$

Ambos índices de correlación ( $\rho$  y  $\zeta$ ) bajo la nula de que son cero poseen una distribución  $t_{n-2}$ . Esto es si  $r$  corresponde a la estimación numérica del índice entonces  $t \sim t_{n-2}$ :

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Sin embargo, cuando  $H_0 : \rho = \rho_0$ , se utiliza la transformación de Fisher (correspondiente a la función `fisher()`) que implica que el estadístico  $z$  se distribuye  $N(\mu_z, \sigma_z^2)$ , donde  $\mu_z = 0,5[\log(1 + \rho_0) - \log(1 - \rho_0)]$  y  $\sigma_z^2 = (n - 3)^{-1}$ .

$$z = \frac{1}{2} \log \left( \frac{1+r}{1-r} \right)$$

**Definición 2.20.** Kendall: mide el nivel de concordancia y discordancia de dos variables. Por ejemplo al tener 2 pares ordenados de una distribución conjunta como  $(x_1, y_1)$  y  $(x_2, y_2)$ , se dirá que son concordantes si ambos miembros de un par son mayores que los respectivos miembros del otro par. Por otro lado, los pares serán discordantes si  $x_1 > x_2$  y  $y_1 < y_2$  o  $x_1 < x_2$  y  $y_1 > y_2$ . Se define como sigue:

$$\tau_K = P[(x_1 - x_2)(y_1 - y_2) > 0] - P[(x_1 - x_2)(y_1 - y_2) < 0]$$

Alternativamente:  $\tau_K = 4 \int \int F(x,y)f(x,y)dxdy - 1$ .

**Definición 2.21.** Copula ( $C$ ): es una función que permite que dos variables definidas independientes conformen una distribución conjunta. De este modo,  $C : [0, 1]^n \rightarrow [0, 1]$ , lo que equivale a decir que  $F(x, y) = C[F_X(x), F_Y(y)]$ .

---

<sup>9</sup>Ver código en C.1.3

**Distribuciones de Interés**

**Definición 2.22.** Se dice que las va  $X_1$  y  $X_2$  tiene una distribución Normal bivariada si su fdp es  $f(x_1, x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp(-0,5q)$ , donde  $q$  se define como:

$$q = \frac{1}{1-\rho^2} \left[ \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{x_1 - \mu_1}{\sigma_1} \right) \left( \frac{x_2 - \mu_2}{\sigma_2} \right) + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right].$$

Notemos que  $\mu_i$  corresponden a  $E(X_i)$ ,  $\sigma_i^2$  a  $V(X_i)$  y  $\rho$  corresponde al coeficiente de correlación.

**Teorema 2.3.** Si  $X_1$  y  $X_2$  se distribuyen normal bivariado. Entonces se cumple que si  $\rho = 0 \iff X_1$  y  $X_2$  independientes.

Prueba: Notemos que utilizando el Teorema 2.2 sólo debemos probar una parte de la igualdad. Por otro lado, al hacer  $\rho = 0$  en la Definición 2.22 y utilizando un poco de álgebra se llega a:  $f(x_1, x_2) = f(x_1)f(x_2)$ , que es el resultado de independencia.

**Definición 2.23.** Distribución Normal Multivariada: corresponde a (2.15) donde  $x$  corresponde a un vector que incluye a las  $n$  va con  $\mu$  vector de esperanza y  $\Sigma$  matriz de varianza y covarianza.

$$f(x) = \frac{1}{|\Sigma|\sqrt{(2\pi)^n}} \exp \left[ -\frac{1}{2}(x - \mu)' \Sigma^{-1}(x - \mu) \right] \quad (2.15)$$

**Teorema 2.4.** Si  $X_1$  y  $X_2$  son dos va independientes y cada una tiene una distribución chi-cuadrada con  $v_1$  y  $v_2$  grados de libertad respectivamente, entonces:  $Y = X_1 + X_2$  también tiene una distribución chi-cuadrada con  $v_1 + v_2$  grados de libertad.

Prueba: Notemos que  $m_Y(t) = \log E[\exp(tY)]$  y dado que  $Y = X_1 + X_2$ . Además  $X_1$  y  $X_2$  son independientes, entonces:  $E[g(X_1)g(X_2)] = E[g(X_1)]E[g(X_2)]$ . Así  $m_Y = m_{X_1} + m_{X_2}$ . Utilizando la FGM de una chi-cuadrado se demuestra que  $Y$  también es una chi-cuadrado y sus grados de libertad son  $v_1 + v_2$ :

$$m_Y(t) = \underbrace{\left[ -\frac{v_1}{2} \log(1-2t) \right]}_{m_{X_1}(t)} + \underbrace{\left[ -\frac{v_2}{2} \log(1-2t) \right]}_{m_{X_2}(t)} = -\frac{(v_1 + v_2)}{2} \log(1-2t)$$

**Teorema 2.5.** Sea  $Z$  una va normal estándar y  $X$  una va chi-cuadrada con  $v$  grados de libertad. Si  $Z$  y  $X$  son independientes, entonces la va:  $T = Z/\sqrt{X/v}$  tiene una distribución  $t$  de Student con  $v$  grados de libertad y su fdp es (2.16). De la Figura 2.6 notaremos que la distribución  $t$  es similar a una  $Z$ . Pero para

grados de libertad pequeños la  $t$  tiene las colas más altas: este resultado será útil cuando se desea hacer inferencia sobre un parámetro de la distribución. La función correspondiente en Excel es `distr.t.inv()`, que devuelve el valor  $t$  de una distribución  $t_v$  donde  $v$  son los grados de libertad. Acá el argumento de la **probabilidad** corresponde a la masa que se reparte simétricamente entre ambas colas de la distribución y es contabilizado de derecha a izquierda. Por ejemplo para una probabilidad de 5% con 7 grados de libertad el resultado es 2.365, valor correspondiente a un  $t$  de tabla que acumula a la derecha 2.5% de probabilidad.<sup>10</sup>

$$f(t; v) = \frac{\Gamma[(v + 1)/2]}{\Gamma(v/2)\sqrt{\pi v}} \left[ 1 + \left( \frac{t^2}{v} \right) \right]^{-\frac{v+1}{2}} \quad (2.16)$$

Podemos definir  $\tilde{z} = t\sqrt{x/v}$  y  $x = x$ , así  $J = \sqrt{x/v}$ . Adicionalmente, se tiene que  $X$  y  $Z$  son independientes.

Por lo tanto,  $f(x, \tilde{z}) = f(x)f(\tilde{z})$  esto es:

$$f(x)f(\tilde{z}) = \left[ \frac{1}{\Gamma(v/2)2^{v/2}} x^{v/2-1} \exp\left(-\frac{x}{2}\right) \right] \left[ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\tilde{z}^2\right) \right]$$

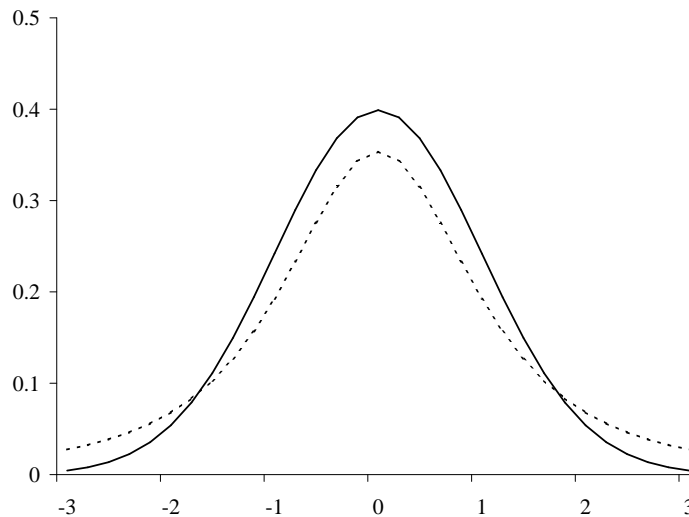


Figura 2.6: Distribución  $N(0, 1)$  y  $t_3$  (línea punteada)

**Teorema 2.6.** Sean  $X_1$  y  $X_2$  son dos va independientes y cada una tiene una distribución chi-cuadrada con  $v_1$  y  $v_2$  grados de libertad respectivamente, entonces la va  $F = \frac{X_1/v_1}{X_2/v_2}$  tiene una distribución  $F$  con  $v_1$  y  $v_2$  grados de libertad en el numerador y denominador, respectivamente. La función de Excel asociada es `distr.f.inv()` que devuelve el valor  $f$  de una  $F(p, q)$  donde  $p$  y  $q$  son los grados de libertad del numerador

<sup>10</sup>Cuando  $v = 1$  la distribución se conoce como Cauchy y tiene la característica que no posee esperanza.

y denominador respectivamente. En este caso el área de probabilidad se acumula de derecha a izquierda. Por ejemplo para un área de 0.1 a la derecha con  $p = 2$  y  $q = 5$  el valor es de 3.78.

**Proposición 2.10.** Derive las fdp de las distribuciones  $t$  y  $F$ .

## 4. Aplicación: Mercados Financieros

### Definiciones Básicas

Comenzaremos esta sección con definiciones básicas, para luego insertar la administración de un portafolio y la frontera eficiente.

**Definición 2.24.** Un activo constituye un bien que otorga a su portador un beneficio, que asumiendo una función de utilidad monetaria lo expresaremos en pesos. Por otro lado, impondremos que todos los activos pueden ser caracterizados con 2 parámetros: el riesgo y el retorno<sup>11</sup>.

**Definición 2.25.** “Valor del Tiempo”: siguiendo con el dicho del tiempo es oro diremos que el pago por una unidad de tiempo corresponde a la tasa de interés ( $R$ ). De este modo para hacer comparables 2 unidades monetarias que se ubican a una unidad de tiempo de distancia es necesario corregirlas con  $R$ . Por ejemplo, un depósito a 30 días valdrá el próximo mes su valor inicial más un premio:  $D_{t+1} = D_t(1 + R)$ .

**Definición 2.26.** Interés: Básicamente existen 2 categorías de interés (1) Lineal, lo que significa que para llevar una unidad monetaria dos períodos adelante los hacemos como  $D_{t+2} = D_t(1 + 2R)$  y (2) Compuesto o Capitalizable, en el cual se devengan interés sobre interés, lo que pone el caso anterior como  $D_{t+1} = D_t(1 + R)^2$ .

**Definición 2.27.** Tasa Capitalizable Continua (TCC): se define como  $r = \log(1 + R)$ .

Notemos que utilizando una AT(1) podemos mostrar que cuando  $R$  es pequeña entonces  $r \approx R$ . Esto es:  $f(R) \approx f(R_0) + f'(R_0)(R - R_0)$ . Consideremos para el caso de  $R_0 = 0$  (debido a que una tasa de 5% sería 0.05) y  $f(R) = \log(1 + R)$ , sería como:

$$r = \log(1 + R) \approx \log(1 + R_0) + \left( \frac{1}{1 + R_0} \right) (R - R_0) = R$$

<sup>11</sup>En realidad el supuesto que se realiza es que sólo 2 momentos son suficientes para caracterizar un activo.

**Definición 2.28.** *Tasa Real: es aquella tasa que incluye el efecto inflacionario. Vale decir para una tasa nominal  $I$  y una inflación  $\Pi$ , la relación sería como (2.17). considerando TCC la relación se transforma en  $r = i - \pi$ . Si las tasas son pequeñas, entonces se puede decir que  $R = I - \Pi$ .*

$$(1 + R) = \frac{(1 + I)}{(1 + \Pi)} \quad (2.17)$$

**Definición 2.29.** *El flujo de períodos valorizados en pesos de hoy corresponde al concepto de “Valor Actual Neto”(VAN). Este es definido como el exceso que se produce entre el valor presente (VP) de todos los flujos futuros a las tasas correspondientes y la inversión inicial ( $I$ ), vale decir:  $VAN = -I + VP$ . Para entender el concepto de VP, haremos un ejemplo donde suponemos que un bono entrega cupones semestrales fijos durante 3 períodos, si descontamos cada período a la misma tasa el VP, o máximo precio a pagar por ese bono, será como (2.18). En el caso de una perpetuidad la relación es más simple ( $P = C/R$ ).*

$$P = \frac{C}{1 + R} + \frac{C}{(1 + R)^2} + \frac{C}{(1 + R)^3} \quad (2.18)$$

Pero en el caso de tener  $n$  pagos el VP podrá ser escrito como (2.19). El VAN por su parte es utilizado en la evaluación de proyectos. Por ejemplo si se compra un taxi se puede obtener el flujo en cada período del tiempo (cada mes) y obtener el VP de acuerdo a la tasa de descuento relevante. Al considerar la inversión inicial obtendremos el VAN. Queda claro que todo proyecto que presente un VAN positivo debiera realizarse, sin embargo por el principio de escasez se puede dar prioridad a los proyectos más rentables. Una medida de rentabilidad es la TIR.

$$P = \frac{C}{R} \left[ 1 - \frac{1}{(1 + R)^n} \right] \quad (2.19)$$

**Definición 2.30.** *“Tasa Interna de Retorno”(TIR) corresponde al valor de una tasa fija (relevante) que se hace que VAN sea cero. Notemos que este problema corresponde a la solución de encontrar las raíces de un polinomio.*

### Precio de un Activo

Partiendo que un activo  $i$  se valoriza en  $t$ , por tanto su precio ( $P_{it}$ ) se valoriza como sigue (2.20), donde la variable  $D_{t+1}$  es el dividendo esperado, se utiliza en el caso de una acción, sin embargo lo más importante en la valoración corresponde al precio futuro.

$$P_{i,t} = \frac{P_{i,t+1} + D_{t+1}}{1 + R_{it}} \quad (2.20)$$

Por otro lado  $R_{it}$  corresponde a la tasa de retribución que se le exige al activo  $i$ . Dicha tasa tiene una distribución con una media (retorno esperado) y una varianza (riesgo). Es usual en Finanzas mencionar este último en términos de su desviación estándar.

