# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"UNA CARACTERIZACION DE DISTRIBUCIONES BASADA EN LA ESPERANZA CONDICIONAL DE FUNCIONES DE ESTADISTICOS DE ORDEN"

# T E S I S

QUE PARA OBTENER EL TITULO DE

**A C T U A R I O**

P R E S E N T A :

**ANTONIO SORIANO FLORES**

DIRECTOR DE TESIS: M. en C. HUGO VILLASEÑOR HERNADEZ



FACULTAD DE CIENCIAS
UNAM

2005

VNIVERIDAD NACIONAL AVFNºMA DE MEXICO

**ACT. MAURICIO AGUILAR GONZÁLEZ**
**Jefe de la División de Estudios Profesionales de la Facultad de Ciencias**
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito:

"Una Caracterización de Distribuciones Basada en la Esperanza Condicional de Funciones de Estadísticos de Orden"

realizado por Antonio Soriano Flores

con número de cuenta 09802626-3 , quien cubrió los créditos de la carrera de: Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

| | |
|---|---|
| Director Propietario | M. en C. Hugo Villaseñor Hernández |
| Propietario | Dr. Vicente Ángel Soriano Ramírez |
| Propietario | Act. Noe Moacyr Vallejo González |
| Suplente | Act. Jaime Vázquez Alamilla |
| Suplente | Act. Gerardo Rubio Hernández |

**Consejo Departamental de**
Matemáticas

Act. Jaime Vázquez Alamilla

FACULTAD
CONSEJO DEPARTAMENTAL DE
MATEMATICAS

**Agradecimientos**

A Dios por haberme dejado llegar a este punto tan importante de mi vida.

A la UNAM por darme la oportunidad de estar en sus aulas y formar parte de esta gran insitución.

A mis Maestros y Sinodales ya que gracias a ellos he podido salir adelante con sus consejos y ayuda que me brindaron hasta el último momento de mis estudios.

A mis Padres por su esfuerzo y cariño que me brindaron a lo largo de mi educación

Con todo cariño a Venice por su apoyo y compañía durante mis estudios. Recuerda que siempre serás una persona muy especial para mi.

A mis amigos de toda la vida: Roman, Enrique, Eduardo, Ernesto, Sergio, etc. Gracias a todos ellos por brindarme su valiosa amistad , les deseo que tengan un futuro lleno de éxito.

# Índice general

# Introducción

El presente trabajo aborda el problema de la caracterización de funciones de distribución el cual consiste en identificar de manera única a cada distribución basándose en alguna propiedad que ésta cumpla. Un claro ejemplo de una caracterización lo podemos ver con los seres humanos los cuales estamos identificados de manera única por el ADN o por la huella digital de nuestro dedo pulgar.

Una caracterización de las funciones de distribución ya muy conocida es aquella que se define a partir de la función característica $\phi_X(t)$.Esta caracterización, basada en el concepto de la transformada de Fourier de la función de distribución, es sin duda muy importante para la probabilidad y estadística ya que es usada tanto para calcular momentos como para encontrar la función de distribución de transformaciones de variables aleatorias.

Este trabajo pretende presentar una caracterización de funciones de distribución basada en la esperanza condicional de funciones de estadísticos de orden con lo cual se intenta resolver el problema de encontrar la distribución de donde provino una muestra aleatoria una vez que es conocida la forma que toma la esperanza condicional de los estadísticos de orden definidos a partir de ella.

Para llevar a cabo la caracterización que se pretende, el trabajo es dividido en 3 capítulos de los cuales los 2 primeros se dedicarán a demostrar algunos teoremas y motivar las definiciones de conceptos necesarios para desarrollar todo el capítulo 3 el cual se dedicará a la caracterización.

En el primer capítulo se expondrá todo lo relacionado con los estadísticos de orden iniciando con su definición y terminando con las fórmulas para encontrar la densidad de estos estadísticos tanto en forma conjunta como marginal. Es de destacar que esto sólo es hecho para el caso en que la distribución de donde proviene la muestra es absolutamente continua, ya que la caracterización que se pretende estudiar sólo involucra a este tipo de distribuciones.

En el segundo capítulo se aborda el tema de esperanza condicional en donde se trata de explicar y sobre todo motivar todos los conceptos que hay detrás de este tema tales como probabilidad condicional, densidad condicional, esperanza, etc. Así pues, también se explican algunas propiedades del operador esperanza condicional que serán posteriormente utilizadas en el capítulo tercero.

Será en el capítulo 3, en donde utilizando todas las herramientas desarrolladas en los capítulos precedentes, que se presentarán un total de 6 teoremas con sus respectivas demostraciones que nos ayudarán a llevar acabo la caracterización que se desea.

Por último se presentan algunos ejemplos en donde se caracteriza a las funciones de distribución Weibull, Uniforme y Pareto, así como algunos ejercicios que muestran otras aplicaciones que se pueden obtener de los teoremas del capítulo 3.

# Capítulo 1

# Estadísticos de Orden

## 1.1. ¿Qué es un estadístico?

El problema principal de la inferencia estadística es tratar de dar una aseveración lo más precisa posible acerca del comportamiento general de una población mediante las observaciones en una muestra obtenida a partir del proceso del muestreo.

Para la estimación de parámetros siempre llevamos a cabo una selección de elementos al azar de la población obteniendo con ello una muestra aleatoria que es, bajo el enfoque de la *Estadística Clásica*, la única fuente de información para poder tomar la decisión final [1] .

Cuando obtenemos $X_1, X_2, \ldots, X_n$, una muestra aleatoria de una población, se asume que cada $X_i$ es una variable aleatoria independiente de las restantes e idénticamente distribuida según la distribución de probabilidades que sigue la población.

Parte del proceso que llevamos a cabo para interpretar los datos obtenidos y tomar decisiones es el de definir funciones de la muestra aleatoria con el objetivo de obtener *estimadores* de cierto parámetro desconocido.

$$\hat{\theta} = g_1(X_1, X_2, \ldots X_n)$$

Es claro que siempre tendremos un sin fin de maneras de definir estimadores, tantos como funciones de la muestra podamos definir, sin embargo el trabajo de la inferencia estadística es encontrar aquel estimador, es decir aquella función de la muestra aleatoria, que tenga un mayor número de propiedades para la estimación del parámetro desconocido.

[1] En el enfoque de la Estadística Bayesiana otra fuente de información es el conocimiento subjetivo que se tiene de la población, el cual se ve reflejado en la distribución inicial o a priori del parámetro

Las funciones de variables aleatorias provenientes de una muestra aleatoria son muy importantes para la inferencia y, es por ello, que surge el concepto de "*Estadístico*" para el cual se tiene la siguiente definición:

**Definición 1.1 (Estadístico).** *Un estadístico es una función de variables aleatorias , la cual no debe depender de algún parámetro desconocido*[2] .

**Ejemplo 1.1.** Supongamos que obtenemos $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño $n$ proveniente de una población que sigue una distribución $Bernoulli(\theta)$.

El problema de inferencia comienza en tratar de encontrar cual es el valor verdadero del parámetro desconocido $\theta$.

Con ayuda de la muestra podemos definir funciones tales como:

$$t_1 = f_1(X_1, X_2, ...X_n) = \sum_{i=1}^{n} X_i/n$$

$$t_2 = f_1(X_1, X_2, ...X_n) = \theta \sum_{i=1}^{n} X_i$$

Sin embargo, se observa que $t_2$ no cumple con las condiciones para ser llamado estadístico por el hecho de que la función que lo define está dependiendo del parámetro desconocido $\theta$. De ahí que digamos que $t_2$ no sea un estadístico mientras que $t_1$ sí lo es.

En resumen, podemos decir que la muestra aleatoria obtenida a partir de una población es la materia prima con la que cuenta el investigador que quiere inferir algún comportamiento general de ésta, mientras que el *estadístico* es la materia prima ya transformada en una herramienta que servirá para llevar a cabo el proceso de inferencia.

## 1.2. Definición de estadístico de orden

En la sección anterior se definió lo que se entiende por un estadístico. Por lo que ahora introduciremos el concepto de *estadísticos de orden* así como algunas definiciones importantes para la Estadística que surgen a consecuencia de éstos.

[2] En cursos más avanzados de estadística la definición formal de estadístico es diferente ya que se pide una función *medible* con el objeto de que el Estadístico definido a partir de esta función sea una nueva variable aleatoria. Sin embargo, en el presente trabajo no introduciremos elementos de *Teoría de la medida* por lo que la definición aquí dada está dando por hecho que la función que define al Estadístico es *medible* (Para mayor información acerca de la teoria de la probabilidad vista desde el enfoque de la teoria de la medida puede consultar el libro de ASH [7] )

Los *estadísticos de orden* juegan sin duda un papel muy importante tanto para la estadística paramétrica como la no paramétrica. De hecho, algunas de sus propiedades no dependen de la función de distribución de la población de donde fue obtenida la muestra aleatoria (Puede ver ejemplos de estas propiedades aplicadas en la Estadística no Paramétrica en la obra de Naray C. [5] en el capítulo 2).

Supongamos que tenemos un experimento aleatorio que sigue una distribución de probabilidad especificada por la función $F(\cdot)$. Enseguida repetimos el experimento $n$ veces obteniendo con ello una muestra aleatoria $X_1, X_2, \ldots, X_n$.

A partir de esto definimos lo siguiente:

$$Y_{1:n} = min\{X_1, X_2, \ldots, X_n\}$$

$$Y_{2:n} = min\{\{X_1, X_2, \ldots, X_n\} - \{Y_{1:n}\}\}$$

$$Y_{3:n} = min\{\{X_1, X_2, \ldots, X_n\} - \{Y_{2:n}, Y_{1:n}\}\}$$

$$Y_{4:n} = min\{\{X_1, X_2, \ldots, X_n\} - \{Y_{3:n}, Y_{2:n}, Y_{1:n}\}\}$$
$$\vdots$$
$$Y_{n-1:n} = min\{\{X_1, X_2, \ldots, X_n\} - \{Y_{n-2:n}, Y_{n-3:n}, \ldots, Y_{1:n}\}\}$$

$$Y_{n:n} = min\{\{X_1, X_2, \ldots, X_n\} - \{Y_{n-1:n}, Y_{n-2:n}, \ldots, Y_{1:n}\}\} = max\{X_1, X_2, \ldots, X_n\}$$

Dado lo anterior nos damos cuenta de tres propiedades importantes para estas funciones definidas a partir de la muestra aleatoria:

1. Cada $Y_{i:n}$ es un estadístico según la definición anteriormente dada.
2. Estos estadísticos definidos son tales que ordenan la muestra de menor a mayor magnitud, logrando con ello que $Y_{k:n}$ sea la k-ésima observación de la muestra aleatoria ordenándola según sus magnitudes.
3. Debe ser claro además que no hay independencia entre estos estadísticos ya que si $Y_{j:n} > k$ entonces, $Y_{j+a:n} > k$.

La definición formal quedaría como sigue:

**Definición 1.2 (Estadístico de Orden).** Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria. de tamaño $n$ de un fenómeno aleatorio especificado por la función de distribución $F(\cdot)$. Ordenemos cada colección de valores $x_1, x_2, \ldots, x_n$ de $X_1, X_2, \ldots, X_n$ en orden creciente de magnitudes tales que obtenemos los valores $y_{1:n}, y_{2:n}, \ldots, y_{n:n}$ (conocida como la muestra ordenada), donde $y_{1:n} = min\{x_1, x_2, \ldots, x_n\}$, $y_{2:n}$ es el siguiente $x_i$ en orden de magnitud $, \ldots,$ y $y_{n:n} = max\{x_1, x_2, \ldots, x_n\}$. La variable aleatoria $Y_{k:n}$ $(k = 1, 2, \ldots, n)$ función de

$X_1, X_2, \ldots, X_n$, la cual toma un valor $y_{k:n}$ en cada posible secuencia de $x_1, x_2, ..., x_n$, es llamado el k-ésimo estadístico de orden (Al número $k$ se le conoce como la magnitud de dicho estadístico de orden).

A partir de la definición anterior surgen los conceptos de *Rango Muestral*, *Mediana Muestral* y *Rango-Medio Muestral* los cuales son usados a menudo por parte de la Estadística Descriptiva como medidas de tendencia central y dispersión de la población de donde fue obtenida la muestra aleatoria.

**Definición 1.3 (Mediana Muestral, Rango Muestral, Rango-Medio Muestral).** Sea $Y_{1:n} \leq Y_{2:n} \leq \ldots \leq Y_{n:n}$ los estadísticos de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una función de distribución $F(\cdot)$.

i) La *Mediana Muestral* se define como:

$$Mediana = \begin{cases} Y_{\frac{n+1}{2}:n} & \text{si } n \text{ es impar,} \\ \\ (Y_{\frac{n}{2}:n} - Y_{\frac{n}{2}+1:n})/2 & \text{si } n \text{ es par.} \end{cases}$$

ii) El *Rango Muestral* se define como:

$$R = Y_{n:n} - Y_{1:n} = max\left\{X_1, X_2, \ldots, X_n\right\} - min\left\{X_1, X_2, \ldots, X_n\right\}$$

iii) El *Rango-Medio Muestral* se define como:

$$R_M = \frac{Y_{n:n} - Y_{1:n}}{2}$$

Las definiciones anteriores implican directamente que la Mediana Muestral, el Rango Muestral y Rango-Medio Muestral son nuevamente *estadísticos* y que por lo tanto pueden ser utilizados para llevar a cabo algún tipo de inferencia.

## 1.3. Distribución de los estadísticos de orden

Para saber qué estadístico es mejor que otro, es necesario conocer las propiedades de cada uno y elegir aquél que más se adecue a nuestras necesidades. Con el objeto de conocer mejor estas propiedades es necesario saber cuando menos cuáles son sus momentos centrales y si es posible saber qué tipo de distribución está siguiendo.

Podemos motivar el estudio de la distribución de un estadístico de la siguiente manera. Sea $X$ una variable aleatoria que nos modela alguna característica numérica de algún fenómeno aleatorio. En muchas ocasiones nos enfrentamos al problema de no saber completamente cuál es la función de distribución de esta variable aleatoria. Supongamos que sabemos que la variable aleatoria sigue una distribución $F_X(x;\theta)$ pero con el parámetro $\theta$ desconocido. Para obtener más información acerca de la distribución de $X$, nosotros podemos repetir el experimento $n$ veces de manera independiente y definir $X_i$ la v. a. de la i-ésima repetición del experimento. Entonces $X_1, X_2, \ldots, X_n$ será una muestra aleatoria de la distribución en estudio con la cual se pretenderá estimar el valor de $\theta$ a través de una función de ella. Ahora suponga que definimos un estadístico $Y = u(X_1, X_2, \ldots, X_n)$ para el cual podemos encontrar su función de densidad de probabilidad $G_Y(y)$ y supongamos que esta función nos muestra que existe una probabilidad considerable de que $Y$, al realizar los experimentos, tenga un valor cercano al parámetro desconocido. Así, una vez que el experimento ha sido repetido en $n$ ocasiones se obtiene una muestra observada $x_1, \ldots, x_n$ la cual al ser evaluada en la función de la muestra $u(x_1, x_2, \ldots, x_n)$ nos dará un número conocido el cual esperamos, bajo las condiciones arriba expuestas, sea cercano al valor desconocido del parámetro $\theta$.

Observe que de lo anterior para saber qué tan cercano será el valor estimado al parámetro desconocido $\theta$, es necesario conocer la forma de la distribución del estadístico que lo define, para así tener una medida de probabilidad de qué tan cercana será nuestra estimación.

La distribución de probabilidad de un estadístico es un problema que hay que afrontar en la inferencia y, por lo tanto, el objetivo se convierte en buscar métodos capaces de resolver este problema.

Para nuestro caso, el problema consiste en encontrar la distribución que siguen cada uno de nuestros estadísticos de orden, tanto de forma individual (densidad marginal) como de forma conjunta.

Hasta este momento no hemos hecho distinción entre la variable aleatoria absolutamente continua y la discreta ya que todos las definiciones pueden ser hechas para ambas clases de variables aleatorias. A continuación enunciaremos nuestro primer teorema el cual nos da la distribución del estadístico $Y_{\alpha:n}$ conociendo la función de distribución de la población de donde fue obtenida la muestra aleatoria y el cual será válido para ambos tipos de variables aleatorias. (*Absolutamente continuas y discretas*).

**Teorema 1.1 (Distribución del estadístico $Y_{\alpha:n}$).** *Sea $Y_{1:n} \leq Y_{2:n} \leq \ldots \leq Y_{n:n}$ los estadísticos de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una función de distribución $F(\cdot)$ conocida. Entonces la función de distribución acumulativa de $Y_{\alpha:n}$, $\alpha = 1, 2, \ldots, n$ está dada por:*

$$F_{Y_{\alpha:n}}(y) = \sum_{j=\alpha}^{n} \binom{n}{j} [F(y)]^j [1 - F(y)]^{n-j} \tag{1.1}$$

*Demostración.* Sea $y$ fija y arbitraria y definamos $Z_i = I_{(-\infty,y]}(X_i)$.

Dado lo anterior notemos lo siguiente:

- $Z_i$ es una variable aleatoria por estar en función de $X_i$
- $\sum_{i=1}^{n} Z_i =$ número de $X_i's$ tales que son menores o iguales a $y$
- $Z_i$ sigue una distribución $Bernoulli(F(y))$
- $\sum_{i=1}^{n} Z_i$ sigue una distribución $Binomial\,(n, F(y))$

Ahora calculemos $F_{Y_{\alpha:n}}(y)$

$$F_{Y_{\alpha:n}}(y) = P\left[Y_{\alpha:n} \leq y\right] = P\left[\sum_{i=1}^{n} Z_i \geq \alpha\right] = \sum_{j=\alpha}^{n} \binom{n}{j} \left[F(y)\right]^j \left[1 - F(y)\right]^{n-j}$$

□

Note que la clave de la prueba anterior es la equivalencia entre los eventos $\{Y_\alpha \leq y\}$ y $\{\sum_{i=1}^{n} Z_i \geq \alpha\}$. Esto pasa porque si el $\alpha$ estadístico de orden es menor o igual a $y$, entonces forzosamente el número de $X_i's$ menores o iguales a $y$ es más grande o incluso igual a $\alpha$, y viceversa.

Enseguida deducimos el siguiente corolario

**Corolario 1.1.**

$$F_{Y_{n:n}}(y) = \left[F(y)\right]^n \tag{1.2}$$

$$F_{Y_{1:n}}(y) = 1 - \left[1 - F(y)\right]^n \tag{1.3}$$

*Demostración.* La demostración de esto último consiste en hacer de la ecuación (1.1) $\alpha =$ n y $\alpha = 1$ respectivamente. □

### 1.3.1. Distribución de los estadísticos de orden (caso absolutamente continuo)

En esta parte únicamente nos enfocaremos al caso en que la muestra aleatoria es obtenida de una población que sigue una distribución acumulativa absolutamente continua por lo que supondremos que los elementos de la muestra $X_i$ tienen una función de densidad $f(\cdot)$ continua la cual es obtenida tras derivar a su respectiva función de distribución.

Ahora nos enfrentaremos al problema no sólo de encontrar las distribuciones de los estadísticos de orden, sino que también sus respectivas densidades conjuntas y marginales. El propósito de este estudio más detallado es porque el problema principal de este trabajo versa únicamente acerca de las distribuciones absolutamente continuas.

Lo primero que vamos a deducir son las densidades marginales de $Y_{1:n}$ y $Y_{n:n}$ de las ecuaciones (1.2) y (1.3) pues lo único a lo que procederemos es a derivar respecto a $y$ cada una de las ecuaciones anteriores:

$$f_{Y_{1:n}}(y) = \frac{d}{dy} F_{Y_{1:n}}(y) = n\left[1 - F(y)\right]^{n-1} f(y) \tag{1.4}$$

$$f_{Y_{n:n}}(y) = \frac{d}{dy} F_{Y_{n:n}}(y) = n\left[F(y)\right]^{n-1} f(y) \tag{1.5}$$

De igual forma podemos proceder con la ecuación (1.1) para así encontrar la densidad del $\alpha$ estadístico de orden:

$$f_{Y_{\alpha:n}}(y) = \frac{d}{dy} F_{Y_{\alpha:n}}(y) = \frac{d}{dy} \sum_{j=\alpha}^{n} \binom{n}{j} \left[F(y)\right]^{j} \left[1 - F(y)\right]^{n-j} \tag{1.6}$$

Sin embargo, exhibiremos otra demostración, basada en la prueba hecha en el libro de Mood [3], la cual creemos es más fácil de comprender.

**Teorema 1.2 (Densidad del estadístico $Y_{\alpha:n}$).** *Sea $Y_{1:n} \leq Y_{2:n} \leq \ldots \leq Y_{n:n}$ los estadísticos de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una función de distribución $F(\cdot)$ continua y conocida. Entonces la función de densidad $Y_{\alpha:n}$, $\alpha = 1, 2, \ldots, n$ está dada por:*

$$f_{Y_{\alpha:n}}(y) = \frac{n!}{(\alpha - 1)!\,(n - \alpha)!} \left[F(y)\right]^{\alpha-1} \left[1 - F(y)\right]^{n-\alpha} f(y) \tag{1.7}$$

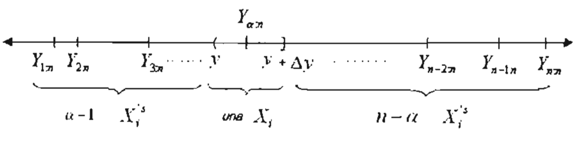
*Demostración.* Sabemos que $f_{Y_{\alpha:n}}(y) = \frac{d}{dy} F_{Y_{\alpha:n}}(y)$.

Entonces por definición de derivada

$$\frac{d}{dy} F_{Y_{\alpha:n}}(y) = \lim_{\Delta y \to 0} \frac{F_{Y_{\alpha:n}}(y + \Delta y) - F_{Y_{\alpha:n}}(y)}{\Delta y} = \lim_{\Delta y \to 0} \frac{P\left[y < Y_{\alpha:n} \leq y + \Delta y\right]}{\Delta y} \tag{1.8}$$

Ahora notemos que el evento $\{y < Y_{\alpha:n} \leq y + \Delta y\}$ es equivalente al evento $\{(\alpha - 1)$ de las $X_i's \leq y$; una $X_i$ en $(y, y + \Delta y]$; $(n - \alpha)$ de las $X_i's > y + \Delta y$ $\}$ (Ver figura 1.1).

Figura 1.1:




Por lo que ahora nuestro problema consiste en encontrar la probabilidad de este segundo evento, el cual se encuentra fácilmente con ayuda de una densidad multinomial.

Al momento de hacer el muestreo, nosotros podemos tomar como un ensayo el hecho de obtener un valor $x_i$ de la variable aleatoria $X$ el cual nos conduce a 3 posibles resultados:

1. $x_i \leq y$ lo cual ocurre con probabilidad $F(y)$
2. $y < x_i \leq y + \Delta y$ lo cual ocurre con probabilidad $F(y + \Delta y) - F(y)$
3. $x_i > y + \Delta y$ lo cual ocurre con probabilidad $1 - F(y + \Delta y)$

Con lo anterior agrupamos a la muestra aleatoria en tres grupos cada uno de los cuales tendrá un número aleatorio de elementos. Si definimos $Z_j$ = número de ocurrencias del resultado $j = 1, 2, 3$ (Es decir $Z_j$ será el número de elementos del grupo j) tendríamos que la distribución conjunta de $Z_1, Z_2, Z_2$ es una multinomial con parámetros

$$(n, F(y), F(y + \Delta y) - F(y), 1 - F(y + \Delta y))$$

.

Como $P\{(\alpha - 1)$ de las $X_i's \leq y$; una $X_i$ en $(y, y + \Delta y]$; $(n - \alpha)$ de las $X_i's > y + \Delta y\} = P[Z_1 = \alpha - 1, Z_2 = 1, Z_3 = n - \alpha]$

Entonces la probabilidad del evento buscado es:

$$\frac{n!}{(\alpha - 1)!(n - \alpha)!1!} F(y)^{\alpha - 1} [F(y + \Delta y) - F(y)]^1 [1 - F(y + \Delta y)]^{n - \alpha}$$

Por lo que sustituyendo esta probabilidad en la ecuación (1.8):

$$f_{Y_{\alpha:n}}(y) = \lim_{\Delta y \to 0} \frac{n!}{(\alpha - 1)!(n - \alpha)!} F(y)^{\alpha - 1} \left[\frac{F(y + \Delta y) - F(y)}{\Delta y}\right]^1 [1 - F(y + \Delta y)]^{n - \alpha}$$

$$f_{Y_{\alpha:n}}(y) = \frac{n!}{(\alpha - 1)!(n - \alpha)!} F(y)^{\alpha - 1} [1 - F(y)]^{n - \alpha} \lim_{\Delta y \to 0} \left[\frac{F(y + \Delta y) - F(y)}{\Delta y}\right]^1$$

Por lo tanto:

$$f_{Y_{\alpha:n}}(y) = \frac{n!}{(\alpha - 1)!(n - \alpha)!} F(y)^{\alpha - 1} [1 - F(y)]^{n - \alpha} f(y)$$

Lo cual termina la demostración. □

Nuestro siguiente paso será dar la densidad conjunta de todos los estadísticos de orden la cual vamos a obtener de manera similar al teorema anterior.