Supondremos, además que existe un activo libre de riesgo con rentabilidad  $F$ , relativamente pequeña. En términos prácticos este activo obedecería a un papel de Tesorería o del Banco Central<sup>12</sup>.

### Portafolios

Siguiendo el dicho popular “es malo poner todos los huevos en la misma canasta”, así también es útil diversificar una inversión en diferentes activos. Por ejemplo si contamos con 1 unidad monetaria y decidimos invertirlo en 2 activos ( $X_1$  y  $X_2$ ) que tienen igual retorno esperado ( $R_1 = R_2 = R$ ) y riesgo ( $\sigma_1 = \sigma_2 = \sigma$ ), de acuerdo a una proporción  $w$ . Nuestro retorno esperado para el portafolio ( $R_p$ ) sería como (2.21) y el riesgo asociado sería (2.22)

$$R_p = E[wX_1 + (1 - w)X_2] = wR_1 + (1 - w)R_2 \quad (2.21)$$

$$\sigma_p^2 = V[wX_1 + (1 - w)X_2] = w^2\sigma_1^2 + 2w(1 - w)\sigma_{12} + (1 - w)^2\sigma_2^2 \quad (2.22)$$

Consideremos la Definición 2.18 y reemplazando en (2.22) quedamos con (2.23), donde el riesgo será menor si  $\rho_{12}$  es negativo. Lo interesante es el intervalo  $[0, 1]$ , ya que en el caso de que  $\rho_{12} = 1$  entonces  $\sigma_p = \sigma$ .

$$\sigma_p^2 = w^2\sigma_1^2 + 2w(1 - w)\rho_{12}\sigma_1\sigma_2 + (1 - w)^2\sigma_2^2 \quad (2.23)$$

### Frontera de Eficiencia

Como vimos en la sección anterior es importante diversificar un portafolio para disminuir el riesgo. Sin embargo, vemos que para una cartera con 2 activos sólo podemos controlar un parámetro, pero nuestro interés son 2 variables: retorno y riesgo. En este sentido no se cumple que exista una herramienta para cada

<sup>12</sup>Los depósitos a plazo y las cuentas de ahorro se encuentran acotada con el límite de garantía estatal a los depósitos.

problema y lo que podemos decir es que optimizar la relación dentro de una frontera.

Supongamos que contamos con  $k$  activos y sus retorno están simbolizados en el vector  $r$ , y su matriz de varianza y covarianza queda definida por  $\Omega$ <sup>13</sup>. Entonces el problema se reduce a encontrar el vector  $w$  con  $k - 1$  asignaciones para lograr maximizar  $w'E(R)$ , sujeto a un valor de riesgo  $\sigma_p^2 = w'E(\Omega)w = \sigma_0^2$ . Con lo anterior se puede formar una frontera de asignación relacionando a diferentes niveles de riesgo distintos retornos esperados.

### Valor en Riesgo

El valor en Riesgo corresponde a un estimado, con cierto grado de confianza, de cuanto puede valor puede perder un portafolio de inversión en un determinado tiempo.

La composición del portafolio puede ser variada, a medida que se incluyen más activos se hace más complejo el análisis pero se disminuyen los riesgos debido a la diversificación. Los grados de confianza típicamente utilizados son el 95 % y 99 %. El horizonte de tiempo usualmente es un día, una semana o un mes. Notamos que la metodología asume que el inversionista no reaccionará recomponiendo el portafolio.

Esta metodología se basa en la probabilidad acumulada. Si tomamos  $\alpha$  como la probabilidad de que la variable de interés ( $X$ ) se ubique por debajo de un cierto valor ( $-L$ ):  $P(X \leq -L) = \alpha$ . En este caso  $X$  representa el retorno de una unidad monetaria sobre un portafolio determinado. Para simplificar el análisis asumiremos que este retorno es una variable aleatoria distribuida normal con cero media y volatilidad  $\sigma\sqrt{h}$ , es decir  $X \sim N(0, \sigma^2h)$ , donde  $h$  representa la constante de anualización, con ello tenemos que

$$P\left(\frac{X}{\sigma\sqrt{h}} \leq -\frac{L}{\sigma\sqrt{h}}\right) = N\left(-\frac{L}{\sigma\sqrt{h}}\right) = \alpha$$

Esto implica que  $L = -\sigma\sqrt{h}N^{-1}(\alpha)$ , es decir el valor límite podría ser computado sobre la base de la volatilidad del portafolio y de un cierto grado de confianza, el cual es caracterizado por  $\alpha$ . Este último corresponde a la cola de la distribución, usualmente  $\alpha = 0,05$  con ello  $N^{-1}(0,05) = -1,645$ , así  $L = 1,645\sigma\sqrt{h}$ . Con ello el Valor en Riesgo (Value at Risk, VaR) del portafolio, se obtiene tomando el valor del portafolio ( $W$ ), así  $VaR = L \times W$ .

La metodología del Valor en Riesgo (VaR) puede aplicarse a otros casos financieros donde la incertidumbre

<sup>13</sup>Notemos que por ejemplo el caso de 2 activos:

$$\Omega = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$



es generada por otra variable aleatoria. Por ejemplo Neriz y Miranda (2005) utilizan el VaR para analizar el efecto de la tasa de interés en las carteras de inversión de Compañías de Seguros de Vida en Chile. A través del Método Delta, la volatilidad de las carteras puede ser aproximada por la sensibilidad del VPN de las carteras multiplicado por la volatilidad de la tasa de interés. Al igual que el precio de un bono ( $P$ ) la sensibilidad del VPN a la tasa de interés es proporcional a la duración modificada ( $D_m$ ). Es decir  $\Delta \%P = -D_m \Delta i \Leftrightarrow \Delta P = -D_m P \Delta i$ .

## 5. Modelo Binomial

El modelo binomial es un sencillo método para caracterizar la dinámica del precio de un activo. El modelo fue originalmente planteado para la valoración de opciones, es decir, como supuesto de como se comporta el precio del activo es posible valorar las opciones. Debido a su simplicidad ha tenido éxito en este campo, debido a que permite de manera muy sencilla la valoración de opciones complejas, las cuales pueden no presentar soluciones factibles bajo la modelación del precio del activo en tiempo continuo. Por lo demás su uso no requiere de conocimiento de cálculo estocástico.

Por ejemplo, la valoración de una opción americana sobre un activo subyacente que es un promedio discreto del precio de una acción que paga dividendos discretos es un gran problema para resolver en tiempo continuo. Si agregamos que el precio de ejercicio crece a una tasa constante y que los factores de descuento intertemporales son también estocásticos entonces el tema se complica mucho más. Algunos autores definen el modelo binomial como un modelo básico de entendimiento de la valoración de opciones. Sin embargo, su aporte puede ser muy importante si consideramos que los modelos de tiempo continuo deben en la mayoría de los casos ser discretizados para su análisis numérico.

### Marco Teórico

Suponemos que el precio del activo se encuentra en un valor inicial  $S_0$ , y existen dos estados de la naturaleza para el próximo período: H y L. En el estado H el precio del activo sube a  $uS_0$ , mientras que en el estado L el precio del activo baja a  $dS_0$ . Es decir, la relación:  $0 < d < 1 < u$ , es al menos esperada.

Cuando el intervalo de tiempo entre el período inicial y el próximo período es importante entonces la relación cambia a  $0 < d < 1 + r < u$ , donde  $r$  es la tasa de interés relevante, esto debido a que un monto del próximo período debe ser descontado. Para efectos estadísticos esta relación no es tan importante, pero para

valoración financiera lo es. La relación entre la tasa de interés y los coeficientes de incremento y reducción es extremadamente para el primer teorema fundamental en finanzas, el cual implica la no existencia de arbitraje.

Debido a que existen dos estados posibles para  $S_1$ , este puede ser entendido como una variable aleatoria. En particular notamos que es una variable aleatoria binomial, cuyo dos primeros momentos teóricos pueden ser calculados como  $E(S_1) = [up + d(1-p)]S_0$  y  $E(S_1^2) = [u^2p + d^2(1-p)]S_0$ . Por otro lado, si contamos con información de los retornos del activo, tal que su media sea  $\mu$  y su volatilidad  $\sigma$ , entonces podemos calibrar los parámetros  $p$ ,  $u$  y  $d$ .

Para una ventana de  $T$  días y un número de pasos del árbol de  $N$  tenemos que:  $u = \exp(\sigma\sqrt{T/N})$ ,  $d = 1/u$  y

$$p = \frac{1}{2} + \frac{1}{2\sigma} \left( \mu - \frac{\sigma^2}{2} \right) \sqrt{\frac{T}{N}} + O\left(\frac{T}{N}\right)$$

Cuando los retornos son cercanos a cero entonces  $p \approx 0,5$ . Esta calibración no es única.

### Volatilidad Intradía

Suponga que  $T = 1$  y  $N = 2$ . Además suponga  $S_0 = 1$  y  $c = \log S_1$ , mientras que  $h$  y  $l$  corresponden a los máximos y mínimos observados durante el día.

Seq	$c$	$h$	$l$
uu	$2 \log(u)$	$2 \log(u)$	0
ud	0	$\log(u)$	0
du	0	0	$-\log(u)$
dd	$-2 \log(u)$	0	$-2 \log(u)$

Notar que  $\log(u) = \sigma\sqrt{0,5}$ . Alfaro y Silva (2008) estudian la volatilidad del IPSA comparando las medidas  $\sigma_{CC}^2 = c^2$  y  $\sigma_{HL}^2 = (h-l)^2$  cuando el retorno es cercano a cero  $p = 0,5$ .

En dicho caso  $E(\sigma_{CC}^2) = \sigma^2$  y  $E(\sigma_{HL}^2) = 1,25\sigma^2$ . Los autores argumentan que el sesgo de HL es condicional al  $N$ .

Para  $N = 2$  los Errores Cuadrático Medios (ECM) son  $ECM(\sigma_{CC}^2) = \sigma^4$  y  $ECM(\sigma_{HL}^2/1,25) = 0,36\sigma^4$ .

Los autores computan el factor de sesgo teórico para  $N \leq 22$  y empírico para el caso del IPSA (2,37). Con ello proponen el índice Volatilidad Intradía del Mercado Accionario (VIMA)

$$VIMA = 1000 \left( \frac{\max}{\min} - 1 \right)$$

### Valoración de Opciones

Originalmente el modelo binomial fue utilizado por Cox, Ross y Rubinstein (1979) para la valoración de opciones.

En el mundo riesgo neutral los agentes no se preocupan por el riesgo por lo que valoración de opciones se considera justa en el sentido competitivo.

En términos más simples bajo el riesgo neutral, las probabilidades asociadas al activo están dictadas por la tasa de interés de mercado.

Así reemplazamos  $\mu$  por  $r$  donde sea posible, haciendo

$$q = \frac{1}{2} + \frac{1}{2\sigma} \left( r - \frac{\sigma^2}{2} \right) \sqrt{\frac{T}{N}} + O\left(\frac{T}{N}\right)$$

El método de valoración es regresivo, iniciándose en la posición final y retrocediendo un paso a la vez.

Cada nodo contiene la valoración de los nodos sucesivos en valor presente, utilizando  $q$  y  $(1 - q)$ .

Para el caso de opciones europeas el resultado es trivial, por ejemplo la call tiene un valor de

$$c = \sum_{i=0}^n \binom{n}{i} q^i (1-q)^{n-i} \max[S_0 u^i d^{n-i} - X, 0]$$

El modelo binomial converge al modelo Black-Scholes cuando  $N \rightarrow \infty$ . Sin embargo su convergencia es cíclica.

## 6. Aplicación: Factor de Descuento Estocástico

En esta sección se analiza el modelo de valoración de activos que utiliza la ecuación de Euler de consumo intertemporal con el factor de descuento derivado de la Tasa Marginal de Sustitución Intertemporal.

### La distribución lognormal

Consideremos que nuestro interés es  $M$ , pero sabemos que  $m = \log(M)$  se distribuye normal con media  $\mu_m$  y varianza  $\sigma_m^2$ . Entonces la FGM para  $m$  corresponde a  $E[\exp(tm)] = \exp(\mu_m t + 0,5\sigma_m^2 t^2)$ . Así  $E[\exp(m)] =$

$\exp(\mu_m + 0,5\sigma_m^2)$  y  $E[\exp(2m)] = \exp(2\mu_m + 2\sigma_m^2)$ . Notamos que  $E[\exp(m)] = E(M)$  y  $E[\exp(2m)] = E(M^2)$ . Con ello  $V(M) = \exp(2\mu_m + 2\sigma_m^2) - \exp(2\mu_m + \sigma_m^2) = \exp(2\mu_m + \sigma_m^2)[\exp(\sigma_m^2) - 1]$ . Por ello se dice que  $M$  se distribuye lognormal con media  $E(M)$  y varianza  $V(M)$ . Considerando  $\sigma(M) = \sqrt{V(M)}$ , entonces  $\sigma(M)/E(M) = \sqrt{\exp(\sigma_m^2) - 1}$ . Expandiendo  $\exp(\sigma^2) = \exp(\sigma_0^2) + \exp(\sigma_0^2)(\sigma^2 - \sigma_0^2)$  y fijando  $\sigma_0^2 = 0$  se tiene que  $\sigma(M)/E(M) \approx \sigma_m$ .