**Teorema 1.3 (Densidad conjunta de los estadísticos de orden $Y_{1:n}, Y_{2:n}, \ldots, Y_{\alpha:n}$).** *La función de densidad conjunta de los estadísticos de orden $Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$ definidos a partir de una muestra aleatoria proveniente de una población con función de distribución $F(\cdot)$ y densidad $f(\cdot)$ está dada por:*

$$f_{Y_{1:n},\ldots,Y_{n:n}}(y_1, y_2, \ldots y_n) = \begin{cases} n! f(y_1) f(y_2) \ldots f(y_n) & -\infty < y_1 < y_2 < \ldots < y_n < \infty \\ \\ 0 & e.o.c \end{cases} \tag{1.9}$$

*Demostración.* Sabemos que:

$$f_{Y_{1:n},\ldots,Y_{n:n}}(y_1, y_2, \ldots y_n) = \lim_{\Delta y_i \to 0} \frac{1}{\prod_{i=1}^{n} \Delta y_i} P\left[y_1 < Y_1 \leq y_1 + \Delta y_1; \ldots; y_n < Y_n \leq y_n + \Delta y_n\right]$$

Pero dada la igualdad entre los siguientes eventos:

- $\{y_1 < Y_1 \leq y_1 + \Delta y_1; \ldots; y_n < Y_n \leq y_n + \Delta y_n\}$
- $\{$ un $X_i \in (\, y_1, y_1 + \Delta y_1 \,]; \ldots;$ un $X_i \in (\, y_n, y_n + \Delta y_n \,] \,\}$

Llegamos a que la densidad conjunta es igual al siguiente límite:

$$= \lim_{\Delta y_i \to 0} \frac{1}{\prod_{i=1}^{n} \Delta y_i} P\left[\text{un } X_i \in (\, y_1, y_1 + \Delta y_1 \,]; \ldots; \text{un } X_i \in (\, y_n, y_n + \Delta y_n \,]\right]$$

Con el mismo argumento de una distribución multinomial podemos calcular la probabilidad de la ecuación anterior obteniendo con eso la siguiente igualdad:

$$= \lim_{\Delta y_i \to 0} \frac{n!}{\prod_{i=1}^{n} \Delta y_i} \left[F\left(y_1 + \Delta y_1\right) - F\left(y_1\right)\right] \ldots \left[F\left(y_n + \Delta y_n\right) - F\left(y_n\right)\right]$$

Luego entonces:

$$= n! \lim_{\Delta y_1 \to 0} \frac{1}{\Delta y_1} \left[F\left(y_1 + \Delta y_1\right) - F\left(y_1\right)\right] \ldots \lim_{\Delta y_n \to 0} \frac{1}{\Delta y_n} \left[F\left(y_n + \Delta y_n\right) - F\left(y_n\right)\right]$$

Con lo cual usando la definición de derivada llegamos a:

$$= n! f(y_1) f(y_2) \ldots f(y_n)$$

Lo cual es lo que se quería demostrar. □

El resultado anterior es de gran importancia pues gracias a la densidad conjunta de todos los estadísticos de orden seremos capaces de obtener cualquier densidad conjunta o marginal simplemente integrando respecto a los variables restantes (Para ver ejemplos de este método puede consultar el libro de Naray C. [5] en las páginas 31-43). De esta manera si queremos encontrar la densidad marginal conjunta de los últimos estadísticos de orden $Y_{k+1:n}, Y_{k+2:n}, \ldots, Y_{n:n}$ con $k < n$ lo único que tenemos que hacer es proceder a integrar la densidad dada por (1.9) respecto a $y_1, y_2, \ldots, y_k$.

$$f_{Y_{k+1:n},\ldots,Y_{n:n}}(y_{k+1}, y_{k+2}, \ldots, y_n) = \int_{-\infty}^{y_{k+1}} \int_{-\infty}^{y_k} \cdots \int_{-\infty}^{y_2} n! \prod_{i=1}^{n} f\left(y_i\right) dy_1 d_2 \ldots dy_k$$

Primero resolvemos la integral con respecto a $y_1$

$$\int_{-\infty}^{y_2} n! \prod_{i=1}^{n} f\left(y_i\right) dy_1 = n! \prod_{i=2}^{n} f\left(y_i\right) \int_{-\infty}^{y_2} f\left(y_1\right) = n! \prod_{i=2}^{n} f\left(y_i\right) F\left(y_2\right)$$

Integrando lo anterior ahora con respecto a $y_2$

$$\int_{-\infty}^{y_3} n! \prod_{i=2}^{n} f(y_i) F(y_2) dy_2 = n! \prod_{i=3}^{n} f(y_i) \int_{-\infty}^{y_3} f(y_2) F(y_2) dy_2$$

Para resolver esta última integral hacemos el cambio de variable $u = F(y_2)$ lo que implica que $du = f(y_2) dy_2$ logrando la siguiente simplificación:

$$\int_{-\infty}^{y_3} f(y_2) F(y_2) dy_2 = \int_0^{F(y_3)} udu = \left.\frac{u^2}{2}\right|_0^{F(y_3)} = \frac{F(y_3)^2}{2}$$

Por lo que la integral con respecto a $y_2$ queda de la siguiente manera:

$$n! \prod_{i=3}^{n} f(y_i) \int_{-\infty}^{y_3} f(y_2) F(y_2) dy_2 = \frac{n!}{2} F(y_3)^2 \prod_{i=3}^{n} f(y_i)$$

Continuando con la integración ahora con respecto a $y_3$ y haciendo un cambio de variable parecido al anterior obtenemos:

$$\int_{-\infty}^{y_4} \frac{n!}{2} F(y_3)^2 \prod_{i=3}^{n} f(y_i) dy_3 = \frac{n!}{3!} F(y_i)^3 \prod_{i=4}^{n} f(y_i)$$

Continuamos con esta estrategia hasta llegar a la última integral con respecto a $y_k$ obteniendo lo siguiente:

$$\int_{-\infty}^{y_{k+1}} \int_{-\infty}^{y_k} \cdots \int_{-\infty}^{y_2} n! \prod_{i=1}^{n} f(y_i) dy_1 dy_2 \ldots dy_k = \int_{-\infty}^{y_{k+1}} \frac{n!}{(k-1)!} F(y_k)^{k-1} \prod_{i=k}^{n} f(y_i) dy_k$$

Esta última integral se resuelve de la misma manera llegando al siguiente resultado:

$$f_{Y_{k+1:n},\ldots,Y_{n:n}}(y_{k+1},\ldots y_n) = \begin{cases} \dfrac{n!}{k!} F(y_{k+1})^k \displaystyle\prod_{i=k+1}^{n} f(y_i) & -\infty < y_{k+1} < \ldots < y_n < \infty \\ \\ 0 & \text{e.o.c} \end{cases} \tag{1.10}$$

Si ahora estuviésemos interesados en la densidad conjunta de los primeros k estadísticos de orden $(k < n)$ procederíamos de la misma manera, sólo que ahora la integración sería con respecto a $y_{k+1}, y_{k+2}, \ldots, y_n$ quedando su desarrollo de la siguiente manera:

$$f_{Y_{1:n},\ldots,Y_{k:n}}(y_1, y_2, \ldots, y_k) = \int_{y_k}^{\infty} \int_{y_{k+1}}^{\infty} \cdots \int_{y_{n-1}}^{\infty} n! \prod_{i=1}^{n} f(y_i) dy_n dy_{n-1} \ldots dy_{k+1}$$

$$= n!\prod_{i=1}^{k} f(y_i)\int_{y_k}^{\infty} f(y_{k+1})dy_{k+1}\ldots\int_{y_{n-1}}^{\infty} f(y_n)dy_n$$

$$= n!\prod_{i=1}^{k} f(y_i)\int_{y_k}^{\infty} f(y_{k+1})dy_{k+1}\ldots\int_{y_{n-2}}^{\infty} \left[1-F\left(y_{n-1}\right)\right] f\left(y_{n-1}\right)dy_{n-1}$$

Continuando nuevamente con la misma estrategia del ejemplo anterior llegamos a:

$$f_{Y_{1:n},\ldots,Y_{k:n}}(y_1,\ldots y_k) = \begin{cases} \dfrac{n!}{(n-k)!}\left[1-F\left(y_k\right)\right]^{n-k}\displaystyle\prod_{i=1}^{k} f\left(y_i\right) & -\infty < y_1 < \ldots < y_k < \infty \\ \\ 0 & \text{e.o.c} \end{cases} \tag{1.11}$$

Similarmente se puede verificar, ya sea mediante el procedimiento de integración o por el de igualdad entre eventos, que la densidad conjunta de:

$$Y_{j_1:n}, Y_{j_1+j_2:n}, Y_{j_1+j_2+j_3:n}, \ldots, Y_{j_1+j_2+j_3+\ldots+j_k:n}$$

está dada por:

$$\frac{n!}{(j_1-1)!\ldots(j_k-1)!\,(n-j_1-\cdots-j_k)!} f\left(y_{j_1}\right)\ldots f\left(y_{j_1+\ldots+j_k}\right)\left[F\left(y_{j_1}\right)\right]^{j_1-1}$$

$$\left[F\left(Y_{j_1+j_2}\right)-F\left(y_{j_1}\right)\right]^{j_2-1}\ldots\left[1-F\left(y_{j_1+\ldots+j_k}\right)\right]^{n-j_1-\ldots-j_k} \tag{1.12}$$

Donde el dominio es: $y_{j_1} < y_{j_1+j_2} < \ldots, < y_{j_1+j_2+\ldots+j_k}$ (Está fórmula puede ser consultada en Naray C. [5] página 35)

La ecuación anterior es importante y la más general ya que podemos obtener a partir de esta fórmula la densidad conjunta de cualquier cantidad de estadísticos. Así por ejemplo, si queremos la densidad del $\alpha$-ésimo estadístico de orden lo único que tenemos que hacer es tomar de la ecuación (1.12) $j_1 = \alpha$ y $k = 1$ para así ver que llegamos a la misma ecuación que habíamos encontrado en la expresión (1.7).

Como ejemplo de lo general de la fórmula anterior veremos que se puede obtener fácilmente la densidad conjunta de $Y_{\alpha:n}$ y $Y_{\beta:n}$ mediante unas pequeñas sustituciones en la ecuación (1.12). Para ello, primero supondremos que $\alpha < \beta$ y haremos $j_1 = \alpha$, $j_2 = \beta - \alpha$ y $k = 2$ para así obtener que:

$$f_{Y_{\alpha:n},Y_{\beta:n}}(y_\alpha, y_\beta) = \frac{n!F(y_\alpha)^{\alpha-1}\left[F(y_\beta)-F(y_\alpha)\right]^{\beta-\alpha-1}\left[1-F(y_\beta)\right]^{n-\beta} f(y_\alpha)f(y_\beta)}{(\alpha-1)!(\beta-\alpha-1)!(n-\beta)!} \tag{1.13}$$

Donde el dominio es $-\infty < y_{\alpha:n} < y_{\beta:n} < \infty$

Por lo que si hacemos $\alpha = 1$ y $\beta = n$ obtendríamos la densidad conjunta del mínimo y el máximo la cual es necesaria al momento de querer obtener la densidad del rango muestral.

$$f_{Y_{1:n},Y_{n:n}}(y_1, y_n) = n(n-1)[F(y_n - F_{y_1})]^{n-2} f(y_1) f(y_n) \tag{1.14}$$

## 1.4. Ejemplos

**Ejemplo 1.2.** Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de un fenómeno aleatorio especificado por la función de densidad dada por:

$$f_X(x) = e^{-x} I_{\{0<x<\infty\}}(x)$$

Encontrar la función de densidad para $Y_{1:n}, Y_{n:n}$ y para el rango muestral $(Y_{n:n} - Y_{1:n})$

*Solución.* Primero obtenemos la función de distribución integrando la respectiva función de densidad

$$F(x) = \int_0^x f_X(y)dy = \int_0^x e^{-y}dy = 1 - e^{-x}$$

Luego por el teorema (1.2)

$$f_{Y_{\alpha:n}}(y_\alpha) = \frac{n!}{(\alpha-1)!\,(n-\alpha)!} \left[F(y_\alpha)\right]^{\alpha-1} \left[1 - F(y_\alpha)\right]^{n-\alpha} f(y_\alpha)$$

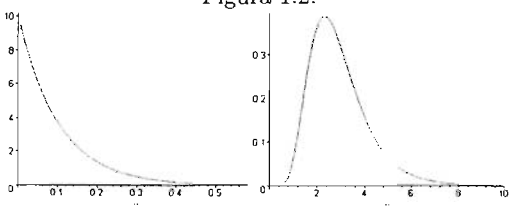
Por lo que haciendo $\alpha = 1$ y $\alpha = n$ obtenemos que:

$$f_{Y_{1:n}}(y_1) = ne^{-ny_1} I_{\{0<y_1<\infty\}}(y_1)$$

$$f_{Y_{n:n}}(y_n) = n\left(1 - e^{-y_n}\right)^{n-1} e^{-y_n} I_{\{0<y_n<\infty\}}(y_n)$$

Las gráficas de estas dos densidades se pueden ver a continuación para el caso en que n=10:

Figura 1.2:

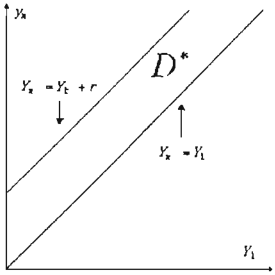


Para obtener la densidad del rango primero encontramos su distribución:

$$F_R(r) = P[Y_{n:n} - Y_{1:n} \leq r] = P[Y_{n:n} \leq Y_{1:n} + r] = P[(Y_{n:n}, Y_{1:n}) \in D^*)]$$

Donde $D^*$ es una región en $\mathbb{R}^2$ delimitada por las rectas $y_n = y_1$ y $y_n = y_1 + r$ en el primer cuadrante del plano cartesiano (Ver figura 1.3).

Figura 1.3:




Por lo tanto tenemos:

$$F_R(r) = \int\int_{D^*} f_{Y_{1:n}Y_{n:n}}(y_1, y_n) dy_n dy_1$$

Sustituyendo la densidad conjunta del mínimo y del máximo dado en ( 1.14):

$$F_R(r) = \int_0^{\infty} \int_{y_1}^{y_1+r} n(n-1)\left\{e^{-y_1} - e^{-y_n}\right\}^{n-2} e^{-y_1} e^{-y_n} dy_n dy_{1n}$$

Integrando por el método de sustitución:

$$F_R(r) = (1 - e^{-r})^{n-1}$$

Derivando obtenemos que la densidad del rango es:

$$f_R(r) = \begin{cases} (n-1)\left(1 - e^{-r}\right)^{n-2} e^{-r} & r \geq 0 \\ 0 & \text{e.o.c} \end{cases}$$

**Ejemplo 1.3.** Sea $Y_{1:5}, Y_{2:5}, \ldots, Y_{5:5}$ los estadísticos de orden de una muestra aleatoria de tamaño 5 de un fenómeno aleatorio especificado por la función de densidad dada por:

$$f_X(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & \text{e.o.c} \end{cases}$$

Probaremos que $Y_{2:5}/Y_{4:5}$ y $Y_{4:5}$ son variables aleatorias independientes.