### La ecuación de Euler

Consideremos un activo con precio  $P_t$ , cuyo valor es  $X_{t+1}$  el próximo período. Entonces se establece que su valor viene determinado por  $P_t = E_t(M_{t+1}X_{t+1})$ , donde  $M_{t+1}$  es el factor de descuento estocástico.

En particular el caso de un bono de descuento también satisface la ecuación. Si  $B_{1t}$  es el precio entonces  $B_{1t} = E_t(M_{t+1}1) = E_t(M_{t+1})$ , lo que implica que  $1/(1+Y_{1t}) = E_t(M_{t+1})$ , donde  $Y_{1t}$  es la tasa de descuento.

Escribiendo Euler en retornos y utilizando la definición de covarianza tenemos que  $1 = E_t[M_{t+1}(1 + R_{i,t+1})] = Cov_t(M_{t+1}R_{i,t+1}) + E_t(M_{t+1})E_t(1 + R_{i,t+1})$ . Utilizando el resultado de bonos de descuento se tiene que  $(1 + Y_{1t}) = (1 + Y_{1t})Cov_t(M_{t+1}R_{i,t+1}) + E_t(1 + R_{i,t+1})$ , lo que con algo de álgebra se traduce en

$$\frac{E_t(R_{i,t+1}) - Y_{1t}}{\sigma_t(R_{i,t+1})} \equiv S(R_{i,t+1}) = -(1 + Y_{1t})\sigma_t(M_{t+1})\rho = -\rho \frac{\sigma_t(M_{t+1})}{E_t(M_{t+1})}$$

Esto implica que el índice de Sharpe (mencionado con  $S$  en la ecuación previa) se relaciona negativamente con la correlación entre el factor de descuento estocástico y el retorno del activo ( $\rho$ ). Finalmente  $|S(R_{i,t+1})| = |\rho|\sigma_t(M_{t+1})/E_t(M_{t+1}) \leq \sigma_t(M_{t+1})/E_t(M_{t+1})$ , pues  $M_{t+1} \geq 0$  y  $|\rho| \leq 1$ .

Para el caso en que  $u'(C) = C^{-\gamma}$  tenemos que  $M_{t+1} = \beta(C_{t+1}/C_t)^{-\gamma}$ . Considerando  $c = \log(C)$  el resultado es  $m_{t+1} = \log(\beta) - \gamma(c_{t+1} - c_t)$ . Tomando la desviación estándar y los resultados previos:  $|S(R_{i,t+1})| \leq \sigma_t(M_{t+1})/E_t(M_{t+1}) \approx \sigma_m = \gamma\sigma(\Delta c_{t+1})$ .

### Exceso de retorno y tasa libre de riesgo

Notemos que  $(1+Y_{1t}) = 1/E_t(M_{t+1})$  lo que se traduce en  $y_{1t} = \log(1+Y_{1t}) = -\log E_t(M_{t+1}) = -\mu_m - 0,5\sigma_m^2$ . Dado  $m_{t+1} = \log M_{t+1} = \log(\beta) - \gamma\Delta \log C_{t+1}$ , y definiendo  $g_{t+1} = \log C_{t+1} - \log C_t$  tenemos que

$$y_{1t} = -\log(\beta) + \gamma E_t(g_{t+1}) - 0,5\gamma^2 Var_t(g_{t+1})$$

Esto entrega una fórmula para la tasa libre de riesgo en función del crecimiento del consumo.

En el caso general se tiene que  $1 = E_t[M_{t+1}(1 + R_{i,t+1})]$ . De este modo  $0 = E_t(m_{t+1} + r_{i,t+1}) + 0,5Var_t(m_{t+1} + r_{i,t+1}) = \mu_m + E_t(r_{i,t+1}) + 0,5(\sigma_i^2 + 2\sigma_{mr} + \sigma_m^2)$ , donde  $\sigma_i$  corresponde a la volatilidad del activo y  $\sigma_{mr}$  su covarianza con el factor de descuento estocástico. Notemos que  $\sigma_{mr} = E[(m - \mu_m)(r - E(r))] = \gamma Cov(g, r)$ , es decir la covarianza entre el retorno y el crecimiento del consumo ponderado por  $\gamma$ . Ordenando las ecuaciones tenemos que

$$E_t(r_{i,t+1} - y_{1t}) + 0,5Var_t(r_{i,t+1}) = \gamma Cov_t(g_{t+1}, r_{i,t+1})$$

Es decir, el exceso de retorno se encuentra positivamente relacionado a la covarianza entre crecimiento del consumo y el retornos del activo.

## Análisis Descriptivo

### 1. Los Datos

Existen dos grandes categorías de variables: cualitativas y cuantitativas. Las primeras de ellas se refieren a una característica de un objeto como su color, la segunda son eminentemente números (conjunto ordenado). Además una variable cuantitativa puede ser de naturaleza discreta o continua, así discreto es el número de hijos de una familia y continuo es la cantidad de litros de agua que puede contener un recipiente.

Por último otra forma de clasificar la información es a través de la dimensión temporal, es decir en Series de Tiempo, Corte Transversal o Panel. La primera de ellas se refiere a una variable que se observa a través del tiempo como la inflación, el tiempo y el PIB de una nación. Corte Transversal corresponde a una fotografía que se toma en un momento del tiempo, así por ejemplo la distribución de ingresos de una nación, el salario que perciben las mujeres que trabajan en Santiago en el año 2002 y los promedios de notas de los alumnos de una clase. El caso de panel, es un poco más costoso ya que se refiere a una observación o unidad que es seguida en el tiempo, por ejemplo los salarios que reciben un grupo de egresados de la universidad.

En el desarrollo de nuestro trabajo trataremos en general con variables que son de carácter numérico, por ello es importante revisar algunas funciones descriptivas.

#### 1. Clasificación de Estacionariedad

Otra clasificación de las variables algo más avanzada guarda relación con su estacionariedad. En términos simples se dirá que una variable tratada como serie de tiempo es estacionaria si presenta una media incondicional, vale decir que su valor esperado no depende del tiempo. Vulgarmente, una variable estacionaria no crece.

Además una variable estacionaria presenta autocovarianzas que no dependen del tiempo. En la Sección ?? se verán estos temas en profundidad. Por ahora diremos que este último requisito implica que la variable presenta una varianza estable a través del tiempo.

## 2. Primeras Definiciones

**Definición 3.1.** *Dos conceptos principales que se abordarán continuamente son la población y la muestra. La primera corresponde a una colección de toda la información posible que caracteriza a un fenómeno. La muestra es, por otro lado, un subconjunto de la población, por tanto bajo el concepto de representatividad podemos suponer que las características de la población se ven de algún modo reflejadas en una muestra. Posteriormente veremos como se pueden construir dichas muestras.*

**Definición 3.2.** *Media Muestral es el Promedio Aritmético simple. La función correspondiente en Excel es `promedio()`.*

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.1)$$

**Definición 3.3.** *Mediana (Med) es la observación ubicada en el centro de la colección de observaciones ordenadas de menor a mayor. En el caso de número pares es el promedio de los valores centrales. La función correspondiente en Excel es `mediana()`.*

**Definición 3.4.** *Moda (Mod) es el valor de la observación que más se repite.*

**Definición 3.5.** *Varianza Muestral es una medida de dispersión de los datos, con respecto a su valor medio. La raíz de ella ( $s$ ) se conoce como Desviación Estándar. La función correspondiente en Excel son `var()` para la varianza y `desvest()` para  $s$ .*

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.2)$$

**Definición 3.6.** *Desviación Media (DM), corresponde a otra medida de dispersión. La función correspondiente en Excel es `desvprom()`.*

$$DM = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}| \quad (3.3)$$

**Definición 3.7.** *Momentos Muestrales: se definen como (3.4) para los centrados en cero y como (3.5) para los centrados en la media<sup>1</sup>.*

$$\hat{\theta}_r = \frac{1}{n} \sum_{i=1}^n x_i^r \quad (3.4)$$

$$\hat{\mu}_r = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^r \quad (3.5)$$

---

<sup>1</sup>Ver código en C.1.2

### 3. Conceptos Técnicos

**Definición 3.8.** *Variación Porcentual:* se base en índices y lo lleva a una tasa, como se observa en (3.6). Tiene la utilidad de hacer comparable 2 variables, como son por ejemplo el ingreso de 2 sectores económicos. Es muy “raro” hablar de cambios porcentuales en el caso de tasas.

$$\Delta \%Y = \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \frac{Y_t}{Y_{t-1}} - 1 \quad (3.6)$$

Por otro lado si consideramos  $y_t = \log Y_t$  y la variación en un pequeño intervalo de tiempo, entonces podemos escribir la variación porcentual como una derivada logarítmica como se muestra en (3.7).

$$\dot{y} = \frac{dy}{dt} = \frac{d \log Y}{dt} = \left( \frac{dY}{dt} \right) \frac{1}{Y} \quad (3.7)$$

Otra forma de presentar el problema es que  $\dot{y}$  corresponde a la TCC de  $\Delta \%Y$ . Consideremos que  $Y_t = Y_{t-1}(1 + G)$ , donde  $G$  es tasa de crecimiento. Tomando  $g$  como TCC de  $G$  podemos probar que:

$$g = \log \left( \frac{Y_t}{Y_{t-1}} \right) = \log \left[ \frac{Y_{t-1}(1 + G)}{Y_{t-1}} \right] = \log(1 + G)$$

**Ejemplo 3.1.** *Considere que Producto Interno Bruto (PIB) de un país es 3600 unidades monetarias (um) en el año  $t$  y 4500 um en el año  $t+1$ . De este modo la variación porcentual en un año es:  $\Delta \%PIB = \Delta \%Y = (4500 - 3600)/3600 = 0,25$ , es decir 25 %. Bajo el uso logarítmico tenemos:  $\dot{y} = 8,4118 - 8,1887 \approx 0,22$ , que corresponde a  $\log(1 + 0,25) \approx 0,22$*

**Definición 3.9.** *Promedio Móvil Simple:* nos permite suavizar una serie. Consiste simplemente en tomar para cada período el promedio de los  $J - 1$  períodos para atrás. Vale decir:

$$S_t = \frac{1}{J} \sum_{j=0}^{J-1} Y_{t-j} \quad (3.8)$$

**Definición 3.10.** *Promedio Móvil Centrado:* análogo al anterior. Pero, se utilizan valores futuros, por ello se requiere previamente de una predicciones de la serie. Dichas predicciones se realizan a través de la metodología ARIMA, desarrollada por Box y Jenkins. Adicionalmente se pueden utilizar ponderadores decrecientes a medida que los datos se alejan temporalmente. Esta metodología se encuentra desarrollada en el programa X12, que utiliza el Departamento de Estadística de EE.UU.



**Ejemplo 3.2.** Ejemplos para trimestres móviles serían  $S_t = (Y_t + Y_{t-1} + Y_{t-2})/3$ ,  $S_t = (Y_t + Y_{t-1} + Y_{t+1})/3$  y  $S_t = (3 \cdot Y_t + 2Y_{t-1} + Y_{t-2})/6$ .

**Definición 3.11.** *Interpolación Lineal:* a menudo no se cuenta con información para cada uno de los períodos por tanto se completa con un crecimiento lineal. Considerando los datos  $X_{t_1}$  y  $X_{t_2}$ , entonces cada dato  $j \in (t_1, t_2)$ , puede ser escrito como:

$$X_j = \left( \frac{X_{t_2} - X_{t_1}}{t_2 - t_1} \right) + X_{j-1} \quad (3.9)$$

**Ejemplo 3.3.** Suponga que desea interpolar 2 datos entre  $a1=10$  y  $a4=40$ . Para ello puede generarse  $a2=a1+(a4-a1)/3$  y copiarse en  $a3$ , ya que se han fijado las celdas de las “esquinas”. Los resultados son 20 y 30.

**Definición 3.12.** *Interpolación Geométrica:* del mismo modo que la anterior interpolación se requiere completar los datos que se ubican entre 2 valores. En este caso se utiliza un crecimiento exponencial. Considerando los datos  $X_{t_1}$  y  $X_{t_2}$ , entonces cada dato  $j \in (t_1, t_2)$ , puede ser escrito como:

$$X_j = X_{j-1} \left( \frac{X_{t_2}}{X_{t_1}} \right)^{\frac{1}{t_2 - t_1}} \quad (3.10)$$

Tomamos los logaritmos en minúsculas como  $x_{t_i} = \log X_{t_i}$  podemos escribir:

$$X_j = \exp \left( x_{j-1} + \frac{x_{t_2} - x_{t_1}}{t_2 - t_1} \right)$$

#### 4. Volatilidad del Retorno de un Activo

Sea  $S_t$  el precio cierre de un activo en el momento  $t$ , luego  $R_{t+1}$  corresponderá al retorno que se asumirá tiene una distribución incondicional caracterizada por una normal con media  $\mu_d$  y varianza  $\sigma_d^2$ . Esto puede ser descrito como (3.11) donde  $\phi$  es una normal estándar.

$$R_{t+1} = \frac{S_{t+1} - S_t}{S_t} = \mu_d + \sigma_d \phi \quad (3.11)$$

**Definición 3.13.** *Efecto Escala:* en el caso de utilizar datos diarios los momentos muestrales que estiman a  $\mu_d$  y  $\sigma_d^2$  son válidos para un base diaria. Para simplificar el análisis se adoptará la formalidad anual tomando

$h$  como la fracción de un día sobre un año (vale decir  $1/h \approx 20 \times 12 = 240$ ). De este modo los parámetros anuales corresponderán a  $\mu = \mu_d/h$  y  $\sigma = \sigma_d/\sqrt{h}$ .

**Definición 3.14.** *Rentabilidad:* consideraremos la rentabilidad promedio ajustada por el efecto escala como sigue:

$$\mu = \frac{1}{nh} \sum_{i=1}^n R_i = \frac{1}{h} \bar{R}$$

**Definición 3.15.** *Volatilidad 1:* una medida simple de volatilidad del retorno puede ser obtenida de la desviación estándar (3.12).

$$\sigma_1^2 = \frac{1}{(n-1)h} \sum_{i=1}^n (R_{t+i} - \bar{R})^2 \quad (3.12)$$

**Definición 3.16.** *Volatilidad 2:* análogamente se puede trabajar con las desviaciones en módulo (3.13).

$$\sigma_2 = \frac{1}{nh} \sum_{i=1}^n |R_{t+i} - \bar{R}| \quad (3.13)$$

Estas medidas pueden considerarse como formas simples y eficaces de volatilidad. Sin embargo, como todos los retornos son igualmente ponderados en el tiempo, la medida de se hace menos confiable a medida que  $n$  crece.

**Definición 3.17.** *Volatilidad 3 (Medida Ponderada):* la ponderación exponencial móvil se basa en un multiplicador  $\lambda$  que pertenece al intervalo  $(0, 1)$ . De este modo, la medida (3.14) permite ponderar eficientemente las observaciones.

$$\sigma_3^2 = \sum_{i=1}^n \frac{\lambda^{i-1}}{\sum_{j=1}^n \lambda^{j-1}} (R_{t+i} - \bar{R})^2 = \frac{(1-\lambda)}{(1-\lambda^n)} \sum_{i=1}^n \lambda^{i-1} (R_{t+i} - \bar{R})^2 \quad (3.14)$$

Una forma para encontrar el  $\lambda$  apropiado es a través de la minimización de  $W$  de la expresión (3.15), donde  $X_{t+1} = \lambda X_t + (1-\lambda)R_t$ , con la condición inicial  $X_1 = R_1$ .

$$W = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_{t+i} - X_{t+i})^2} \quad (3.15)$$

## 1. Definiciones

**Definición 4.1.** *“Test de Hipótesis”*: debido a que conocemos cuál es la distribución que caracteriza a una estadística podemos realizar una comparación entre el valor de tabla de una distribución y el valor muestral. Sin embargo aún no hemos definido una regla para determinar cuando la comparación es aceptable. La idea entonces es definir primero una hipótesis nula ( $H_0$ ) que implica suponer el comportamiento de un parámetro de la distribución, esta  $H_0$  se contrasta con una  $H_1$  que es un espacio del complemento del parámetro.

Tomemos el ejemplo sobre un parámetro  $\theta$ . Los tipos posibles de pruebas son  $H_0 : \theta = \theta_0$  contra: (1)  $H_1 : \theta = \theta_1$ : en este caso corresponde realizar un test de Razón de Verosimilitud y (2)  $H_1 : \theta > \theta_0$  ( $H_1 : \theta < \theta_0$ ): el área relevante se ubica a la derecha (izquierda).

Pese a lo anterior, se requiere un criterio adicional, que define el tamaño del área relevante.

**Definición 4.2.** *“Error tipo I”*( $\alpha$ ): es la probabilidad de rechazar  $H_0$ , dado que esta es verdadera.

**Definición 4.3.** *“Error tipo II”*( $\beta$ ): es la probabilidad de no rechazar  $H_0$ , dado que es falsa.

**Definición 4.4.** *“Intervalo de Confianza”*(IC): se define como el área que se construye para una determinada estadística en la cual se tiene la certeza de que no se está cometiendo un  $\alpha$  determinado. Así un IC se asocia a una región donde se ubica el parámetro que estima la estadística.

**Definición 4.5.** *“Función de Poder”*(FP): se define como  $1 - \beta$ , por tanto corresponde a la probabilidad de rechazar  $H_0$  cuando esta es falsa. Permite ver cuan eficiente es una estadística para identificar valores erróneos obtenidos de una muestra.

## 2. Propiedades de las Estadísticas

Cuando se tiene una muestra más una estadística podemos encontrar una estimación puntual. Sin embargo, lo que nos interesa es obtener propiedades deseables sobre un conjunto de estadísticas.

**Definición 4.6.** “Error Cuadrático Medio”(ECM): sea  $T$  cualquier estadística sobre un parámetro desconocido  $\theta$ .  $ECM(T) = E(T - \theta)^2$ .<sup>1</sup>

**Definición 4.7.** “Insesgamiento”: se dice que una estadística  $T$  es insesgada para  $\theta$  si  $E(T) = \theta$ . Esta corresponde a un definición “finita” de convergencia en probabilidad.<sup>2</sup>

Siempre es posible ordenar a las estadísticas de acuerdo a su menor ECM. Sin embargo, en términos prácticos es preferible una estadística insesgada. Esto último tiene gran relevancia cuando el propósito es realizar predicciones utilizando la estadística.

**Teorema 4.1.** “Cota de Crámer-Rao”(CR): define una varianza mínima para una estadística insesgada, esta es: . Así cuando la estadística cumple con esta cota se dirá que es “eficiente”.

$$V(T) = \left\{ nE \left[ \left( \frac{\partial \log f(X; \theta)}{\partial \theta} \right)^2 \right] \right\}^{-1} \quad (4.1)$$

**Teorema 4.2.** “Factorización de Neyman”: define la separación de la función de verosimilitud (distribución conjunta) con el fin de definir si una estadística es “suficiente”.

**Proposición 4.1.** Trabajando con una población independiente con media  $\mu$  y varianza  $\sigma^2$ . Presente las propiedades de la estadística  $\bar{x}$ .

Esta sección presenta cómo podemos derivar resultados a partir de una muestra. En la primera parte se entregan definiciones de muestras finitas<sup>3</sup>. El resto de las secciones está diseñado para formalizar la Teoría de Grandes Muestras. La idea es apoyar en teoremas la distribución asintótica del promedio muestral ( $\bar{x}$ ) y de la estadística de la varianza ( $s^2$ ).

<sup>1</sup>Notemos que convergencia en media cuadrática implica que el ECM tiende a cero conforme aumenta el tamaño de la secuencia.

<sup>2</sup>Notemos, que para toda estadística insesgada el ECM corresponde a la varianza.

<sup>3</sup>Básicamente trabajaremos con muestras independientes, por tanto nuestro muestreo será con reemplazo.

### 3. Muestreo Finito

**Definición 4.8.** *Estadística es una función de va que busca una distribución sobre el parámetro en estudio. O de otro modo, corresponde a la función que puede ser computada con una muestra. La estadística de mayor interés corresponderá al promedio aritmético simple.*

**Teorema 4.3.** *Sea el set de va  $\{X_i\}_{i=1}^n$  independientes con FGM  $\{m_{X_i}(t)\}_{i=1}^n$  y un set de coeficientes  $\{a_i\}_{i=1}^n$ . Entonces para una nueva va  $Y = \sum_{i=1}^n a_i X_i$ , tendremos que su FGM será:  $m_Y(t) = \prod_i m_{X_i}(ta_i)$ .*

**Proposición 4.2.** *Realice las siguientes demostraciones:*

1. *El teorema anterior.*
2. *Para un set de va  $\{X_i\}_{i=1}^n$  que distribuyen normales e independientemente, vale decir  $N(\mu_i, \sigma_i^2)$ . Demuestre  $Y = \sum_{i=1}^n a_i X_i$  se distribuye  $N(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$ .*

### 4. Modos de Convergencia

Suponiendo que se cuenta con una secuencia de variables aleatorias como:  $\{X_i\}_{i=1}^n$ , definiremos 4 modos de convergencia.

**Definición 4.9.** *una secuencia de variables aleatorias se dice que converge en probabilidad a una va  $X$  si:  $\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$ , para todo  $\epsilon > 0$ . Su notación será  $\text{plim} X_n = X$  o  $X_n \xrightarrow{P} X$ .*

**Definición 4.10.** *de la definición anterior se dirá que una estadística es consistente si converge en probabilidad al parámetro que estima.*

**Definición 4.11.** *una secuencia de variables aleatorias se dice que converge en media cuadrática a  $X$  si:  $\lim_{n \rightarrow \infty} E(X_n - X)^2 = 0$ . Su notación será  $X_n \xrightarrow{M} X$ .*

**Definición 4.12.** *una secuencia de variables aleatorias se dice que converge en distribución a  $X$  si la función de distribución de  $X_n$  ( $F_n$ ) converge a  $F$  de  $X$  para todo punto continuo de  $F$ . Su notación será  $X_n \xrightarrow{d} X$ .*

Por ejemplo, la distribución  $t_v$  converge a una  $N(0,1)$  conforme  $v \rightarrow \infty$ . Esta prueba es sencilla pues  $T = Z/\sqrt{X/v}$ , donde  $Z \sim N(0,1)$  y  $X \sim \chi_v^2$ . Por tanto la distribución de interés es  $X$ . Se sabe que  $E(X) = v$  y que  $V(X) = 2v$ . De este modo, si consideramos  $Y = X/v$ , se tiene que  $E(Y) = 1$  y  $V(Y) = 2/v$ . Así cuando  $v \rightarrow \infty \Rightarrow Y \xrightarrow{d} 1 \Rightarrow t \xrightarrow{d} z$ .

**Definición 4.13.** una secuencia de variables aleatorias se dice que converge casi seguro a  $X$  si:  $P[\omega | \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)] = 1$ . Su notación será  $X_n \xrightarrow{as} X$ .

**Teorema 4.4.** Chebyshev.  $E(X_n^2) \rightarrow 0 \implies X_n \xrightarrow{P} X$ . Lo anterior permite expresar que:  $X_n \xrightarrow{M} X \implies X_n \xrightarrow{P} X$ .

**Teorema 4.5.** la estadística  $\bar{x}$  obtenida de una población independiente con media finita  $\mu$  y varianza  $\sigma^2$  es un estimador consistente. Esto es fácil de ver, ya que  $E(\bar{x}) = \mu$  y  $V(\bar{x}) = n^{-1}\sigma^2$ , entonces  $\bar{x}$  converge en media cuadrática a  $\mu$  y por tanto  $\text{plim}\bar{x} = \mu$ .

**Teorema 4.6.**  $X_n \xrightarrow{P} X \implies X_n \xrightarrow{d} X$ . El teorema se cumple en ambas direcciones si  $X$  es una constante, vale decir si  $X_n \xrightarrow{d} \alpha \implies X_n \xrightarrow{P} \alpha$ .

**Teorema 4.7.**  $X_n \xrightarrow{as} X \implies X_n \xrightarrow{P} X$ .

## 5. Leyes de los Grandes Números

Dado una secuencia de variables aleatorias  $\{X_i\}_{i=1}^n$ , y tomando  $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ . Un ley de los grandes números (LGN) especifica las condiciones bajo las cuales  $\bar{X}_n - E(\bar{X}_n)$  converge a 0 de forma casi segura o en probabilidad.

**Teorema 4.8.** “Kolmogorov 1”: Sea  $\{X_i\}_{i=1}^n$  un set independiente con varianza  $V(X_i) = \sigma_i^2$ . Si  $\sum_{i=1}^{\infty} \sigma_i^2/i^2 < \infty$ , entonces  $\bar{X}_n - E(\bar{X}_n) \xrightarrow{as} 0$ .

**Teorema 4.9.** “Kolmogorov 2”: Sea  $\{X_i\}_{i=1}^n$  un set iid (independiente e idénticamente distribuido). Luego una condición necesaria y suficiente para que  $\bar{X}_n \xrightarrow{as} \mu$  es que  $E(X_i)$  exista y sea igual a  $\mu$ .

Los teoremas anteriores son complementarios, ya que las condiciones que requieren cada uno de ellos son distintas y no se existe una que sea superior a la otra. Otro ejemplo lo proporciona Markov.

**Teorema 4.10.** “Markov” Sea  $\{X_i\}_{i=1}^n$  un set independiente y suponiendo que  $E(|X_i|^{1+\delta}) < M < \infty$  para algún  $\delta > 0$  y algún  $M < \infty$ . Luego  $\bar{X}_n - E(\bar{X}_n) \xrightarrow{P} 0$ .