*Solución.* Obtenemos la función de distribución integrando la densidad

$$F_X(x) = \int_0^x f_X(y)dy = \int_0^x 3y^2 dy = x^3 \qquad 0 < x < 1$$

Encontramos la densidad conjunta de $Y_{2:5}$ y $Y_{4:5}$ haciendo $n = 5$, $\alpha = 2$ y $\beta = 4$ en la ecuación (1.13)

$$f_{Y_{2:5},Y_{4:5}}(y_2, y_4) = \begin{cases} 1080 y_2^5 y_4^2 [y_4^3 - y_2^3][1 - y_4^3] & 0 < y_2 < y_4 < 1 \\ 0 & \text{e.o.c} \end{cases}$$

Definimos $U$ y $V$ como variables aleatorias dadas por:

$$U = \frac{Y_{2:5}}{Y_{4:5}} \qquad\qquad V = Y_{4:5}$$

Tenemos que demostrar que U y V son independientes.

Encontramos la densidad conjunta de U y V mediante el teorema de cambio de variables [3]

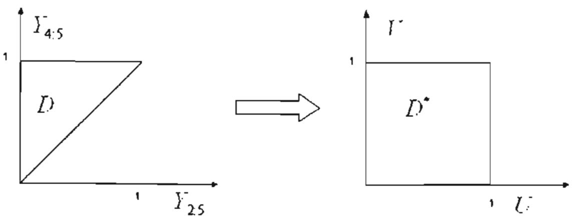
Las funciones inversas y el respectivo jacobiano son:

$$Y_{2:5} = W_1(U,V) = UV \qquad\qquad Y_{4:5} = W_2(U,V) = V$$

$$|J| = \begin{vmatrix} \frac{\partial}{\partial U} W_1(U,V) & \frac{\partial}{\partial V} W_1(U,V) \\ \frac{\partial}{\partial U} W_2(U,V) & \frac{\partial}{\partial V} W_2(U,V) \end{vmatrix} = \begin{vmatrix} V & U \\ 0 & 1 \end{vmatrix} = V$$

La transformación del dominio puede verse en la figura ( 1.4).

Figura 1.4:




[3] El teorema de cambio de variable puede ser consultado en el libro "*Introduction to Mathematical Statistics*" [2] en las páginas 132-147 en donde además podrá encontrar una serie de ejemplos.

Entonces bajo la transformación anterior la densidad conjunta de U y V es:

$$f_{U,V}(u,v) = f_{Y_{2:5},Y_{4:5}}(W_1(u,v), W_2(u,v))\,|J| = 1080(uv)^5v^3[v^3-(uv)^3][1-v^3] \qquad (u,v)\in D^*$$

$$f_{U,V}(u,v) = 1080(uv)^5v^6[1-u^3][1-v^3] = \underbrace{1080u^5[1-u^3]}_{g(u)}\cdot\underbrace{v^6[1-v^3]}_{h(v)} \qquad (u,v)\in D^*$$

Por lo tanto:

$$f_{U,V}(u,v) = \begin{cases} g(u)\cdot h(v) & (u,v)\in D^* \\ 0 & \text{e.o.c} \end{cases}$$

Finalmente, como $f_{U,V}(u,v)$ se expresa como el producto de una función que sólo depende de $u$ por otra que sólo depende de $v$ aunado con el hecho de que la región $D^*$ es tal que el dominio de la función no depende de variables sino que únicamente de constantes, entonces se prueba que en efecto $U$ y $V$ son variables aleatorias independientes.[4]

**Ejemplo 1.4.** Encuentre la probabilidad de que el rango muestral de una muestra aleatoria de tamaño 5 proveniente de una distribución uniforme $(0,1)$ sea mayor a $1/2$

*Solución.*

$$f_X(x) = 1 \qquad 0<x<1 \Rightarrow F_X(x) = \int_0^x 1dx = x \qquad 0<x<1$$

Procedemos a encontrar la función de distribución del rango, para ello nuevamente utilizaremos la densidad conjunta del mínimo y el máximo la cual encontramos sustituyendo en la ecuación (1.14)

$$f_{Y_{1:5},Y_{5:5}}(y_1,y_5) = 20[y_5-y_1]^3 \qquad 0<y_1<y_5<1$$

Entonces

$$F_R(r) = P[Y_{5:5}-Y_{1:5}\le r] = P[Y_{5:5}\le Y_{1:5}+r] = \int\int_{D^*} f_{Y_{1:5},Y_{5:5}}(y_1,y_5)dy_1dy_5$$

Donde $D^*$ es la región dibujada en la figura (1.5):
Entonces

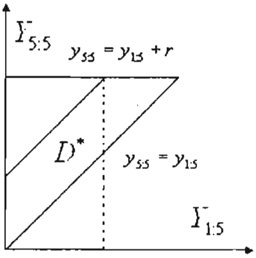
$$F_R(r) = \int_0^{1-r}\int_{y_1}^{r+y_1} 20(y_5-y_1)^3dy_5dy_1 + \int_{1-r}^{1}\int_{y_1}^{1} 20(y_5-y_1)^3dy_5dy_{1:1} = r^5$$

Por lo tanto

$$F_R(r) = r^5 \qquad 0<r<1 \Rightarrow f_R(r) = 5r^4 \qquad 0<r<1$$

[4] En esta parte hacemos uso de un teorema que nos garantiza que $X$ y $Y$ son independientes sí y sólo si la densidad puede ser expresada como el producto de una función que sólo dependa de $X$ por otra que sólo dependa de $Y$, ambas definidas en un dominio con extremos constantes. Este teorema así como algunos ejemplos pueden ser consultados en el libro de Hogg [2] en su página 81.

Figura 1.5:




Entonces la probabilidad buscada es:

$$P[R > \frac{1}{2}] = 1 - F_R\left(\frac{1}{2}\right) = 1 - \left(\frac{1}{2}\right)^5 = \frac{31}{32} = 0{,}96875$$

Este último resultado nos dice que al obtener una muestra aleatoria de tamaño 5 proveniente de una distribución uniforme (0,1), entonces la diferencia del máximo menos el mínimo será mayor a $\frac{1}{2}$ con probabilidad 0.96875.

# Capítulo 2

# Esperanza Condicional

En este segundo capítulo se aborda el tema de Esperanza Condicional, el cual es el segundo paso para poder llevar a cabo la caracterización de funciones de distribución que pretendemos en este trabajo . Para comenzar a motivar la definición de esperanza condicional daremos antes una breve introducción a herramientas necesarias para poder definir este concepto. Así pues primeramente se verá el concepto de probabilidad condicional para posteriormente motivar la definición de función de densidad condicional discreta o continua de una variable aleatoria $X$ dado otra variable aleatoria $Y$. Posteriormente, se define el concepto de esperanza (no condicional) para que una vez comprendido éste, se motive de una manera lógica el concepto principal de este capítulo.

## 2.1. Probabilidad Condicional

Dentro de la introducción al cálculo de probabilidades surge el concepto de **probabilidad condicional** el cual aparece para dar una solución al problema de medir la probabilidad de una evento A dado la ocurrencia de un evento B. Un ejemplo típico de esta situación se observa cuando nosotros llevamos a cabo un muestreo sin reemplazo de una cierta población. Por ejemplo, supongamos que tenemos una urna que tiene 8 bolas rojas y una bola negra, de la cual procedemos a obtener una muestra por medio de un muestreo sin reemplazo, esto es iremos sacando una bola tras otra sin regresarla a la urna, una buena pregunta por lo cual estaríamos interesados es saber cuál es la probabilidad de que se obtenga la bola negra dado que se han obtenido 2 bolas rojas.

En estos casos tenemos el problema de calcular la probabilidad de eventos relativos a cada parte del experimento condicionadas a lo que haya ocurrido en las primeras partes de éste. A las probabilidades de este tipo se les conoce como **probabilidades condicionales** y se les denota como $P(A|B)$ ('La probabilidad de que ocurra A dado que ocurrió B").

Una probabilidad condicional tiene un sentido muy claro en los casos en que se trata de la probabilidad de un evento que depende de una segunda parte de un experimento condicionada a lo que ocurra en la primera parte de éste. Es natural preguntar por la probabilidad de que obtengamos una bola negra dado que han salido 2 rojas, sin embargo, al dar una definición general de probabilidad condicional sería entonces necesario dar una medida de probabilidad para la ocurrencia de que se obtengan dos bolas rojas al principio dado que la tercer bola fue negra, lo cual quizás no suena muy lógico para asignarle una medida de probabilidad. En el caso general de un experimento aleatorio es posible definir el concepto de probabilidad condicional de un evento $A$ dada la ocurrencia de un evento $B$ o bien la de $B$ dado $A$.

Decir que un evento $B$ ocurre equivale a decir que ocurre alguno de los posibles resultados que lo componen, en otras palabras, los posibles resultados del experimento aleatorio al momento de condicionar no son ya todos los elementos del espacio muestral sino que únicamente aquellos que conforman el evento $B$, esto quiere decir que al momento de condicionar un evento lo que hacemos es reducir el número de resultados posibles a todos aquellos que estén contenidos dentro del evento al cual se está condicionando los resultados.

Tenemos el siguiente problema: supongamos que restringimos a $\Omega$ a la ocurrencia de un evento $B$ y que queremos medir la probabilidad de un evento $A$. ¿Qué relación existe entre la probabilidad de $A$ calculada con relación a todo el espacio muestral con la probabilidad del evento $A$ con relación al espacio muestral reducido?. Para tratar de resolver esta pregunta notemos que obviamente estas dos probabilidades no pueden ser iguales ya que por ejemplo, dado que ocurre $B$, la ocurrencia de $B$ se vuelve segura por lo que la probabilidad de $B$ dado $B$ debe ser definida como 1, aquello que no necesariamente ocurriría cuando se tome la medida de probabilidad respecto a todo $\Omega$ del evento $B$. Otro aspecto importante es que si el evento $A$ no intercecta al evento $B$ ($B \cap A = \phi$), entonces la definición de probabilidad condicional deber de ser tal que asigne 0 a este evento. Así pues, dado que condicionamos al evento $B$, es claro que para que el evento $A$ ocurra es necesario que ocurra el evento $B \cap A$, lo que implica que se debe tener que $P(A|B) = P(B \cap A|B)$, de manera que el problema se reduce únicamente a calcular la probabilidad de eventos que son subconjuntos del espacio restringido.

De acuerdo al significado que queremos darle a la probabilidad de un evento como una medida de que tan fácilmente ocurre éste al realizar el experimento aleatorio, podemos asumir que las probabilidades en el espacio muestral restringido deben asignarse de tal manera que entre ellas guarden la misma proporción que la que guardaban en el espacio muestral original. Lo anterior se escribiría así:

Sea $B$ un evento y $B_1, B_2 \subseteq B$ cualquiera, entonces la definición de probabilidad condicional debe ser tal que cumpla con lo siguiente:

$$\frac{P(B_1)}{P(B_2)} = \frac{P(B_1|B)}{P(B_2|B)}$$

$$\Rightarrow \frac{P(B_1)}{P(B_1|\, B)} = \frac{P(B_2)}{P(B_2|\, B)} = c$$

Como lo anterior es válido para todo $B_1$ y $B_2$ contenidos en $B$, entonces si tomamos $B_2 = B$, de donde por la última igualdad concluimos que $c = P(B)$, haciendo $A = B_1$ y despejando $P(A|\, B)$ obtendríamos:

$$P(A|\, B) = \frac{P(A)}{P(B)} \qquad \forall A \subseteq B$$

Pero como además pedimos que:

$$P(A|\, B) = P(A \cap B|\, B)$$

Y como $A \cap B \subseteq B$ entonces:

$$P(A|\, B) = \frac{P(A \cap B)}{P(B)}$$

Lo cual motiva la siguiente definición de probabilidad condicional.

**Definición 2.1 (Probabilidad Condicional).** Sean $A$ y $B$ eventos y supongamos que la probabilidad de $B$ es positiva. Se define la probabilidad condicional de $A$ dada la ocurrencia de $B$, denotada por $P(A|\, B)$, mediante la fórmula:

$$P(A|\, B) = \frac{P(A \cap B)}{P(B)} \tag{2.1}$$

Es fácil ver que la probabilidad condicional tiene las mismas propiedades que cualquier función de probabilidad pues nuevamente vuelve a ser una medida donde el espacio muestral está siendo restringido por la ocurrencia de un evento.

De la definición dada, se desprenden nuevos conceptos y reglas como la regla del producto, el concepto de independencia estocástica y la regla de Bayes las cuales se mencionan a continuación.

### 2.1.1. Regla del Producto

De la definición de probabilidad condicional se tiene que:

$$P(A \cap B) = P(A|\, B)P(B) \tag{2.2}$$

Relación que es conocida como la regla del producto aplicado a dos eventos. Note que dicha regla es capaz ahora de calcular la probabilidad de intersección de dos eventos en aquellos casos en donde es más fácil encontrar la probabilidad condicional de manera directa que buscar la probabilidad de la intersección de los eventos involucrados.

La regla del producto puede fácilmente generalizarse al caso de $n$ eventos, obteniendo con ello lo que se conoce como la *regla general del producto* la cual en algunos casos se vuelve una herramienta muy útil para el cálculo de probabilidades.

**Proposición 2.1 (Regla del producto gerneralizada).** *Sean $A_1, A_2, A_3, \ldots, A_n$ eventos tales que $P(A_1 \cap A_2 \cap \ldots \cap A_n) > 0$, entonces:*

$$P(\cap_{i=1}^{n} A_i) = P(A_n | A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \ldots P(A_2 | A_1) P(A_1)$$

*Demostración.* Por la regla del producto para dos eventos dada en ( 2.2):

$$P(\cap_{i=1}^{n} A_i) = P(A_n | \cap_{i=1}^{n-1} A_i) P(\cap_{i=1}^{n-1} A_i)$$

Donde los eventos $A$ y $B$ son definidos como $A = A_n$ y $B = \cap_{i=1}^{n-1} A_i$ respectivamente.

De lo anterior volvemos aplicar la regla del producto para dos eventos obteniendo

$$P(\cap_{i=1}^{n} A_i) = P(A_n | \cap_{i=1}^{n-1} A_i) P(A_{n-1} | \cap_{i=1}^{n-2} A_i) P(\cap_{i=1}^{n-2} A_i)$$

Donde ahora los eventos $A$ y $B$ son definidos como $A = A_{n-1}$ y $B = \cap_{i=1}^{n-2} A_i$ respectivamente.

Continuando $n - 3$ veces más obtenemos:

$$P(\cap_{i=1}^{n} A_i) = P(A_n | A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \ldots P(A_2 | A_1) P(A_1)$$

Lo cual es lo que se quería probar □

### 2.1.2. Independencia estocástica

La independencia estocástica es un concepto que se define a partir de la idea de probabilidad condicional de la siguiente manera:

Sabemos que la $P(A|B)$ es una probabilidad que se altera en comparación a $P(A)$ pues la medida está siendo hecha sobre un espacio muestral restringido. Cuando hay diferencia entre estas probabilidades se puede interpretar que el evento $A$ sí está dependiendo del evento $B$ pues se sabe que una vez que ocurrió $B$ la probabilidad de ocurrencia de $A$ se ve modificada. Siguiendo esta lógica, en los casos en que la ocurrencia de un evento $B$ no altere la probabilidad de ocurrencia de un evento $A$ se puede hablar de independencia de la probabilidad de ocurrencia de $A$ dada la ocurrencia de $B$. Es decir si $A$ y $B$ son independientes se tiene que cumplir que:

$$P(A|B) = P(A)$$

Lo cual al ser sustituido en la regla del producto para dos eventos:

$$P(A \cap B) = P(A|B)P(B) = P(A)P(B)$$

Lo cual motiva a definir la independencia estocástica de dos eventos.

**Definición 2.2.** Se dice que dos eventos $A$ y $B$ son estocásticamente independientes si:

$$P(A \cap B) = P(A)P(B)$$

Note que para construir la definición de independencia estocástica se utilizó la interpretación de la probabilidad condicional, no obstante dentro de la definición formal de independencia no aparece ninguna probabilidad condicional involucrada. Lo anterior es hecho así con el propósito de tener una definición que no imponga algún tipo de restricción a los eventos, problema que sucede con la probabilidad condicional en donde se pide que el evento al cual se esté condicionado tenga probabilidad positiva.

### 2.1.3. Regla de Bayes

Al motivar la definición de probabilidad condicional se utilizó el problema de tratar de calcular la probabilidad de un evento $A$ que ocurre en un segundo experimento dado la ocurrencia de un evento B en un primer experimento. Sin embargo, la definición de probabilidad condicional es general y, en la misma situación descrita, tiene sentido calcular la probabilidad de ocurrencia del evento $B$ dado la ocurrencia del evento $A$.

Así pues, la definición de probabilidad condicional nos permite calcular tanto $P(A|B)$ como $P(B|A)$ obteniendo con ello la siguiente relación

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

Esta ecuación conocida como la **Regla de Bayes** es muy útil en la resolución de problemas en donde se facilita el cálculo de una probabilidad condicional ya sea la de $A$ dado $B$ o la de $B$ dado $A$ para encontrar la opuesta.

## 2.2. Distribuciones Condicionales

Ahora abordaremos el tema de distribuciones condicionales. Para ello primero lo definiremos para variables aleatorias del tipo discreto a partir de la probabilidad condicional de dos eventos.

Sea $X_1$ y $X_2$ dos variables aleatorias del tipo discreto en el espacio de probabilidad $(\Omega, \mathcal{A}, P)$. Queremos motivar la definición de la densidad de probabilidad de $X_2 | X_1 = x_1'$, para ello sea $p_{X_1,X_2}(x_1, x_2)$ su respectiva función de densidad de probabilidad conjunta la cual suponemos será positiva en un subconjunto $\mathcal{D} \in \mathbb{R}^2$ y cero en cualquier otro caso. Denotaremos por $p_{X_1}(x_1)$ y $p_{X_2}(x_2)$ sus respectivas funciones marginales de densidad de probabilidad de $X_1$ y $X_2$.

Definimos a los siguientes eventos:

$$A_1 = \left\{ \omega \in \Omega : X_1(\omega) = x_1' \right\} = X^{-1}(x_1')$$

$$A_2 = \left\{ \omega \in \Omega : X_2(\omega) = x_2' \right\} = X^{-1}(x_2')$$

Donde $x_1'$ debe ser un punto tal que $P(A_1) = P[X_1 = x_1] = p_{X_1}(x_1') > 0$.

Entonces tenemos que por la definición de probabilidad condicional:

$$P(A_2|\, A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{P\left[X_1 = x_1', X_2 = x_2'\right]}{P\left[X_1 = x_1'\right]} = \frac{p_{X_1,X_2}(x_1', x_2')}{p_{X_1}(x_1')}$$

Lo anterior nos está diciendo que la probabilidad de que $X_2 = x_2'$ dado que $X_1 = x_1'$ está dado por el cociente de la densidad de probabilidad conjunta dividida entre la densidad de probabilidad marginal de la variable aleatoria que está condicionando. Como este cociente está proporcionando la probabilidad condicional, entonces se le asigna el símbolo $p_{X_2|X_1}(x_2'|\, x_1')$ lo cual motiva la siguiente definición.

**Definición 2.3 (Densidad de probabilidad condicional).** Sea X y Y dos variables aleatorias del tipo discreto con $p_{X,Y}(x,y)$, $p_X(x)$ y $p_Y(y)$ sus respectivas funciones de densidad de probabilidad conjunta y marginal respectivamente. Sea $y_1 \in \mathbb{R}$ fijo tal que $p_Y(y_1) > 0$. Se define la densidad de probabilidad condicional de $X$ dado $Y = y_1$ denotada por $p_{X|Y}(x|\, y_1)$ como:

$$p_{X|Y}(x|\, y_1) = P\left[X = x|\, Y = y_1\right] = \frac{p_{X,Y}(x, y_1)}{p_Y(y_1)} \tag{2.3}$$

Note que una vez fijo $y_1$, la densidad condicional discreta cumple con ser siempre positiva o igual a cero al ser cociente de dos funciones de densidad de probabilidad, además es fácil probar que cumple con el hecho de ser función de densidad de probabilidad ya que como la densidad marginal de Y se obtiene sumando la conjunta respecto a todos los posibles valores de la variable aleatoria X se tiene que:

$$\sum_x p_{X|Y}(x|\, y_1) = \frac{1}{p_Y(y_1)} \sum_x p_{X,Y}(x, y_1) = \frac{1}{p_Y(y_1)} p_Y(y_1) = 1$$

Lo cual prueba que en efecto la densidad de probabilidad condicional así definida cumple con los requisitos para ser llamada "*densidad de probabilidad*".

Extenderemos la definición anterior ahora para el caso continuo.

**Definición 2.4 (Densidad condicional).** Sea X y Y variables aleatorias del tipo absolutamente continuo y sean $f_{X,Y}(x,y)$ su respectiva función de densidad conjunta. Llamemos $f_X(x)$ y $f_Y(y)$ las densidades marginales. Usaremos el resultado del caso discreto para motivar la definición de la función de densidad condicional de una variable aleatoria absolutamente continua. Sea $y_1 \in \mathbb{R}$ fijo y tal que $f_Y(y_1) > 0$. Se define la densidad condicional de $X$ dado $Y = y_1$ como:

$$f_{X|Y}(x|\, y_1) = \frac{f_{X,Y}(x, y_1)}{f_Y(y_1)} \tag{2.4}$$

Nuevamente es fácil probar que la definición anterior cumple con los requisitos para ser llamada densidad pues para cada $y_1$ fijo y tal que $f_Y(y_1) > 0$ se tiene que:

$$\int_{-\infty}^{\infty} f_{X|Y}(x|\, y_1)dx = \int_{-\infty}^{\infty} \frac{f_{X,Y}(x, y_1)}{f_Y(y_1)}dx = \frac{1}{f_Y(y_1)} \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x, y_1)dx}_{f_Y(y_1)} = 1$$

Además es claro que $f_{X|Y}(x|\, y_1) \geq 0$ por ser un cociente de funciones de densidad.

Observe que cuando $X$ y $Y$ son discretas, entonces la variable aleatoria. $X|\, Y = y_1$ vuelve a ser una variable aleatoria discreta con función de densidad de probabilidad dada por $p_{X|Y}(x|\, y_1)$ mientras que cuando $X$ y $Y$ son variables aleatorias absolutamente continuas, entonces $X|\, Y = y_1$ es una variable aleatoria absolutamente continua con función de densidad dada por $f_{X|Y}(x|\, y_1)$. De donde se calculan las probabilidades fácilmente ya sea sumando o integrando las respectivas funciones para el caso discreto o absolutamente.

Teniendo bien definido lo anterior y recordando hecho que $F(x) = \sum_{x_k \leq x} f_X(x_k)$, entonces resulta natural definir la función de distribución condicional para el caso discreto como:

$$F_{X|Y}(x) = \sum_{x_k \leq x} p_{X|Y}(x_k|\, y_1)$$

Mientras que para el caso absolutamente continuo:

$$F_{X|Y}(x) = \int_{-\infty}^{x} f_{X|Y}(u|\, y_1)du$$

Note que en ambos casos $F_{X|Y}(x) = P\,[X \leq x|\, Y = y_1]$.

## 2.3. El concepto de Esperanza

Sin duda el concepto de esperanza de una v. a. es básico en la teoría de probabilidades, pues su importancia es comparable incluso con el de la probabilidad de un evento. Quizás al principio se desechó esta idea de relación pero fue la historia misma la que probaría que realmente había una estrecha relación entre ambos conceptos ya que una vez que se da la probabilidad como disciplina matemática para su posterior fusión con la teoría de la medida a principios del siglo XX, es cuando se ve el concepto de probabilidad como una medida y la esperanza como una integral de *Lebesgue* respecto a esta misma medida¨(Puede encontrar mucha infomación relacionada con la historia del cálculo de probabilidades en las notas de Miguel A. García Álvarez [4]).

La definición de esperanza hoy en día puede hacerse con el enfoque de teoría de la medida o bien con el enfoque clásico que involucra el promedio ponderado de valores que toma una variable aleatoria. El presente trabajo no pretende adentrarse al modelo probabilístico visto como un espacio de medida por lo que la definición será dada como sigue:

**Definición 2.5 (Esperanza de variables aleatorias discretas).** Sea X una variable aleatoria discreta con función de densidad de probabilidad dada por $p_X(x)$. Supongamos que $x_1, x_2, x_3 \ldots$ son los posibles valores que puede tomar. Se dice que la variable aleatoria $X$ tiene esperanza finita si la serie $\sum_{x_k} |x_k| \, p_X(x_k)$ es convergente y, en ese caso, se define la esperanza de $X$, denotada por $E[X]$, mediante la ecuación

$$E[X] = \sum_{x_k} x_k p_X(x_k) \tag{2.5}$$

Observe que cuando tomamos a $X$ como la función indicadora de un evento $A \in \Omega$:

$$X(\omega) = I_A(\omega) = \begin{cases} 1 & \text{si } \omega \in \text{A} \\ \\ 0 & \text{e.o.c} \end{cases}$$

Entonces:

$$E[X] = 1 * P(A) + 0 * P(A^c) = P(A)$$

Lo que muestra de una manera más directa el por qué el concepto de esperanza y probabilidad de un evento están siempre muy ligados.

**Definición 2.6 (Esperanza de variables aleatorias absolutamente continuas).** Sea X una variable aleatoria absolutamente continua con función de densidad dada por $f_X(x)$ . Se dice que X tiene esperanza finita si la integral $\int_{-\infty}^{\infty} |x| \, f_X(x)dx$ es finita y, en ese caso, se define la esperanza de $X$, denotada por $E[X]$, mediante la ecuación

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{2.6}$$

Observe que se tuvo que hacer una definición para variables aleatorias discretas y otra para variables aleatorias absolutamente continuas, ésto no implica que la esperanza sólo pueda definirse para estos dos tipos de variables aleatorias, sino que puede darse otra definición que involucra más herramienta matemática y la cual es capaz de extender el concepto a cualquier variable aleatoria ya sea continua, absolutamente continua, discreta o mixta. Lo anterior se logra introduciendo el concepto de la integral de Riemann-Stieltjes la cual hace la integral ya no sobre una variable $x$ sino sobre una función. [1]

La definición general de Esperanza que daremos en este trabajo es la siguiente:

**Definición 2.7 (Definición general de Esperanza).** Sea $X$ una variable aleatoria con función de distribución dada por $F_X(x)$. Se dice que $X$ tiene esperanza finita si la integral de Riemann-Stieltjes $\int_{-\infty}^{\infty} |x|\, dF_X$ es finita y, en ese caso, se define la esperanza de $X$, denotada por $E[X]$ como:

$$E[X] = \int_{-\infty}^{\infty} x dF_x \tag{2.7}$$

Es claro que lo anterior es una extensión de las dos definiciones hechas en (2.5) y (2.6) puesto que:

$$E[X] = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x dF_x = \int_{-\infty}^{\infty} x f_X(x) dx & \text{cuando } X \text{ es v.a. absolutamente continua} \\ \\ \displaystyle\int_{-\infty}^{\infty} x dF_x = \sum_{x_k} x_k p_X(x_k) & \text{cuando } X \text{ es v.a. discreta.} \end{cases}$$

Además, esta nueva definición nos dice que cuando X es una v.a. mixta entonces la esperanza de X se calcula sumando en las partes en que $F_X(x)$ tiene discontinuidades e integrando donde $F_X(x)$ es continua.