## 6. Teoremas Centrales del Límite

Lo que nos interesa saber es cuál es la distribución que sigue  $\bar{X}_n$  cuando  $n \rightarrow \infty$ . Ya que por alguna LGN podemos llegar a  $\bar{X}_n - E(\bar{X}_n) \xrightarrow{P} 0$ . Y tomando el teorema 4.6 llegaremos a  $\bar{X}_n - E(\bar{X}_n) \xrightarrow{d} 0$ , mas esta

es una distribución degenerada. Lo anterior ocurre debido a que  $V(\bar{X}_n) \rightarrow 0$  cuando  $n \rightarrow \infty$ . Por tanto, definiendo una nueva secuencia  $Z_n = [V(\bar{X}_n)]^{-1/2}[\bar{X}_n - E(\bar{X}_n)]$ , que posee la gracia que  $V(\bar{Z}_n) = 1$  para todo  $n$ . Un teorema central del límite (TCL) establece las condiciones para que  $Z_n$  converja (en distribución a) una distribución normal estándar.

**Teorema 4.11.** “Lindeberg-Lévy”: Sea  $\{X_i\}_{i=1}^n$  un set iid con  $E(X_i) = \mu$  y  $V(X_i) = \sigma^2$ . Luego  $Z_n \rightarrow N(0, 1)$ .<sup>4</sup>

**Teorema 4.12.** “Liapunov”: Sea  $\{X_i\}_{i=1}^n$  un set independiente con  $E(X_i) = \mu_i$ ,  $V(X_i) = \sigma_i^2$  y  $E[|X_i - \mu_i|^3] = m_{3i}$ . Si:

$$\lim_{n \rightarrow \infty} \left[ \sum_{i=1}^n \sigma_i^2 \right]^{-1/2} \left[ \sum_{i=1}^n m_{3i} \right]^{1/3} = 0$$

Entonces  $Z_n \rightarrow N(0, 1)$ .

**Teorema 4.13.** “Lindeberg-Feller”: Sea  $\{X_i\}_{i=1}^n$  un set independiente con funciones de distribución  $\{F_i\}_{i=1}^n$  y  $E(X_i) = \mu_i$  y  $V(X_i) = \sigma_i^2$ . Definiendo  $C_n = (\sum_{i=1}^n \sigma_i^2)^{1/2}$ . Bajo una condición especial<sup>5</sup> se cumple que  $Z_n \rightarrow N(0, 1)$ .

## 7. Distribución de $s^2$

En la sección anterior investigamos los teoremas que nos permiten establecer una distribución para la estadística  $\bar{X}$ . Considerando esto, hecho que si  $P$  es una normal estándar entonces  $P^2$  corresponde a una chi-cuadrado con 1 grado de libertad y el Teorema 2.4 sobre la suma de distribuciones chi-cuadrado entonces podremos establecer un teorema para  $s^2$ .

**Teorema 4.14.** Para una muestra  $\{X_i\}_{i=1}^n$  proveniente de una población caracterizada por una distribución  $N(\mu, \sigma^2)$  la estadística  $(n-1)\sigma^{-2}s^2$  sigue una distribución chi-cuadrado con  $n-1$  grados de libertad.

<sup>4</sup>Este es el teorema usualmente presentado en los libros de texto.

<sup>5</sup>Para conocer la condición vea Amemiya(1985): “Advance Econometrics”, Harvard University Press.

## Métodos de Estimación: MV y MM

### 1. Máxima Verosimilitud (MV)

Considere una ma independiente proveniente una distribución con fdp  $f(x; \theta)$ , entonces la distribución conjunta  $f(x_1, \dots, x_n; \theta) = f(x_1; \theta)f(x_2; \theta) \cdots f(x_n; \theta)$ . Se dirá que  $\hat{\theta}_{MV}$  es el estimador de máxima verosimilitud para  $\theta$  si corresponde al argumento que maximiza a la probabilidad conjunta. Dicho de otro modo, el estimador de MV es aquel que maximiza la probabilidad de ocurrencia de la muestra, bajo una distribución determinada.

**Definición 5.1.** *La verosimilitud ( $L$ ) o probabilidad conjunta corresponde a la función objetivo a maximizar. Para simplificar se aplica sobre  $L$  una función no decreciente como es el logaritmo, así  $l = \log L$  que implica (5.1) y se dirá  $\hat{\theta}_{MV} = \operatorname{argmax}(l)$ .*

$$l = \sum_{i=1}^n \log f(x_i; \theta) \quad (5.1)$$

#### 1. Propiedades de la Optimización en MV

La optimización de (5.1) se caracteriza porque su gradiente tiene valor esperado cero y porque la varianza del gradiente es igual al negativo del valor esperado del hessiano<sup>1</sup>. Considerando la observación  $i$  se tiene que  $\int f(x_i; \theta) dx_i = 1$ . Derivando esta igualdad con respecto a  $\theta$ :

$$0 = \int \frac{\partial f(x_i; \theta)}{\partial \theta} dx_i = \int \left[ \frac{\partial f(x_i; \theta)}{\partial \theta} \frac{1}{f_i} \right] f_i dx_i = \int \frac{\partial \log f(x_i; \theta)}{\partial \theta} f_i dx_i = E \left[ \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right]$$

Considerando la definición de gradiente:

<sup>1</sup>Notar que como la varianza es una matriz definida positiva entonces el hessiano es definido negativo lo que constituye un requisito para encontrar un máximo.



$$g = \frac{\partial l}{\partial \theta} = \sum_{i=1}^n \frac{\partial \log f(x_i; \theta)}{\partial \theta} \Rightarrow E[g] = E \left[ \sum_{i=1}^n \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right] = \sum_{i=1}^n E \left[ \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right] = 0$$

Tomando la segunda derivada con respecto a  $\theta$  en la ecuación anterior se tiene:

$$\frac{\partial}{\partial \theta} \left[ \int \frac{\partial \log f(x_i; \theta)}{\partial \theta} f_i dx_i \right] = \int \frac{\partial^2 \log f(x_i; \theta)}{\partial \theta \partial \theta^T} f_i dx_i + \int \left( \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right) \left( \frac{\partial f(x_i; \theta)}{\partial \theta} \right)^T dx_i = 0$$

Notar que:

$$\int \left( \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right) \left( \frac{\partial f(x_i; \theta)}{\partial \theta} \right)^T dx_i = \int \left( \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right) \left( \frac{\partial \log f(x_i; \theta)}{\partial \theta} \right)^T f_i dx_i$$

Con algo más de álgebra se llega a:

$$E \left[ \left( \frac{\partial \log f_i}{\partial \theta} \right) \left( \frac{\partial \log f_i}{\partial \theta} \right)^T \right] = -E \left[ \frac{\partial^2 \log f_i}{\partial \theta \partial \theta^T} \right] \Rightarrow E(gg^T) = V(g) = -E(H)$$

## 2. Propiedades del Estimador MV

Es consistente, asintótico normal, asintótico eficiente e invariante.

## 3. Ejemplo: la distribución Poisson

Consideremos que la va  $Y$  se distribuye Poisson con parámetro  $\lambda$ , entonces su fdp corresponderá a:  $p(y_i; \lambda) = \exp(-\lambda) \lambda^{y_i} / y_i! \Rightarrow \log p(y_i) = -\lambda + y_i \log(\lambda) - \log(y_i!)$ . Agregando para las  $n$  observaciones se tiene que  $l = -n\lambda + \log(\lambda) \sum_i y_i - \sum_i \log(y_i!)$ . Al maximizar  $l$  se tiene que:

$$g = \frac{dl}{d\lambda} = -n + \frac{1}{\lambda} \sum_{i=1}^n y_i$$

que al igualarse a cero se obtiene  $\hat{\lambda}_{MV} = \bar{y}$ . Notar que el Hessiano asegura que la estimación corresponde a un máximo, pues todos los  $y_i > 0$ .

$$H = \frac{d^2 l}{d\lambda^2} = -\frac{1}{\lambda^2} \sum_{i=1}^n y_i < 0$$

$\hat{\lambda}_{MV}$  es insesgado, ya que  $E[\hat{\lambda}_{MV}] = E[\bar{y}] = (1/n) \sum_i E(y_i) = \lambda$ .

#### 4. Parámetros condicionados

Resulta conveniente establecer que una población proviene de una distribución definida. Por ejemplo, natural es establecer que el número de hijos ( $Y$ ) de una familia puede ser caracterizado a través de una distribución de Poisson. Sin embargo, no es aplicable establecer para todas las familias el mismo valor esperado de hijos. Una forma de establecer esta diferencia es condicionar el parámetro de la distribución a otras características del individuo ( $X$ ). De este modo, se puede definir convenientemente (ya que con logaritmo permite evitar la restricción  $\lambda_i \geq 0$ )  $\log \lambda_i = \alpha + \beta x_i$ , así el logaritmo de la función de probabilidad será:  $\log p(y_i) = -\exp(\alpha + \beta x_i) + y_i(\alpha + \beta x_i) - \log y_i!$ .

Otra distribución ampliamente utilizada es la Bernoulli, en cuyo caso las funciones para el parámetro de la probabilidad ( $\theta$ )<sup>2</sup> se define preferentemente de dos formas conocidas como logit y probit. En el caso del logit se tiene que  $\theta_i = \exp(\alpha + \beta x_i) / [1 + \exp(\alpha + \beta x_i)]$  que implica un logaritmo de la función de probabilidad igual a:  $\log p(y_i) = y_i(\alpha + \beta x_i) - \log[1 + \exp(\alpha + \beta x_i)]$ . Mientras que para el caso del probit  $\theta_i = \Phi(\alpha + \beta x_i)$ , entonces su función de probabilidad en logaritmo corresponderá a:  $\log p(y_i) = y_i \log[\Phi(\alpha + \beta x_i)] + (1 - y_i) \log[\Phi(-\alpha - \beta x_i)]$ .

## 2. Método de Momentos (MM)

---

<sup>2</sup>Para evitar confusiones se considera  $\theta$  en vez de  $p$ .

# 6

## Tests e Intervalos de Confianza

En este capítulo se presentan los test tradicionales en inferencia estadística y se agrega una sección que explora la construcción de intervalo de confianza a través de la simulación aleatoria de remuestreo y método de Monte Carlo.

### 1. Dótimas

En este apartado se revisan los casos de test de hipótesis que se trabajaron en el curso. Deben crear un archivo de Excel que les permita trabajar sobre los cálculos de estas estadísticas. Además se debe incluir un valor crítico. En lo que sigue del texto se utilizan las funciones descriptivas (3.1) y (3.2).

#### 1. Una muestra

Los resultados derivados para una muestra se derivan de la materia que hemos visto. Para ello haremos el supuesto de que la muestra proviene de una distribución  $N(\mu, \sigma^2)$

#### Casos

1. Haciendo una hipótesis sobre la media con  $\sigma^2$  conocido será presentado como  $H_0 : \mu = \mu_0$ . Si nos basamos en el Teorema 7.9, trabajaremos con una distribución  $\Phi$  y utilizamos la estadística

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

2. Haciendo una hipótesis sobre la varianza ( $H_0 : \sigma^2 = \sigma_0^2$ ). Si nos basamos en el Teorema 7.12, trabajaremos con una distribución Chi-cuadrado con  $n - 1$  grados de libertad  $(\chi_{n-1}^2)^1$  y utilizamos la

---

<sup>1</sup>Esta distribución es asimétrica por tanto debemos tener más cuidado con la cola de la distribución a la que se refiere la Hipótesis alternativa.

estadística

$$c = \frac{(n-1)s^2}{\sigma_0^2}$$

3. Para una hipótesis sobre la media con  $\sigma^2$  desconocido nos basaremos en el Teorema 5.2 y los dos puntos anteriores. Así, trabajaremos con una distribución  $t_{n-1}$ <sup>2</sup> y utilizamos la estadística

$$t = \frac{z}{\sqrt{c/(n-1)}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

## 2. Dos muestras

La idea principal es que cuando se desea testear la diferencia entre las medias se hace a través de una resta, en cambio cuando se quiere verificar diferencia entre las varianzas se realiza a través de un cociente. Diremos que existen 2 poblaciones de va.  $X$  e  $Y$  que son normales e independientes con medias  $\mu_x$  y  $\mu_y$ , las condiciones sobre las varianzas se darán en cada caso. Nuestras muestras serán de tamaños  $n_x$  y  $n_y$ .

### Casos

1. Queremos hacer  $H_0 : \mu_x - \mu_y = \delta$  y conocemos  $\sigma_x$  y  $\sigma_y$ . Basados en una  $\Phi$  utilizamos la estadística

$$z = \frac{(\bar{x} - \bar{y}) - \delta}{\sqrt{\sigma_x^2/n_x + \sigma_y^2/n_y}}$$

Notemos que cuando se cumple que  $\sigma_x = \sigma_y = \sigma_0$  entonces la estadística se simplifica a:

$$z = \frac{(\bar{x} - \bar{y}) - \delta}{\sigma_0 \sqrt{1/n_x + 1/n_y}}$$

2. Análogo al caso con una muestra, nos interesará estudiar cuando las varianzas de las poblaciones son ambas iguales a un valor puntual, diremos  $H_0 : \sigma_x = \sigma_y = \sigma_0$ . Computando  $s_p^2 = (n_x - 1)s_x^2 + (n_y - 1)s_y^2$  y basados en el Teorema 5.1, utilizaremos una distribución  $\chi_{n_x+n_y-2}^2$  y nuestra estadística será:

$$c = \frac{s_p^2}{\sigma_0^2}$$

3. Acá nos interesará tratar  $H_0 : \mu_x - \mu_y = \delta$  sin conocer  $\sigma_x$  y  $\sigma_y$ , pero asumiendo que son iguales entre

---

<sup>2</sup>Simétrica como la normal. Ver Figura 6.

sí. El problema es similar al caso de una muestra y la estadística sigue una distribución  $t_{n_x+n_y-2}$ . Si consideramos  $s_q^2 = s_p^2/(n_x + n_y - 2)$ , entonces el cálculo es:

$$t = \frac{(\bar{x} - \bar{y}) - \delta}{s_q \sqrt{1/n_x + 1/n_y}}$$

4. ¿Será posible creer que  $\sigma_x = \sigma_y$ ?. Utilizando el Teorema 5.3 podemos revisar  $H_0 : \phi\sigma_x^2 = \sigma_y^2$  a través de una distribución  $F_{(n_x-1, n_y-1)}$ <sup>3</sup> con la estadística:

$$f = \frac{\phi s_x^2}{s_y^2}$$

### 3. Ejemplos

1. La resistencia a la rotura de los cables producidos por un fabricante tienen una media de 1800 libras y una desviación ( $\sigma$ ) de 100 libras. Mediante una nueva técnica en el proceso de fabricación se aspira a que esta resistencia pueda ser incrementada. Para ensayar esta aspiración, se ensaya una muestra de 50 cables y se encuentra que su resistencia media es de 1850 libras. ¿Puede mantenerse que, en efecto, hay un aumento de resistencia al nivel de significación del 0.01?

Utilizando el estadístico correspondiente:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1850 - 1800}{100\sqrt{50}} = 3,54$$

Utilizando la función `difmedias(1850;0;100;0;50;0;1800;0.02)=(3.54;2.33;3.54;2.4)`. Siendo el valor del estadístico, seguido del valor normal y finalmente el  $t$ . Todos los valores de la tabla son de dos colas al 2%, dado que el test es para para una cola al 1%. El test es rechazado.

2. Se hizo un examen a dos clases formadas por 40 y 50 estudiantes respectivamente. En la primera clase la puntuación media fue de 74 con una desviación de 8, mientras que en la segunda clase la puntuación media fue de 78 con una desviación de 7. ¿Hay una diferencia significativa entre el resultado de las dos clases al nivel de significancia de 5%?.

---

<sup>3</sup>Recordar que la distribución  $F$  también es asimétrica.

Utilizando el estadístico correspondiente:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{74 - 78}{\sqrt{\frac{8^2}{40} + \frac{7^2}{50}}} = -2,49$$

Utilizando la función `difmedias(74;78;8;7;40;50;0;0.05)` = (-2.49; 1.96; -2.53; 1.99). En esta oportunidad alfa se considera al 5% porque el test se requiere con dos colas. El test es rechazado.

3. Un ensayo sobre resistencia a la rotura de 6 cuerdas fabricadas por una compañía mostró una resistencia media de 7750 lb y una desviación de 145 lb mientras que el fabricante sostenía que la resistencia media de sus cuerdas era de 8000 lb. ¿Se puede admitir que la afirmación del fabricante al 5%?.

Utilizando el estadístico correspondiente:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7750 - 8000}{145\sqrt{6}} = -4,22$$

Utilizando la función `difmedias(7750;0;145;0;6;0;8000;0.1)` = (-4.22; 1.64; -4.22; 2.02). En este caso la hipótesis requiere de una cola por lo que  $\alpha$  corresponde a 10%. El test es rechazado.

4. En el pasado la desviación de los pesos de ciertos paquetes de 40 onzas, llenadas por una máquina era de 0.25 onzas. Una muestra aleatoria de 20 paquetes dio una desviación de 0.32 onzas. ¿Es el aparente incremento de variabilidad significativa al nivel de significancia de 5%?.

Utilizando el estadístico correspondiente:

$$c = \frac{(n-1)s^2}{\sigma_0^2} = \frac{19 \times (0,32)^2}{(0,25)^2} = 31,13$$

Utilizando la función `igualvar(0.32;0;20;0;0.25;0.1)` = (31.13; 10.12; 30.14). En este caso la hipótesis requiere de una cola por lo que  $\alpha$  corresponde a 10%. El test es rechazado pues el estadístico está fuera del intervalo (10.12;30.14).

5. Un instructor tiene dos clases: A y B, en una asignatura específica. La clase A tiene 16 estudiantes en tanto que la clase B tiene 25 estudiantes. En el mismo examen, aunque no hubo diferencia significativa en medias de las calificaciones, la clase A tuvo una desviación de 9 en tanto que la clase B tuvo una

desviación de 12. ¿Podemos concluir al nivel de 1% que la variabilidad de la clase B es mayor que la de la clase A.

Utilizando el estadístico correspondiente:

$$f = \frac{\phi s_x^2}{s_y^2} = \frac{1 \times (9)^2}{(12)^2} = 0,56$$

Utilizando la función `difvar(9;12;16;25;1;0.02)=(0.56;0.30;3.29)`. El test no es rechazado pues el estadístico está dentro del intervalo (0.30;3.29).

## 2. Ajuste de Distribución

En este capítulo la hipótesis nula será  $H_0 : F(x) = F_0(x)$ , que equivale a decir que cierta muestra aleatoria proviene de una distribución definida.

### 1. Prueba Chi-cuadrado

$$w_1 = \sum_{i=1}^n \frac{(o_i - e_i)^2}{e_i}$$

Se utiliza la función `prueba.chi()`, donde el primer argumento corresponde a los datos observados y el segundo a los esperados. La función entrega la probabilidad del valor de  $w_1$ . La corrección de Yates es

$$w_1^* = \sum_{i=1}^n \frac{(|o_i - e_i| - 0,5)^2}{e_i}$$

### 3. Tablas de Doble Entrada

$$w_2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

Puede utilizarse la misma función `prueba.chi()` de Excel.

## 4. Remuestreo

Retomemos el principio que prima sobre una muestra, esto es que se espera sea una buena o aceptable representación de las características de la población.<sup>4</sup> Bajo este concepto regenerar una nueva muestra a partir de la actual significa que cada elemento representativo tiene igual probabilidad de ocurrencia y por tanto la nueva muestra debiera conservar las características de la población. Este concepto se conoce como *bootstrapping* y es altamente utilizado para generar IC sobre estadísticas de las cuales se desconoce la distribución.

### 1. Ejemplos

1. Los siguientes datos provienen de una distribución normal con media  $\mu$  y varianza  $\sigma^2$ . Realizando 200 repeticiones de tamaño 10.

1.3 2.1 0.4 1.3 0.5 0.2 1.8 2.5 1.9 3.2

Solución: Primero se debe generar una tabla como la que aparece abajo. Luego utilizar la función `buscar(aleatorio();Rango F(x);Rango x)` para crear la celda que genera la distribución inversa, vale decir basta colocar como celda objetivo en el menú **Simula**.

$x$	1.3	2.1	0.4	1.3	0.5	0.2	1.8	2.5	1.9	3.2
$p(x)$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
$F(x)$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

2. Genere 50 muestras, cada una de tamaño 25 a partir de una distribución normal con media 60 y desviación estándar 10.

Solución: Basta crear la celda objetivo como: `distr.norm.inv(aleatorio();60;10)`.

## 5. Simulación de Monte Carlo

Cuando se comentó la generación de números aleatorios se introdujo el concepto de generar características poblacionales en base a la evidencia encontrada con la muestra. De esta forma las herramientas presentadas en esta sección son mecanismos auxiliares a la teoría de distribución.

<sup>4</sup>Si esta muestra es obtenida con reemplazo aseguramos que sea un ma independiente.



En un 99% de los casos nuestros estudios estarán basados sobre una muestra. Sin embargo, nos interesaría ir cada vez más cerca de la población. De este modo, lo que se pretende con los *pseudo random number* es generar *poblaciones* sobre la base de parámetros que caractericen a la distribución. Esto es conocido como el método de *Monte Carlo*.<sup>5</sup>

Existen varios mecanismos para la generación de estos números aleatorios, pero por el alcance de este curso utilizaremos los disponibles en el *software* Excel.

### 1. Ejemplo: Simulación de Escenarios

Consideremos la evaluación de un proyecto en el cuál el crecimiento de las ventas ( $g_t$ ) y los costos medios ( $h_t$ ) se consideran variables aleatorias distribuidas normales e independientes. De este modo, las ventas del período  $t$  son proporcionales a las ventas del período anterior:  $V_t = (1 + g_t)V_{t-1}$  y los costos por venta son  $C_t = V_t h_t$ . Considerando una tasa de impuesto de  $\tau$  el flujo por período corresponde a:  $F_t = (V_t - C_t)(1 - \tau)$ , que sumado a una tasa de descuento de  $r$  permite obtener el Valor Presente del proyecto.

$$W = \sum_{i=1}^n \frac{F_t}{(1+r)^i} = \sum_{i=1}^n \frac{(V_t - C_t)(1 - \tau)}{(1+r)^i} \quad (6.1)$$

Para resolver el modelo anterior se requiere de un valor inicial de las ventas ( $V_0$ ). De este modo, (6.1) puede reescribirse como (6.2) donde  $g_t \sim N(\mu_g, \sigma_g^2)$  y  $h_t \sim N(\mu_h, \sigma_h^2)$ .

$$W = (1 - \tau) \sum_{i=1}^n \frac{V_t(1 - h_t)}{(1+r)^i} = V_0(1 - \tau) \sum_{i=1}^n (1 - h_t) \left[ \frac{(1 + g_t)}{(1+r)} \right]^i \quad (6.2)$$

---

<sup>5</sup>Cuando se analicen las distribución de las *estadísticas* se utilizará como medio alternativo *bootstrapping*.

# A

## Apéndice Matemático

### 1. Funciones

**Definición A.1.** *Función* o  $y = f(x)$ : describimos por  $f()$  a una relación que se define desde el conjunto  $X^1$  al conjunto  $Y$ , con la condición que cada elemento de  $X$  o pre-ímagenes posea una y sólo una imagen en el conjunto de  $Y$ . El conjunto  $X$  se conoce como Dominio de la función y a la colección de imágenes de la función se le denomina Recorrido de la función.

**Definición A.2.** *Función Inversa* o  $x = f^{-1}(y)$ : esta función existe si y sólo si cada elemento del conjunto  $Y$  posee una y sólo una pre-imagen en  $X^2$ .

**Definición A.3.** *Función Exponencial* o  $y = \exp(x)$ : esta función tiene como base 2.71828182845905 (según la precisión de Excel), que tiene los siguientes atributos:

1. Es definida positiva, vale decir  $\exp(x) > 0$  para todo  $x$ .
2. La pendiente de la función es igual a la función. Esto quiere decir que  $d\exp(x)/dx = \exp(x)$ .
3. La tasa de crecimiento de la función es creciente e igual a la función. Es decir,  $d^2 \exp(x)/dx^2 = \exp(x)$ .
4. Cuando  $x \rightarrow -\infty$  la función tiende a cero (posee una asíntota en  $y = 0$ ) y cuando  $x \rightarrow +\infty$  la función tiende a  $+\infty$  (explota).

**Definición A.4.** *Función Logarítmica* o  $y = \log(x)$ : corresponde a la función inversa de  $\exp(x)$  y sus tributos son:

1. Existe sólo para los número positivos, vale decir  $\log(x)$  existe si y sólo si  $x > 0$ .
2. La pendiente de la función es igual al inverso de  $x$ . Esto es  $d\log(x)/dx = 1/x$ .

---

<sup>1</sup>En términos generales este conjunto podría tener varias dimensiones.

<sup>2</sup>Para este caso, los conjuntos  $X$  e  $Y$  deben ser unidimensionales.

3. La tasa de crecimiento de la función es decreciente e igual al negativo del cuadrado del inverso de  $x$ .

Es decir,  $d^2 \log(x)/dx^2 = -1/x^2$ .

4. Cuando  $x \rightarrow 0$  la función tiende a  $-\infty$  (posee una asíntota en  $x = 0$ ) y cuando  $x \rightarrow +\infty$  la función tiende a  $+\infty$  (explota).

**Definición A.5.** *Función Gamma:* se define como (A.1). Se cumple además que  $\Gamma(x) = (x-1)\Gamma(x-1)$ , si  $x$  es número enteros tenemos  $\Gamma(x) = (x-1)!$ . Por último  $\Gamma(0,5) = \sqrt{\pi}$ .

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \exp(-t) dt \quad (\text{A.1})$$

**Definición A.6.** *Función Geométrica:* suponemos que  $x$  es una constante, entonces:

$$S(n) = \sum_{i=0}^n x^i = \frac{1-x^{n+1}}{1-x}$$

En el caso de que la sumatoria comience en 1 el resultado es:

$$S(n) - 1 = \sum_{i=1}^n x^i = x \left( \frac{1-x^n}{1-x} \right)$$

**Ejemplo A.1.** *Considere una secuencia (P) de pagos C que están descontados a un factor  $x = 1/(1+R)$  por T períodos. ¿Qué ocurre si  $T \rightarrow \infty$ ?*

Se trata de una función geométrica que parte en 1. Esto es:

$$P = \sum_{i=1}^T \frac{C}{(1+R)^i} = C \sum_{i=1}^T x^i = Cx \left( \frac{1-x^T}{1-x} \right) = \frac{C}{R} \left[ 1 - \frac{1}{(1+R)^T} \right]$$

cuando  $T \rightarrow \infty$  el resultado es  $P = C/R$ .