### 2.3.1. Algunas ideas erróneas

Eventualmente han surgido algunas ideas erróneas con relación al concepto de esperanza. Dos de ellas provienen de que a la esperanza de una variable aleatoria $X$ se le llama también su valor esperado, termino que al decirlo sugiere las dos primeras ideas:

1. La esperanza de una variable aleatoria $X$ coincide con alguno de sus posibles valores.

2. El valor que se espera al realizar el experimento aleatorio es $E[X]$

[1] El tema de la integral de Riemann-Stieltjes es visto en los cursos de Cálculo Integral por lo que se dará por hecho todas las propiedades que tiene dicha integral. Para mayor información puede consultar estas propiedades en las obras de G. Bartle [8] página 212 o la de Rudin [9] pagina 120

Para demostrar que la primer idea es errónea basta considerar el caso de una variable aleatoria $X$ con distribución Binomial de parámetros $(n, p)$. Entonces según nuestra definición de esperanza dado en (2.5) tenemos que:

$$E[X] = \sum_{x=0}^{n} x p_X(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = np$$

Por lo que si consideramos el caso particular en que $X$ se distribuye como una Binomial con parámetros $n = 5$ y $p = \frac{1}{2}$ obtendríamos que $E[X]$=2.5 lo cual mostraría que en efecto la esperanza no coincide con algún valor posible de $X$ ya que como sabemos una variable aleatoria Binomial sólo toma valores enteros.

La segunda de estas ideas erróneas se desecha de igual manera con el ejemplo anterior, pues claro es que si $X(\omega) \neq 2.5 \quad \forall \omega \in \Omega$ entonces al realizar el experimento nunca se esperaría que $X$=2.5, es decir no se puede esperar que la variable aleatoria coincida con un valor que simplemente no puede tomar.

Sin embargo, detrás de esta idea puede haber algo de verdad si precisamos bien qué entendemos con la frase "el valor que se espera obtener". Para aclarar esto consideremos una variable aleatoria del tipo discreta $X$ Bernoulli de parámetro $p = \frac{1}{2}$, la cual sabemos sólo toma un conjunto finito de valores, a saber, 1 con probabilidad $p$ y 0 con probabilidad $1-p$. Supongamos que el experimento aleatorio que modela $X$ es el lanzamiento de una moneda justa haciendo que $X = 1$ cuando obtenemos águila y $X = 0$ cuando obtenemos sol. De este medo repetimos $n$ veces el experimento aleatorio es decir, aventamos $n$ veces la moneda, obteniéndose $x_k$ como valor de $X$ en la realización número $k$. Entonces el promedio de los valores que ha obtenido $X$ (Lo cual denotamos por: $\bar{x}$) en los $n$ experimentos estará dado por:

$$\bar{x} = \frac{x_1 + \ldots + x_n}{n}$$

Por otra parte, es obvio que cada que se realiza el experimento, $X$ sólo tomará uno de sus posibles valores 0 ó 1 respectivamente. Definimos $n_0$ como el número de ocasiones en que la variable aleatoria tomó el valor 0 y $n_1$ como el número de ocasiones en que la variable aleatoria tomó el valor 1 (Es claro que $n_1 + n_2 = n$). Se tiene entonces que el promedio de valores que tomó $X$ en las $n$ experimentos se puede expresar como:

$$\bar{x} = \frac{x_1 + \ldots + x_n}{n} = \frac{n_0 0 + n_1 1}{n} = 0\frac{n_0}{n} + 1\frac{n_1}{n}$$

Ahora note que los cocientes dados por $\frac{n_0}{n}$ y $\frac{n_1}{n}$ es la frecuencia relativa con la que $X$ toma el valor 0 ó 1 respectivamente, de manera que si tomamos la interpretación frecuentista de la probabilidad, se esperaría que, cuando $n$ es grande el cociente $\frac{n_0}{n}$ ha de estar próximo a la probabilidad de obtener sol ($P[X=0]$), mientras que $\frac{n_1}{n}$ próximo a la probabilidad de obtener águila ($P[X=1]$). Obteniéndose así que:

$$\bar{x} \approx 0P[X=0] + 1P[X=1] \underbrace{:=}_{def.} E[X] = \frac{1}{2}$$

Es decir la esperanza nos estaría diciendo que si lanzamos muchas veces la moneda, entonces el "*promedio*" de los valores tomados por la variable aleatoria $X$ será $\frac{1}{2}$, es decir se esperaría que a la larga se obtuviera el mismo número de águilas que de soles.

De esta manera podemos enunciar la siguiente conclusión:

**La esperanza de $X$ representa el promedio de los valores que toma la variable aleatoria cuando el experimento aleatorio es repetido en varias ocasiones**

Queda ahora claro que la esperanza de $X$ no es el valor que se espera obtener para $X$, pero sí el valor que se espera obtener en promedio. Obsérvese además que, independientemente de que se tome el punto de vista frecuentista, la esperanza de una variable aleatoria discreta $X$, de acuerdo a su definición, es un **promedio ponderado** de sus diferentes posibles valores, de tal manera que al valor $x_k$ se le asigna el peso $p_X(x_k)$.

Esta noción de promedio ponderado también puede ser llevada al caso en que $X$ es absolutamente continua pues recordando que la integral puede ser vista como una suma de Riemann se tendría que[2] :

$$E[X] = \int_a^b x f_X(x)dx \approx \sum_{k=1}^{n} x_k f_X(x_k)(x_k - x_{k-1})$$

Lo cual si consideramos que $f_X(x_k)$ es constante en el intervalo $(x_{k-1}, x_k)$ :

$$f_X(x_k)(x_k - x_{k-1}) = F_X(x_k) - F_X(x_{k-1}) = P\left[X \in (x_{k-1}, x_k]\right]$$

Por lo que sustituyendo se tendría que:

$$E[X] \approx \sum_{k=1}^{n} x_k P\left[X \in (x_k, x_{k-1}]\right]$$

De donde nuevamente la esperanza resulta ser un promedio ponderado.

Una tercera idea errónea que suele también tenerse es que la esperanza es igual, o por lo menos cercana, a un valor $x$ de $\mathbb{R}$ para el cual $f_X(x)$ o $p_X(x)$ es un máximo [3] . Esto ciertamente ocurre en algunas variables de más renombre como por el ejemplo la Binomial, Possion o Normal. Sin embargo no siempre es así, simplemente consideremos a una variable aleatoria con distribución exponencial de parámetro $\lambda = 100$, entonces:

$$E[X] = \int_0^{\infty} x\frac{1}{\lambda}exp\left\{-\frac{x}{\lambda}\right\}dx = \lambda = 100$$

Sin embargo al graficar la densidad (ver gráfico 2.1) se observa que el máximo se obtiene cuando $x = 0$, el cual es un valor relativamente lejano a la $E[X]$ dando un claro ejemplo de la posible diferencia de estas dos medias.

[2] Se da por hecho que se hizo una partición del intervalo (a,b) en $n$ subintervalos

[3] El punto x para el cual $f_X(x)$ ó $p_X(x)$ es máximo se le conoce como *Moda* de la distribución

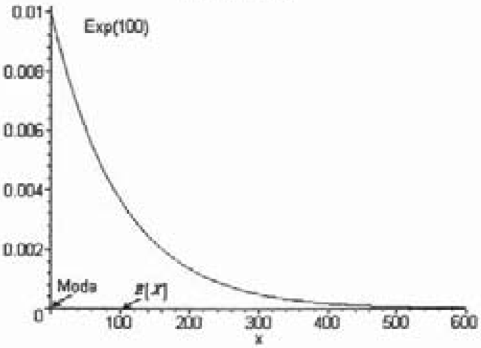
El concepto de esperanza puede generalizarse de una manera clara para el caso de funciones de variables aleatorias de la siguiente manera:

Sea $G\colon \mathbb{R} \to \mathbb{R}$ una función tal que $G(X)$ es una nueva variable aleatoria, entonces por la definición de esperanza de la variable aleatoria $G(X)$ se puede probar que:

$$E[G(X)] = \begin{cases} \int_{-\infty}^{\infty} G(x)dF_x = \int_{-\infty}^{\infty} G(x)f_X(x)dx & \text{cuando } X \text{ es v.a. absolutamente continua} \\ \\ \int_{-\infty}^{\infty} G(x)dF_x = \sum_{x_k} G(x_k)p_X(x_k) & \text{cuando } X \text{ es v.a. discreta.} \end{cases} \tag{2.8}$$

En donde se pide para que la esperanza este bien definida que la integral $\int_{-\infty}^{\infty} |G(x)|\, dF$ sea finita.

Figura 2.1:




Para el caso de vectores aleatorios el concepto anterior puede ser escrito de la siguiente manera. Sea $\mathbf{X}$=$(X_1, X_2, \ldots, X_p)$ un vector aleatorio y $G\colon \mathbb{R}^p \to \mathbb{R}$ entonces se define:

$$E[G(\mathbf{X})] = \begin{cases} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-veces} G(x_1, \ldots, x_p) f_{\mathbf{X}}(x_1, \ldots, x_p)dx_1 \ldots dx_p & \text{caso continuo} \\ \\ \sum_{x_{k_1}} \cdots \sum_{x_{k_p}} G(x_{k_1}, \ldots x_{k_p}) p_{\mathbf{X}}(x_{k_1}, \ldots, x_{k_p}) & \text{caso discreto} \end{cases} \tag{2.9}$$

En donde $f_{\mathbf{X}}$ es la función de densidad conjunta del vector aleatorio $\mathbf{X}$ para el caso continuo y $p_{\mathbf{X}}$ la función de densidad de probabilidad del vector aleatorio en el caso discreto.

## 2.4. Propiedades de la esperanza

El operador esperanza definido en la sección anterior goza de ciertas propiedades que son importantes en las aplicaciones de la estadística. A continuación se listará una serie resultados que nos serán útiles para el presente trabajo:

**Teorema 2.1 (Propiedades del Operador Esperanza).** *Sea $X$ , $Y$ variables aleatorias con esperanza finita y $c$ un número real, entonces se tiene que:*

1. $E[cX] = cE[X]$
2. $E[X+Y] = E[X] + E[Y]$ *(Linealidad)*
3. *Si $X$ y $Y$ son independientes entonces* $E[XY] = E[X]E[Y]$
4. *si $X \leq Y$ entonces* $E[X] \leq E[Y]$ *(Monotonia)*

*Demostración.* Se demostrará en orden cada una de las propiedades anteriores

1. Usando la definición de esperanza para la función de variables aleatorios $(G(X) = cX)$ tenemos:
$$E[cX] = \int_{-\infty}^{\infty} cx dF(x) = c\int_{-\infty}^{\infty} x dF(x) = cE[X]$$
Note que se usa la propiedad de linealidad de la integral de Riemann-Stieltjes

2. Supongamos que X y Y son absolutamente continuas entonces haciendo $G(X,Y) = X+Y$ tenemos:
$$E[X+Y] := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{X,Y}(x,y)dxdy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{X,Y}(x,y)dxdy$$
$$= \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy}_{f_X(x)}\, dx + \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx}_{f_Y(y)}\, dy$$

$$= \int_{-\infty}^{\infty} x f_X(x)dx + \int_{-\infty}^{\infty} y f_Y(y)dy = E[X] + E[Y]$$

La demostración para el caso discreto es análoga cambiando el símbolo de integración por el de suma.

3. Supongamos que X y Y son absolutamente continuas e independientes, entonces haciendo $G(X,Y) = XY$ tenemos:

$$E[XY] := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dxdy$$

Usando la hipótesis de independencia de las variables aleatorias se tiene que $f_{X,Y}(x,y)$ es igual al producto de las marginales $f_X(x)f_Y(y)$ por lo que sustituyendo:

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y)dxdy$$

$$= \int_{-\infty}^{\infty} x f_X(x)dx \int_{-\infty}^{\infty} y f_Y(y)dy = E[X]E[Y]$$

La demostración para el caso discreto es análoga cambiando el símbolo de integración por el de suma

4. Definamos a la variable aleatoria $Z = Y - X$. Entonces por hipótesis obtenemos que:

$$0 \leq Y - X = Z$$

Entonces $Z$ es una v.a. que toma valores positivos por lo que usando las propiedades de la integral de Riemann-Stieltjes:

$$E[Z] = \int_{0}^{\infty} z dF_z \geq 0 \Rightarrow E[Y - X] \geq 0$$

Finalmente usando la linealidad del operador esperanza obenemos:

$$E[Y] - E[X] \geq 0 \Rightarrow E[X] \leq E[Y]$$

□

## 2.5. El concepto de Esperanza Condicional

Hasta este punto hemos expuesto todas las ideas relacionadas con el tema principal de este capítulo, lo anterior fue hecho con el objetivo de tener todas las herramientas necesarias para motivar de una manera fácil una buena definición de la esperanza condicional dado un evento $B$ al cual denotaremos como $E[X \mid B]$ [4] .