## 2. Diferenciación

**Definición A.7.** *Derivada de una función:* se define como la función resultado de aplicar un cambio marginal a la función, vale decir:

$$\frac{df}{dx} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Notaremos como  $f'(x) = df/dx$  y también pueden ser definidas para órdenes superiores. Por ejemplo la segunda derivada es:

$$\frac{d^2 f(x)}{dx^2} = \lim_{h \rightarrow 0} \frac{f'(x+h) - f'(x)}{h}$$

Considerando  $f(x)$  y  $g(x)$  dos funciones, podemos escribir las propiedades de las derivadas:

1. Para  $h(x) = f(x) + g(x)$ , entonces  $h'(x) = f'(x) + g'(x)$ .
2. Para  $h(x) = f(x)g(x)$ , entonces  $h'(x) = f'(x)g(x) + f(x)g'(x)$ .
3. Para  $h(x) = f(x)/g(x)$  y  $g(x) \neq 0$ , se tiene que  $h'(x) = [f'(x)g(x) - f(x)g'(x)]/[g(x)^2]$ .
4. Para  $h(x) = f[g(x)]$ , entonces  $h'(x) = [df/dg]g'(x)$  (Regla de la Cadena).

**Definición A.8.** *Aproximación de Taylor de orden  $n$  o  $AT(n)$ : Sea  $f(x)$  una función real  $n$  veces continuamente diferenciable en el intervalo  $I = [a, b]$ , con la  $n$ -ésima derivada denotada por  $f^{(n)}(x)$  y sea  $x_0 \in I$ . Se dice  $f(x)$  puede aproximarse según Taylor de orden  $n$ , como se muestra en (A.2).*

$$f(x) \approx f(x_0) + \sum_{k=1}^n \frac{(x-x_0)^k}{k!} f^{(k)}(x_0) \tag{A.2}$$

**Ejemplo A.2.** *Demuestre que la función exponencial puede ser escrita como una serie infinita, esto es:*

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Consideremos una aproximación de Taylor de orden infinito (por tanto la relación es exacta) con  $x_0 = 0$ .

$$\exp(x) = \exp(x_0) + \sum_{k=1}^{\infty} \frac{(x-x_0)^k}{k!} \exp(x_0) = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!} = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

### 3. Optimización

Esta sección pretende reordenar la teoría de la maximización de función y relacionarla con la práctica con el uso del PC.

En una primera etapa se estudian las condiciones de Kuhn-Tucker para resolver problemas no lineales y posteriormente la solución iterativa con el algoritmo de búsqueda de Newton. Programación Clásica Se trata

de maximizar o minimizar una función continua eligiendo valores del vector que pertenecen a un conjunto factible que es cerrado. Por el Teorema de *Weierstrass*, se nos asegura que existe una solución a ese problema.

Un algoritmo constituye un mecanismo por el cual de forma iterativa se llega a una solución. La idea es que partiendo de un valor inicial  $\theta_0$  el se defina un nuevo valor:  $\theta_1 = \theta_0 + \lambda_0 \Delta_0$ , donde  $\Delta_0$  corresponde a la dirección y  $\lambda_0$  al tamaño del paso. De este modo se detiene el mecanismo de búsqueda una vez que bajo algún criterio  $\theta_t \simeq \theta_{t+1}$ . Nuestro interés estará en analizar sólo los algoritmos que se basan en el Método del Gradiente.

**Definición A.9.** “*Ascendente*”: considerando al gradiente como:  $g$  y al Hessiano como:  $H$ . Entonces:

$$\theta_{t+1} = \theta_t - \frac{g'_t g_t}{g'_t H_t g_t} g_t.$$

**Definición A.10.** “*Newton*”: considera la expansión de la CPO en serie de Taylor de primer orden.<sup>3</sup> Esto es:

$$\frac{\partial F}{\partial \theta} \simeq g^0 + H^0(\theta - \theta^0) = 0$$

Despejando podemos escribir:  $\theta_{t+1} = \theta_t - H_t^{-1} g_t$ .

---

<sup>3</sup>También puede ser derivado de Newton-Raphson: algoritmo de búsqueda de raíces de una función como  $y = f(x)$ . Por definición:

$$f'(x) = \frac{df(x)}{dx} = \frac{y_1 - y_0}{x_1 - x_0}$$

Imponiendo  $y_1 = 0$ ,  $f'(x)$  en  $x_0$  y reordenando tenemos que:

$$x_1 = x_0 - \frac{y_0}{f'(x_0)}$$

# B

## Microsoft Excel

### 1. Introducción

El actual `Excel` es el nieto de `VisiCal`, un programa desarrollado para Apple II por estudiantes de la escuela de negocios de Harvard en los años 80. Dicho programa revolucionó a la industria con el concepto de planilla de cálculos con filas y columnas. Posteriormente se incluyeron a la planilla de cálculos gráficos, que permitieron representar con mayor claridad y precisión los resultados de la planilla. Hablamos de `Lotus 1-2-3` desarrollado para IBM.

Aunque `Excel` es un programa relativamente amigable está diseñado para trabajos relativamente sencillos. Al aumentar la información en un archivo pueden producirse error de cálculo notoriamente importantes. Del mismo modo, la precisión en el número por lo general se ubica en los 32-bit<sup>1</sup> con lo que podrían generarse diferencias numéricas al realizar operatorias en diferentes órdenes.

Por otro parte, se debe tener conciencia que el programa fue diseñado como una solución amigable para la contabilidad y no para la estadística ni para las matemáticas. Por ello operaciones de mayor complejidad matemática requieren ser chequeadas con otros programas más especializados.

### 1. Funciones Útiles

Una pequeña lista de funciones en `Excel` incluye:

1. `desref()`: permite referirse a una celda en particular. Su primer argumento es una celda de posición, el segundo y tercero indican cuántas filas y columnas nos moveremos. Los argumentos cuarto y quinto nos permiten entregar referencias que incluyan más de una celda.
2. `esnumero()`: permite saber si una expresión contiene un número o no.

---

<sup>1</sup>Este punto es condicional al equipo donde se encuentre instalado el programa.

## 2. Tablas Dinámicas

El menú **Datos** ofrece un amplio manejo de los datos como si fuese un pequeño administrador de base de datos. Por un lado, los filtros permite que el computador ordene los tipos de elementos presentados en la hoja de cálculo. Por otro las **Tablas Dinámicas** ofrecen la posibilidad de entregar resultados en tablas de dos entradas.

## 3. Herramientas para el análisis

Con esta se aumentan las funciones matemáticas y financieras <sup>2</sup>. Además se inserta una opción llamada **Análisis de Datos** que trabaja eficientemente.

## 4. Ejemplos

1. Suponga que cuenta con datos en la columna A y desea obtener en la columna B promedios móviles.

Utilizando esta fórmula: `promedio(desref(A4;-2;0;3;1))` puede hacerlo.

2. Suponiendo que se cuenta con los siguientes datos:

País	PIB	Año
A	100	1950
A	120	1960
B	150	1950
B	200	1960

Podemos generar una tabla de dos entradas que presente los años (filas) y países (columnas) utilizando **Tablas Dinámicas**.

3. Suponga que contamos con una serie de datos diarios de los años 1999 al 2001. Para cada uno de ellos poseemos la variable de interés y la **fecha**. A partir de esta información podemos calcular los promedios mensuales creando una nueva serie a partir de la función `100año(fecha)+mes(fecha)`. Utilizando la tabla dinámica basamos nuestros cambios sobre la base de esta nueva variable.
4. Considerando el archivo  `analisis.xls`, hoja `ej2`, se requiere graficar el promedio y la desviación estándar mensual para el TCO.

---

<sup>2</sup>También puede ser cargada Herramientas para el análisis VBA.

Para ello se ubica al lado de valor un código que permita identificar rápidamente un mes dentro de un año. Para la fecha ubicada en la posición **a3**, la función será **mes(a3)\*10000+año(a3)**. Luego vamos al menú de Datos y de allí a Tablas dinámicas. Seleccionamos la nueva variable y el TCO. Es necesario que cada serie tenga asociado el nombre en una casilla, por tanto la selección parte en la fila 2. Una vez obtenidos los resultados agregados de las tablas dinámicas, podemos graficar en 2 ejes.

## 2. Gráficos

En este apartado se discute sobre algunos gráficos útiles para la presentación de información.

1. Gráfico de Barras: esta opción genera histogramas. Al graficar las frecuencias de los intervalos y luego en las opciones del formato de las series se deja el ancho del rango igual a 0.

**Ejemplo B.1.** *Considere que posee la siguiente tabla de frecuencia: Genere un histograma.*

2. Gráfico de Dispersión: este tipo de gráficos permite relacionar dos variables y se utiliza para la búsqueda de alguna forma funcional que mejor las aproxime. La incorporación de un rótulo en cada uno de los puntos, que no corresponda a alguna de las series presentadas en los ejes, no es una tarea muy sencilla, aunque siempre se encuentra la alternativa de agregar un texto. Si los rótulos son números y estos no son muy distintos se puede utilizar el gráfico de burbujas, colocando la serie del rótulo como tamaño de la burbuja.

**Ejemplo B.2.** *Considere la siguiente tabla con datos.*

3. Gráfico Tridimensional: Excel nos permite graficar una función del tipo  $y = f(x_1, x_2)$ . Para ello debemos generar una matriz donde el argumento sea la función evaluada en las combinaciones  $(x_1, x_2)$ .

**Ejemplo B.3.** *Consideremos la función  $y = x_1^2 - x_2^2$ . Para ello anotaremos en la celda **b1=-2**, y en la celda **c1=b1+0.05**. Copiando esta fórmula hacia la derecha hasta **cd1**. Similarmente copiamos en **a2=-2** y en **a3=a2+0.05** hasta **a82**.*

*Colocamos la función en la celda **b2=\$a2^2-(b\$1^2)**. Y la copiamos hasta la posición **cd82**. Manteniendo el área seleccionada vamos a insertar gráfico y seleccionamos superficie.*

4. Gráfico de Cotizaciones: de una línea más financiera estos gráficos son utilizados para observar la volatilidad dentro del día de las observaciones.



### 3. Optimización

Se ubican en el menú de **Herramientas...**

#### Buscar Objetivo

Consideremos un polinomio función de una variable  $x$  ( $p(x)$ ). Sabemos que se tienen tantas raíces como grados puede tener el polinomio.

Por ejemplo si queremos saber cuáles son las raíces de  $x^2+2x-3$ , podemos factorizar en  $(x+3)(x-1)$  con lo que las soluciones son 1 y -3. Anotaremos en **a1=0** (punto a cambiar) y en **a2=+a1^2+2\*a1-3** (celda objetivo). El valor que se requiere es cero. El resultado es 1 (0.999917). Luego cambiamos **a1=-1** y encontramos la raíz -3 (-3.000023). En este caso el ejercicio es más fácil resolverlo a través de la ecuación de las soluciones de una cuadrática, que se establece para todo polinomio de la forma  $ax^2 + bx + c$  como se presenta en (B.1)

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2 \pm \sqrt{2^2 - 4(-3)}}{2} = \frac{-2 \pm 4}{2} = -1 \pm 2 \quad (\text{B.1})$$

En el apartado (B.4) se muestra la forma de crear una fórmula para solucionar cuadrática.

#### Solver

Este complemento<sup>3</sup> permite la optimización de una función, con restricciones sobre la base de algunos parámetros dentro del archivo. Si esta opción no esta disponible puede ser cargado el **Complemento** asociado a él. Los complementos son rutinas adicionales que carga Excel sólo cuando le es necesario de manera de hacer más eficiente el uso de la aplicación.

### 4. Programando en Excel

En esta sección vamos a construir funciones en Excel, para ello entraremos en la programación de VBA (Visual Basic for Application). Este programa viene asociado a Excel y aparece cuando se presiona **Alt+F11**. Los módulos son los espacios donde se coloca el código de la programación. Los formularios son diseños que pueden utilizarse para crear menús especiales. En este apartado no se trabajará este punto.

<sup>3</sup>No todas las versiones de Excel pueden cargar los complementos. Para ello debe reinstalar el programa con estas opciones o copiar los complementos en la carpeta Macros que se ubica dentro de Office y esta dentro Microsoft Office en la carpeta de Archivo de Programas.

**Ejemplo B.4.** *Genere la docima z para 1 muestra.*

Presionando Alt+F11 aparece VBA, luego se insertar un Modulo, que es un archivo de texto. En el se escribe el siguiente codigo:

```
Function Docima_Z(Promedio, Mu, Sigma,N)
    Docima_Z =Sqr(N)*(Promedio - Mu)/Sigma
End Function.
```

La funcion puede ser usada insertandola desde “Definidas por el Usuario”.

**Ejemplo B.5.** *Generar una funcion que nos permita encontrar las raices de una cuadratica de la funcion  $ax^2 + bx + c$  en el caso que ellas sean reales.*

Se agrega un argumento (**nro**) que nos indica cual es la solucion que estamos encontrando.

```
Function cuadratica(a, b, c, nro)
    d = Sqr(b*b - 4 * a * c)
    If nro = 1 Then
        cuadratica = (-b + d) / (2 * a)
    Else
        cuadratica = (-b - d) / (2 * a)
    End If
End Function
```

Las funciones anteriormente definidas operan dentro del archivo donde se crea (llamandolas directamente por su nombre) o solo cuando el archivo esta abierto (insertandolas de las “Definidas por el usuario”). Existen dos formas de solucionar este problema:

1. Crear un archivo personal (**funciones.xls**) que se abra cada vez que se inicie Excel, debe ubicarse en la carpeta **iniciar**, que se ubica en el subdirectorio **Office** de **Microsoft Office**. Es importante notar que la generacion de macros desde el ayudante de Excel se guardan en esta direccion en un archivo denominado (**Personal.xls**).
2. Generar un complemento y ubicarlo dentro de la direccion con que se inicia Excel. La diferencia del complemento es que las funciones pueden ser llamadas directamente por su nombre. Para crear el complemento se debe guardar el archivo como tal.