Recordemos que si $X$ es una variable aleatoria discreta que toma valores $x_1, x_2, \ldots$, entonces su esperanza es igual a $\sum_k x_k P[X = x_k]$. De manera que ahora resulta lógico definir a la esperanza condicional como:

$$E[X| B] = \sum_k x_k P[X = x_k \mid B] = \sum_k x_k \frac{P[\{X = x_k\} \cap B]}{P(B)}$$

$$= \frac{1}{P(B)} \sum_k x_k P[\{X = x_k\} \cap B] = \frac{1}{P(B)} \sum_k x_k P[XI_B = x_k] = \frac{1}{P(B)} E[XI_B]$$ [5]

Ahora observe que $E[XI_B]$ no depende de la forma que tiene $X$ (absolutamente continua o discreta) por lo que se desprende la siguiente definición de manera general:

**Definición 2.8 (Esperanza condicional dado un evento).** Sea $B$ cualquier evento de probabilidad positiva y $X$ una variable aleatoria de esperanza finita [6] . Se define la esperanza condicional de $X$ dada la ocurrencia del evento $B$, $E[X| B]$, como:

$$E[X \mid B] = \frac{E[XI_B]}{P(B)} \tag{2.10}$$

**La interpretación que daremos al número $E[X| B]$ será el promedio de los valores que toma la variable aleatoria X cuando el experimento aleatorio es repetido en varias ocasiones suponiendo en cada repetición que el evento $B$ ocurrió.**

La anterior definición es una generalización de la definición de probabilidad condicional la cual mostramos a continuación.

Supongamos que tenemos dos eventos $A$ y $B$, $B$ con probabilidad positiva, entonces por definición sabemos que:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

---

[4] En esta parte siempre supondremos que el evento al cual se está condicionando tiene probabilidad positiva

[5] En esta igualdad no estamos dando por hecho que $P[XI_B = 0] = P[\{X = 0\} \cap B]$ ya que la igualdad no necesariamente es cierta, sin embargo, las igualdades arriba mencionadas son válidas ya que cuando $x_k = 0$ entonces se tiene que $x_k P[\{X = x_k\} \cap B] = 0 = x_k P[XI_B = x_k]$

[6] Note que al pedir que X tenga esperanza finita se asegura que $E[X \mid B]$ exista ya que $E[|XI_B|] \leq E[|X|]$ lo cual es consecuencia inmediata de que $|XI_B| \leq |X|$

Si ahora tomamos a $X(\omega) = I_A(\omega)$ entonces:

$$E[X \mid B] := \frac{E[I_A I_B]}{P(A)} = \frac{E[I_{A\cap B}]}{P(A)} = \frac{P(A \cap B)}{P(A)} := P(A \mid B)$$

Lo que muestra el como con la definición de esperanza condicional se obtiene la probabilidad de un evento $A$ dada la ocurrencia de otro evento $B$.

## 2.6. Esperanza condicional en el caso discreto

Ahora daremos la definición de la esperanza condicional de $X$ dado que ocurrió el evento $\{Y = y\}$ para el caso en que tanto $X$ como $Y$ son variables aleatorias de tipo discreto.

Como ya tenemos la definición de esperanza condicional de una variable aleatoria dado un evento B con probabilidad positiva, entonces para seguir siendo consistentes con esta definición resulta natural definir a $B$ como el evento $\{Y = y\}$ quedando la esperanza condicional de la siguiente forma:

$$E[X \mid Y = y] = \frac{E[X I_{Y=y}]}{P[Y = y]} = \sum_k \frac{x_k P[X = x_k, Y = y]}{P[Y = y]}$$

Ahora note que el cociente que aparece en la última expresión no es más que la densidad condicional de $X$ dado $Y = y$ evaluada en $x_k$ por lo que al sustituirlo, nuestra definición quedaría de la siguiente forma [7] :

**Definición 2.9.** Sea $X$ y $Y$ variables aleatorias del tipo discreto y sea $y$ tal que $P[Y = y] > 0$, entonces se define la esperanza condicional de $X$ dado $Y = y$, denotado por $E[X \mid Y = y]$ como:

$$E[X \mid Y = y] = \sum_k x_k p_{X \mid Y}(x_k \mid y) \tag{2.11}$$

La definición anterior arroja las siguiente observaciones:

1. La esperanza condicional de $X$ dado $Y = y$ nuevamente puede ser pensado como un promedio ponderado en donde a cada $x_k$ ahora se le asigna el peso de la probabilidad condicional de $X = x_k$ dado $Y = y$
2. Al calcular la esperanza condicional nos queda una expresión que depende de $y$ por lo que a veces se acostumbra a escribir $g(y) = E[X \mid Y = y]$, lo cual nos indica que la esperanza condicional cambia dependiendo de que valor se le asigne a $y$.
3. Un punto importante que se desprende de lo anterior es la definición de $E[X \mid Y]$ como una v.a. aleatoria que es función de Y ($g(Y) = E[X \mid Y]$).

[7] Se da por hecho que el evento $Y = y$ tiene probabilidad positiva

## 2.7. Esperanza condicional en el caso absolutamente continuo

Ahora pensaremos a $X$ y $Y$ como variables aleatorias del tipo absolutamente continuo y nuevamente el problema será definir una expresión para la esperanza condicional de $X$ dado $Y = y$.

Motivados por la definición en el caso discreto resulta ahora natural definir este concepto de esperanza condicional poniendo en lugar de la función de densidad de probabilidad condicional del caso discreto a la función de densidad continua del caso absolutamente continuo, además para seguir siendo consistentes se cambiará el símbolo de suma por el de integración.

**Definición 2.10.** Sea $X$ e $Y$ variables absolutamente continuas e $y$ un número real tal que $f_Y(y) > 0$. Se define la Esperanza condicional de $X$ dado $Y = y$ como:

$$E[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X\mid Y}(x \mid y) dx \tag{2.12}$$

Las observaciones que surgen tras esta definición son similares al caso discreto, pues nuevamente la esperanza condicional puede ser vista como un promedio ponderado viendo a la integral como una suma de Riemann. Note además que nuevamente la esperanza condicional queda en función del valor que toma $y$.

Estas definiciones también pueden ser llevadas para el caso de funciones de variables aleatorias y de vectores aleatorios de la siguiente forma:

Sea $X_1, \ldots, X_p$, $Y$ variables aleatorias y $G\colon \mathbb{R}^p \to \mathbb{R}$ una función tal que $G(X_1, \ldots, X_p)$ es nuevamente una variable aleatoria con esperanza finita, entonces se puede probar que:

$$E[G(\mathbf{X}) \mid Y = y] = \begin{cases} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-veces} G(x_1, \ldots, x_p) f_{\mathbf{X}\mid Y}(x_1, \ldots, x_p \mid y) dx_1 \ldots dx_p & \text{caso continuo} \\ \\ \sum_{x_{k_1}} \cdots \sum_{x_{k_p}} G(x_{k_1}, \ldots x_{k_p}) p_{\mathbf{X}\mid Y}(x_{k_1}, \ldots, x_{k_p} \mid y) & \text{caso discreto} \end{cases} \tag{2.13}$$

En donde, para que todo esté bien definido pediremos que $y$ sea un número real tal que $f_Y(y) > 0$ en el caso absolutamente continuo o $p_Y(y) = P[Y = y] > 0$ en el caso discreto.

## 2.8. Propiedades de la esperanza condicional

La esperanza condicional para variables aleatorias discretas o absolutamente continuas definida anteriormente goza de las propiedades descritas en el teorema siguiente:

**Teorema 2.2 (Propiedades de la esperanza condicional).** *Sea $X, Y, Z$ variables aleatorias discretas o absolutamente continuas con esperanzas finitas y $c, y \in \mathbb{R}$ tal que $f_Y(y) > 0$ si $Y$ es absolutamente continua o $p_Y(y) > 0$ si $Y$ es discreta. Entonces se cumple lo siguiente:*

1. $E[cX \mid Y = y] = cE[X \mid Y = y]$
2. $E[X + Z \mid Y = y] = E[X \mid Y = y] + E[Z \mid Y = y]$
3. *Si $X$ y $Y$ son independientes, entonces* $E[X \mid Y = y] = E[X]$
4. $E[H(X)G(Y) \mid Y = y] = G(y)\, E[H(X) \mid Y = y]$

*Demostración.* La prueba se llevará a cabo en orden y únicamente para el caso en que $X$,$Y$, y $Z$ son variables aleatorias absolutamente continuas. (*El inciso 4 no es demostrado en este trabajo por su complejidad y falta de herramientas para su demostración* sin embargo la demostración puede ser consultada en "Probability and Measure Theory"[7] en la página 220)

1. Sea $G(X) = cX$, entonces por definición tenemos:

$$E[cX \mid Y = y] := \int_{-\infty}^{\infty} cx f_{X|Y}(x|\, y)dx = c\int_{-\infty}^{\infty} x f_{X|Y}(x|\, y)dx := cE[X \mid Y = y]$$

2. Sea $G(X) = X + Z$, entonces por definición tenemos:

$$E[X + Z \mid Y = y] := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x + z) f_{X,Z|Y}(x, z \mid y)dxdz$$

$$= \int_{-\infty}^{\infty} \frac{x}{f_Y(y)} \underbrace{\int_{-\infty}^{\infty} f_{X,Z,Y}(x, y, z)dz}_{f_{X,Y}(x,y)}\, dx + \int_{-\infty}^{\infty} \frac{z}{f_Y(y)} \underbrace{\int_{-\infty}^{\infty} f_{X,Z,Y}(x, y, z)dx}_{f_{Z,Y}(z,y)}\, dz$$

$$= \int_{-\infty}^{\infty} x \underbrace{\frac{f_{X,Y}(x, y)}{f_Y(y)}}_{f_{X|Y}(x|y)} dx + \int_{-\infty}^{\infty} z \underbrace{\frac{f_{Z,Y}(z, y)}{f_Y(y)}}_{f_{Z|Y}(z|y)} dz = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)dx + \int_{-\infty}^{\infty} z f_{Z|Y}(z \mid y)dz$$

$$:= E[X \mid Y = y] + E[Z \mid Y = y]$$

3. Supongamos que tanto X como Y son independientes, lo que implicaría que la densidad conjunta de $X$ e $Y$ es igual al producto de sus respectivas marginales ($f_{X,Y}(x,y) = f_X(x)f_Y(y)$). Entonces tendríamos lo siguiente:

$$E[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X\mid Y}(x \mid y)dx$$

$$= \int_{-\infty}^{\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)} \underbrace{=}_{hip} \int_{-\infty}^{\infty} x f_X(x) \frac{f_Y(y)}{f_Y(y)} dx = E[X]$$

□

Observe que una consecuencia importante que obtenemos de las propiedades anteriores es que:

$$E[G(X) \mid X = t] = G(t)$$

## 2.9. Algunos Resultados Importantes

En esta sección daremos algunos resultados importantes relacionados con la densidad condicional de los estadísticos de orden y esperanza condicional dado un evento, mismos que serán utilizados en el capítulo 3 de este trabajo y nos serán de gran ayuda para poder llevar a cabo la caracterización de las funciones de distribución que pretendemos.

Los dos primeros teoremas que se tratan a continuación están basados en la prueba expuesta en la obra de H.A. David [1] en las páginas 265-278.

**Teorema 2.3 (Distribución condicional de $Y_{r:n}| Y_{k:n} = t$ con $k < r$).** *Sea $Y_{1:n} \leq Y_{2:n} \leq \ldots \leq Y_{n:n}$ los estadísticos de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una función de distribución $F(\cdot)$ continua y sea $1 \leq k < r \leq n$. Entonces $Y_{r:n}| Y_{k:n} = t$ tiene la misma distribución que el $(r-k)$ estadístico de orden de una muestra aleatoria de tamaño $n-k$ proveniente de un distribución dada por:*

$$G(x) = \begin{cases} \dfrac{F(x) - F(t)}{1 - F(t)} & x \geq t \\ \\ 0 & x < t \end{cases} \tag{2.14}$$

*Demostración.* Por definición tenemos que:

$$f_{Y_{r:n}|Y_{k:n}}(y_r| \, y_k = t) = \frac{f_{Y_{r:n},Y_{k:n}}(y_r, t)}{f_{Y_{k:n}}(t)}$$

Por lo que usando la fórmula (1.13) para la densidad conjunta de dos estadístico de orden y la formula (1.7) para la densidad del estadístico de orden $k$ tenemos que lo anterior es igual a:

$$= \frac{n!F(t)^{k-1}\left[F(y_r)-F(t)\right]^{r-k-1}\left[1-F(y_r)\right]^{n-r}f(t)\,f(y_r)\,(k-1)!\,(n-k)!}{(k-1)!\,(r-k-1)!\,(n-r)!n!F(t)^{k-1}\left[1-F(t)\right]^{n-k}f(t)}$$

Lo cual tras eliminar términos que multiplican y dividen queda de la siguiente forma:

$$f_{Y_{r:n}|Y_{k:n}}(y_r|\,y_k=t) = \frac{(n-k)!\left[F(y_r)-F(t)\right]^{r-k-1}\left[1-F(y_r)\right]^{n-r}f(y_r)}{(r-k-1)!\,(n-r)!\left[1-F(t)\right]^{n-k}} \tag{2.15}$$

Así pues, hemos probado hasta esta parte que la densidad condicional $Y_{r:n}|\,Y_{k:n}=t$ está dada por la ecuación (2.15) por lo que para terminar con la prueba del teorema lo que resta es verificar que en efecto esta ecuación es igual la densidad del $(r-k)$ estadístico de orden de una muestra aleatoria de tamaño $(n-k)$ proveniente de la distribución $G(x)$

Sea $X_1',\ldots,X_{n-k}'$ una muestra aleatoria de $G(x)$. Aplicando la fórmula (1.7) para encontrar el $(r-k)$ estadístico de orden de una muestra de tamaño $(n-k)$ tenemos:

$$g_{Y_{r-k:n-k}}(y_{r-k}) = \frac{(n-k)!G(y_{r-k})^{r-k-1}\left[1-G(y_{r-k})\right]^{n-k-r+k}G'(y_{r-k})}{(r-k-1)!\,(n-k-r+k)!}$$

Como:

$$G'(y_{r-k}) = \left(\frac{F(y_{r-k})-F(t)}{1-F(t)}\right)' = \frac{f(y_{r-k})}{1-F(t)}$$

$$\left[1-G(y_{r-k})\right]^{n-r} = \left[\frac{1-F(y_{r-k})}{1-F(t)}\right]^{n-r}$$