# C

## Código en VBA

En este apéndice se presenta código en *Visual Basic for Application* el cual puede ser utilizado en Excel. Para mejor uso del código se recomienda definir las variables y generar un complemento de Excel.

Los códigos han sido probados, pero existe la posibilidad que presenten algunos errores.

### 1. Funciones

#### 1. Función de Densidad Beta

Asume que  $x$  es una variable aleatoria que se distribuye  $\text{Beta}(\alpha, \beta)$  tal que  $0 \leq x \leq 1$ .

```
Function D.Beta(x, alfa, beta)
    Dim y, num, den as Double
    y = (x ^ (alfa - 1)) ((1 - x) ^ (beta - 1))
    num = Application.GammaLn(alfa + beta)
    den = Application.GammaLn(alfa)+Application.GammaLn(beta)
    D_Beta = y * Exp(num-den)
End Function
```

#### 2. Función Momento Muestral

Requiere una serie de datos (DatosX), el r-ésimo momento (R) y si es centrado (1) o no (0).

```
Function MOMENTO(DatosX, R, Centrado)
    Dim x()
    N = Application.Count(DatosX)
    M = Application.Average(DatosX)
    la = 1
    ReDim x(1 To N)
    If Centrado <>1 Then
        la = 0
    End If
    For i = 1 To N
```

```

        x(i) = Application.Power((Application.Index(DatosX, i) - la * M), R)
    Next i
    MOMENTO = Application.Average(x())
End Function

```

### 3. Correlación Bivariada Spearman

Sus argumentos son las series  $X$  e  $Y$  que se desea comparar. Su resultado es una matriz de  $3 \times 1$ , que entrega la estimación puntual del estadístico ( $\zeta$ ), el valor  $t$  bajo  $H_0 : \zeta = 0$  y su  $p$ -value.

```

Function SPEARMAN(Matriz1, Matriz2)
    Dim y1(), y2(), result()
    N = Application.Count(Matriz1)
    ReDim y1(1 To N)
    ReDim y2(1 To N)
    ReDim result(1 To 3, 1 To 1)
    For i = 1 To N
        y1(i) = Application.Rank(Application.Index(Matriz1, i), Matriz1, 1)
        y2(i) = Application.Rank(Application.Index(Matriz2, i), Matriz2, 1)
    Next
    result(1, 1) = Application.Pearson(y1, y2)
    k = (N - 2) / (1 - (result(1, 1) * result(1, 1)))
    result(2, 1) = result(1, 1) * Sqr(k)
    result(3, 1) = Application.TDist(Abs(result(2, 1)), N - 2, 2)
    SPEARMAN = result()
End Function

```

### 4. Función Test sobre Medias

Sus argumentos son la información de las series  $X$  e  $Y$ , su diferencia teórica ( $\delta$ ) y el error tipo I ( $\alpha$ ). Su resultado es una matriz de  $4 \times 1$  con el  $z$ , su  $p$ -value, el  $t$  y su  $p$ -value. La función también es válida en el caso de que se desee hacer  $H_0 : \mu = \mu_0$ .

```

Function TESTMEDIAS(MediaX, MediaY, DesvX, DesvY, nX, nY, delta, alfa)
    Dim result()
    ReDim result(1 To 4, 1 To 1)
    nZ = Application.Max(nY, 1)
    Nume = (MediaX - MediaY) - delta
    Den1 = Sqr((DesvX * DesvX) / nX + (DesvY * DesvY) / nZ)
    sp2 = (nX - 1) * (DesvX * DesvX) + (nZ - 1) * (DesvY * DesvY)
    sq = Sqr(sp2 / (nX + nZ - 2))
    Den2 = sq * Sqr(1 / nX + nY / (nZ * nZ))
    result(1, 1) = Nume / Den1
    result(2, 1) = Application.NormInv(1 - alfa / 2, 0, 1)

```

```

    result(3, 1) = Nume / Den2
    result(4, 1) = Application.TInv(alfa, nX + nZ - 2)
    TESTMEDIAS = result()
End Function

```

## 5. Función Test sobre Varianzas

Sus argumentos son la información de las series  $X$  e  $Y$  y el error tipo I. Testea  $H_0 : \sigma_X = \sigma_Y$  cuando  $\sigma_0 = 0$  y  $H_0 : \sigma_X = \sigma_0$  en el otro caso (una muestra). Su resultado es una matriz de  $3 \times 1$  con el valor del estadístico y su IC.

```

Function TESTVARS(DesvX, DesvY, nX, nY, sig0, alfa)
    Dim result()
    ReDim result(1 To 3, 1 To 1)
    la = 1
    If nY <> 0 Then
        la = 0
    End If
    nZ = Application.Max(nY, 2)
    aux1 = (nX - la) * (DesvX * DesvX)
    aux2 = (1 - la) * nX * (DesvY * DesvY) + la * (sig0 * sig0)
    aux3 = Application.GammaInv(alfa / 2, (nX - 1) / 2, 2)
    aux4 = Application.GammaInv(1 - alfa / 2, (nX - 1) / 2, 2)
    aux5 = Application.FInv(1 - alfa / 2, nX - 1, nZ - 1)
    aux6 = 1 / Application.FInv(1 - alfa / 2, nZ - 1, nX - 1)
    result(1, 1) = aux1 / aux2
    result(2, 1) = la * aux3 + (1 - la) * aux5
    result(3, 1) = la * aux4 + (1 - la) * aux6
    TESTVARS = result()
End Function

```

## 6. Función de Autocorrelación

Computa la función de autocorrelación.

```

Private Function AC(Data, Lag)
    Dim base()
    Dim x()
    Dim y()
    N = Application.Count(Data)
    p = N - Lag
    ReDim base(1 To N)
    ReDim x(1 To p)
    ReDim y(1 To p)
    m = Application.Average(Data)

```

```

For i = 1 To N
    base(i) = Application.Index(Data, i) - m
Next
For i = 1 To p
    x(i) = base(Lag + i)
    y(i) = base(i)
Next
AC = Application.SumProduct(x(), y()) / Application.SumSq(base())
End Function

```

## 7. Función de Autocorrelación Parcial

Computa la función de autocorrelación parcial.

```

Private Function PAC(Data, Lag)
    Dim base()
    Dim y()
    Dim x()
    N = Application.Count(Data)
    p = N - Lag
    ReDim base(1 To N)
    ReDim y(1 To p)
    ReDim x(1 To Lag, 1 To p)
    For i = 1 To N
        base(i) = Application.Index(Data, i)
    Next i
    For i = 1 To p
        y(i) = base(i + Lag)
        For j = 1 To Lag
            x(j, i) = base(i + Lag - j)
        Next j
    Next i
    PAC = Application.LinEst(y(), x(), 1, 0)
End Function

```

## 2. Rutinas

### 1. Histograma

El siguiente código puede ser implementado, utilizando una macro adicional que le entregue los parámetros solicitados.

```

Sub Histo(Entrada, Clases, Salida, Mini, Maxi)

```

```

If Salida = ""Then
    salf = ActiveCell.Row
    salc = ActiveCell.Column
    Exit Sub
Else
    salf = Range(Salida).Row
    salc = Range(Salida).Column
    hsal1 = InStr(1, Salida, "!", 1)
    hsal2 = Left(Salida, hsal1 - 1)
    Worksheets(hsal2).Activate
End If
largo = (Maxi - Mini) / Clases
Cells(salf, salc).Activate
'Display results
ActiveCell = "Intervalo"
ActiveCell.Offset(0, 1) = "Inferior"
ActiveCell.Offset(0, 2) = "Superior"
ActiveCell.Offset(0, 3) = "Punto Medio"
ActiveCell.Offset(0, 4) = "Frecuencia"
For k = 1 To Clases
    ActiveCell.Offset(k, 0) = k
    ActiveCell.Offset(k, 1) = Mini + (k - 1) * largo
    ActiveCell.Offset(k, 2) = Mini + k * largo
    ActiveCell.Offset(k, 3) = (ActiveCell.Offset(k, 1) + ActiveCell.Offset(k, 2)) / 2
Next k
rf = Range(ActiveCell.Offset(1, 2), ActiveCell.Offset(Clases, 2))
re = Application.Frequency(Range(Entrada), rf)
For k = 1 To Clases
    ActiveCell.Offset(k, 4).Formula = Application.Index(re, k)
Next k
End Sub

```

## 2. Monte Carlo y Remuestreo

El siguiente código requiere una celda generadora de números aleatorios. En el caso del Monte Carlo esta puede ser producida por una la función de probabilidad inversa, mientras que para el caso de remuestreo un algoritmo de remuestreo con reemplazo puede ser fácilmente implementado.

```

Sub Simular(Filas, Columnas, Entrada, Salida)
    entf = Range(Entrada).Row
    entc = Range(Entrada).Column
    salf = Range(Salida).Row
    salc = Range(Salida).Column
    hent1 = InStr(1, Entrada, "!", 1)
    hent2 = Left(Entrada, hent1 - 1)
    hsal1 = InStr(1, Salida, "!", 1)
    hsal2 = Left(Salida, hsal1 - 1)
    If Salida = ""Then

```

```

    Exit Sub
Else
    Worksheets(hsal2).Activate
End If
Cells(salf, salc).Activate
For i = 0 To Filas - 1
    For j = 0 To Columnas - 1
        ActiveCell.Offset(i, j) = Worksheets(hent2).Cells(entf, entc)
    Next j
Next i
End Sub

```

### 3. Variación Porcentual

Muestra la variación porcentual de los extremos de una selección. En el caso de que la serie este en una columna representa el cociente entre el elemento de la fila mayor dividido por elemento de la fila menor. Si la serie está en una fila, representa el cociente entre el elemento de la columna de la derecha sobre la columna izquierda.

```

Sub VariacionPorc()
    Dim Deci, IniFila, IniColu, TotFila, TotColu As Integer
    Dim Fact, Form0, Form1, Form2, Form3 As Double
    Deci = 3
    Fact = 10 ^ 3
    IniFila = Selection.Row
    IniColu = Selection.Column
    TotFila = Selection.Rows.Count
    TotColu = Selection.Columns.Count
    Form0 = Cells(IniFila, IniColu)
    Form1 = Cells(IniFila + TotFila - 1, IniColu)
    Form2 = Cells(IniFila, IniColu + TotColu - 1)
    If TotFila = 1 And TotColu >1 Then
        Form3 = 100 * (Form2 / Form0 - 1)
    ElseIf TotColu = 1 And TotFila >1 Then
        Form3 = 100 * (Form1 / Form0 - 1)
    Else
        Application.StatusBar =
        Exit Sub
    End If
    Form3 = Round(Fact * Form3) / Fact
    Application.StatusBar = "Variación =" & Form3 & "%"
End Sub

```



#### 4. Interpolación

Permite generar datos interpolados. Para su aplicación se requiere seleccionar un área a interpolar incluyendo los extremos de los datos efectivos. El código reconoce si la interpolación debe realizarse entre columnas o entre filas. Además sombrea en amarillos los datos interpolados y utiliza fórmulas. Sin embargo, el código no distingue si existen datos entre los extremos y los reemplaza sin posibilidad de recuperar los valores originales.

```

Sub Interpolado()
    Dim IniFila, IniColu, TotFila, TotColu As Integer
    Dim Form0, Form1, Form2, Form3, Form4 As String
    IniFila = Selection.Row
    IniColu = Selection.Column
    TotFila = Selection.Rows.Count
    TotColu = Selection.Columns.Count
    Form0 = "-R"& IniFila & "C"& IniColu
    Form1 = "(R"& (IniFila + TotFila - 1) & "C"& IniColu
    Form2 = "(R"& IniFila & "C"& (IniColu + TotColu - 1)
    Form3 = Form1 & Form0 & ")/"& (TotFila - 1) & "+R[-1]C"
    Form4 = Form2 & Form0 & ")/"& (TotColu - 1) & "+RC[-1]"
    If TotFila = 1 And TotColu >1 Then
        For i = 1 To TotColu - 2
            Cells(IniFila, IniColu + i).FormulaR1C1 = Form4
        Next i
        Range(Cells(IniFila, IniColu + TotColu - 2),Cells(IniFila, IniColu + 1)).Select
        With Selection.Interior
            .ColorIndex = 6
            .Pattern = xlSolid
        End With
        Cells(IniFila, IniColu + TotColu - 1).Select
    ElseIf TotColu = 1 And TotFila >1 Then
        For i = 1 To TotFila - 2
            Cells(IniFila + i, IniColu).FormulaR1C1 = Form3
        Next i
        Range(Cells(IniFila + TotFila - 2, IniColu),Cells(IniFila + 1, IniColu)).Select
        With Selection.Interior
            .ColorIndex = 6
            .Pattern = xlSolid
        End With
        Cells(IniFila + TotFila - 1, IniColu).Select
    Else
        Exit Sub
    End If
End Sub

```