Entonces tenemos que la densidad es de la forma:

$$g_{Y_{r-k:n-k}}(y_{r-k}) = \frac{(n-k)!\left[F(y_{r-k})-F(t)\right]^{r-k-1}\left[1-F(y_{r-k})\right]^{n-r}f(y_{r-k})}{(r-k-1)!\,(n-r)!\left[1-F(t)\right]^{r-k-1}\left[1-F(t)\right]^{n-r}\left[1-F(t)\right]}$$

$$g_{Y_{r-k:n-k}}(y_{r-k}) = \frac{(n-k)!\left[F(y_r)-F(t)\right]^{r-k-1}\left[1-F(y_r)\right]^{n-r}f(y_r)}{(r-k-1)!\,(n-r)!\left[1-F(t)\right]^{n-k}}$$

Lo cual es lo mismo a lo que habíamos llegado en la ecuación (2.15). Por lo tanto hemos probado que la densidad de $Y_{r:n}|\,Y_{k:n}=t$ es igual a la densidad del $r-k$ estadístico de orden de una muestra de tamaño $n-k$ proveniente de una distribución dada por $G(x)$ □

**Teorema 2.4 (Distribución condicional de $Y_{r:n}|\,Y_{k:n}=t$ con $r<k$).** *Sea $Y_{1:n}\leq Y_{2:n}\leq\ldots\leq Y_{n:n}$ los estadísticos de orden de una muestra aleatoria $X_1,X_2,\ldots,X_n$ de una función de distribución $F(\cdot)$ continua y sea $1\leq r<k\leq n$. Entonces $Y_{r:n}|\,Y_{k:n}=t$ tiene la misma*

*distribución que el r-ésimo estadístico de orden de una muestra aleatoria de tamaño $k-1$ proveniente de una distribución dada por:*

$$G(x) = \begin{cases} \frac{F(x)}{F(t)} & x \leq t \\ \\ 1 & x > t \end{cases} \tag{2.16}$$

*Demostración.* Por definición de densidad condicional tenemos que:

$$f_{Y_{r:n}|Y_{k:n}}(y_{r:n}|\, y_k = t) = \frac{f_{Y_{r:n},Y_{k:n}}(y_r, t)}{f_{Y_{k:n}}(t)}$$

Ahora recordando que $r < k$ y usando la fórmula para la densidad conjunta para dos estadísticos de orden tenemos que:

$$f_{Y_{r:n}|Y_{k:n}}(y_r \mid y_k = t) = \frac{(k-1)!F(y_r)^{r-1}\left[F(t) - F(y_r)\right]^{k-r-1} f(y_r)}{(r-1)!(k-r-1)!F(t)^{k-1}} \tag{2.17}$$

Por lo que ahora basta probar que la expresión (2.17) es igual a la expresión obtenida al encontrar la distribución del r-ésimo estadístico de orden de una muestra aleatoria de tamaño $k-1$ proveniente de la distribución dada por $G(x)$.

Sea $X_1', \ldots, X_{k-1}'$ una muestra aleatoria de $G(x)$. Entonces aplicando la formula (1.7) tenemos:

$$g_{Y_{r:k-1}}(y_r) = \frac{(k-1)!G(y_r)^{r-1}\left[1 - G(y_r)\right]^{k-r-1} G'(y_r)}{(r-1)!(k-r-1)!}$$

Pero por hipótesis, sabemos que

$$G(y_r) = \frac{F(y_r)}{F(t)}$$

$$[1 - G(y_r]^{k-r-1} = \left\{\frac{F(t) - F(y_r)}{F(t)}\right\}^{k-r-1}$$

$$G'(y_r) = \frac{f(y_r)}{F(t)}$$

Por lo que sustituyendo obtenemos que:

$$g_{Y_{r:k-1}}(y_r) = \frac{(k-1)!F(y_r)^{r-1}\left[F(t) - F(y_r)\right]^{k-r-1} f(y_r)}{(r-1)!(k-r-1)!F(t)^{k-1}} \tag{2.18}$$

Finalmente como la ecuación ( 2.17) es igual a la ecuación (2.18) se prueba que en efecto la densidad de $Y_{r:n}|\, Y_{k:n} = t$ $(r < k)$ es igual a la densidad del r-ésimo estadístico de orden de una muestra aleatoria de tamaño $k-1$ proveniente de una distribución dada por $G(x)$ $\square$

**Teorema 2.5.** *Sea $X$ variable aleatoria absolutamente continua con esperanza finita. Entonces:*

$$\lim_{n\to-\infty} E[X \mid X > n] = E[X]$$

*Demostración.* Por la definición dada en (2.8) tenemos:

$$\lim_{n\to-\infty} E[X \mid X > n] = \lim_{n\to-\infty} \frac{E[XI_{\{X>n\}}]}{P[X>n]} = \lim_{n\to-\infty} \frac{\int_n^\infty xf_X(x)dx}{P[X>n]}$$

$$= \frac{\lim_{n\to-\infty}\int_n^\infty xf_X(x)dx}{\lim_{n\to-\infty} P[X>n]} = \frac{\int_{-\infty}^\infty xf_X(x)dx}{\lim_{n\to-\infty} P[X>n]} = E[X]$$

Lo anterior se justifica haciendo $A_n = \{X > n\}$ (observe que $A_n$ converge a $\Omega$ cuando $n \to -\infty$) lo que implicaría que:

$$\lim_{n\to-\infty} P[X > n] = \lim_{n\to-\infty} P[A_n] = P[\Omega] = 1$$

Para ver con más detalle este paso, puede consultar "*Probability and Measure Theory*" [7] en la página 9. □

Observe que de manera análoga, se tiene que:

$$\lim_{n\to\infty} E[X \mid X < n] = E[X]$$

**Teorema 2.6.** *Sean $\alpha$ y $\beta$ los extremos izquierdo y derecho de una distribución absolutamente continua $F$. Sea $g : [\alpha, \beta) \to \mathbb{R}$ una función con primera derivada continua tal que $\lim_{x\to\beta} g(x) = \infty$ y supongamos que $E[g(X)] < \infty$. Entonces*

$$\int_\alpha^\beta g'(x)[1-F(x)]dx = E[g(X)] - g(\alpha)$$

*Demostración.*

$$\int_\alpha^\beta g'(x)[1-F(x)]dx = \int_\alpha^\beta g'(x)P[X>x]dx = \int_\alpha^\beta g'(x)\left\{\int_x^\beta f_X(y)dy\right\}dx = \int_\alpha^\beta\int_x^\beta g'(x)f_X(y)dydx$$

Intercambiando el orden de integración

$$= \int_\alpha^\beta\int_\alpha^y g'(x)f_X(y)dxdy = \int_\alpha^\beta f_X(y)\left\{\int_\alpha^y g'(x)dx\right\}dy$$

Utilizando el Teorema Fundamental del Cálculo para la integral respecto a x

$$= \int_\alpha^\beta f_X(y)\{g(y) - g(\alpha)\}\,dy = \underbrace{\int_\alpha^\beta g(y)f_X(y)dy}_{E[g(X)]} - g(\alpha)\underbrace{\int_\alpha^\beta f_X(y)dy}_{1}$$

Por lo tanto

$$\int_\alpha^\beta g'(x)[1 - F(x)]dx = E[g(X)] - g(\alpha)$$

□

**Teorema 2.7.** *Sean $\alpha$ y $\beta$ los extremos izquierdo y derecho de una distribución absolutamente continua $F$. Sea $g : [\alpha, \beta) \to \mathbb{R}$ una función con primera derivada continua tal que $\lim_{x\to\beta} g(x) = \infty$ y supongamos que $E[g(X)] < \infty$. Entonces*

$$\lim_{x\to\beta} g(x)[1 - F(x)] = 0$$

*Demostración.* Integrando por partes:

$$\int_\alpha^\beta g'(x)[1 - F(x)]dx = g(x)[1 - F(x)]\Big|_\alpha^\beta + \underbrace{\int_\alpha^\beta g(x)f_X(x)dx}_{E[g(X)]}$$

$$\Rightarrow \int_\alpha^\beta g'(x)[1 - F(x)]dx = \lim_{x\to\beta} g(x)[1 - F(x)] - g(\alpha) + E[g(X)]$$

Entonces usando el teorema anterior:

$$E[g(X)] - g(\alpha) = \lim_{x\to\beta} g(x)[1 - F(x)] - g(\alpha) + E[g(X)]$$

$$\Rightarrow 0 = \lim_{x\to\beta} g(x)[1 - F(x)]$$

□

Note que como $[1 - F(x)]^n \leq [1 - F(x)]$ para todo $n \in N$ entonces:

$$\Rightarrow \lim_{x\to\beta} g(x)[1 - F(x)]^n = 0 \qquad \forall n \in N$$

De la misma forma que en los teoremas anteriores se puden probar los siguientes resultados:

**Teorema 2.8.** *Sean $\alpha$ y $\beta$ los extremos izquierdo y derecho de una distribución absolutamente continua $F$. Sea $g : (\alpha, \beta] \to \mathbb{R}$ una función con primera derivada continua tal que $\lim_{x\to\alpha} g(x) = \infty$ y supongamos que $E[g(X)] < \infty$. Entonces*

$$\int_{\alpha}^{\beta} g'(x)F(x)dx = g(\beta) - E[g(X)]$$

**Teorema 2.9.** *Sean $\alpha$ y $\beta$ los extremos izquierdo y derecho de una distribución absolutamente continua $F$. Sea $g : (\alpha, \beta] \to \mathbb{R}$ una función con primera derivada continua tal que $\lim_{x\to\alpha} g(x) = \infty$ y supongamos que $E[g(X)] < \infty$. Entonces*

$$\lim_{x\to\alpha} g(x)F(x) = 0$$

Y nuevamente de esto último se concluye que:

$$\lim_{x\to\beta} g(x)F(x)^n = 0 \qquad \forall n \in N$$

# Capítulo 3

# Una caracterización de distribuciones

Hasta estos momentos se han expuesto todas la herramientas necesarias para poder comenzar la caracterización de funciones de distribución por medio de la esperanza condicional de sus estadísticos de orden. De esta manera, este capítulo contará con varios teoremas y corolarios que nos dirán bajo qué condiciones podremos construir una función de distribución, sabiendo la forma general que toma la esperanza condicional de los estadísticos de orden.

Comenzaremos con el teorema (3.1) el cual nos dice que si tenemos la esperanza condicional de la forma:

$$E\left[g\left(Y_{r:n}\right) \mid Y_{k:n}=t\right]=g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j} \qquad r<k \qquad \forall t$$

Entonces bajo adecuadas condiciones dadas al dominio y forma de la función $g$ podemos afirmar que la distribución, que siguen las variables aleatorias que dieron origen a los estadísticos de orden, es:

$$F(x)=1-exp\left\{-\frac{1}{c}\left[g(x)-g(\alpha)\right]\right\}$$

Posteriormente se mencionarán otras formas de esperanza condicional que también caracterizarán a estas mismas funciones de distribución.

Más adelante, también se mencionan los teoremas para el caso en que F es de la forma:

$$F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$$

Y finalmente terminamos con algunos ejemplos y aplicaciones de todos estos teoremas.

## 3.1. Teoremas de caracterización

Las demostraciones de los teoremas (3.1),(3.2) y (3.3) que a continuación se presentan están basadas en un artículo que publicó L.Y. Ouyang [11] en 1996 por lo que estas demostraciones seguirán el mismo razonamiento que él siguió complementándolo con algunos pasos que omitió como obvios.

**Teorema 3.1.** *Sea $c \in \mathbb{R} - \{0\}$, consideremos $\alpha \geq -\infty$ y $\beta \leq \infty$ los extremos izquierdo y derecho de una función de distribución $F(\cdot)$ de una variable absolutamente continua $X$ con $0 < F(x) < 1$ $\forall x \in (\alpha, \beta)$, y sea $g\colon [\alpha, \beta) \to \mathbb{R}$ una función continua sobre el intervalo $[\alpha, \beta)$ con primera derivada continua sobre $(\alpha, \beta)$ y tal que $g(x) \to^{+}_{-} \infty$ cuando $x \to \beta^{-}$ si $c > 0$ o $c < 0$ respectivamente. Si $E[g(X)] < \infty$, entonces:*

$$E\left[g\left(Y_{r:n}\right) | Y_{k:n} = t\right] = g(t) + c\sum_{j=k}^{r-1} \frac{1}{n-j} \Leftrightarrow F(x) = 1 - exp\left\{-\frac{1}{c}\left[g(x) - g(\alpha)\right]\right\}$$

*con $1 \leq k < r \leq n$ , $\forall t \in (\alpha, \beta)$*

*Demostración.* Usando el resultado del teorema (2.3) que nos da la densidad condicional de $Y_{r:n} \mid Y_{k:n} = t$ con $k < r$ y por definición de esperanza condicional para el caso absolutamente continuo dado en (2.13) tenemos:

$$E\left[g\left(Y_{r:n}\right) | Y_{k:n} = t\right] = A_1 \int_t^{\beta} g\left(y_r\right) \left\{\frac{\left[F\left(y_r\right) - F(t)\right]^{r-k-1}\left[1 - F\left(y_r\right)\right]^{n-r}}{\left[1 - F(t)\right]^{n-k}}\right\} dF\left(y_r\right) \quad (3.1)$$

donde

$$A_1 = \frac{(n-k)!}{(r-k-1)!\,(n-r)!}$$

Dado lo anterior, comenzaremos la prueba con la parte de suficiencia.

- $\Leftarrow$:]

Suponiendo que $F(x) = 1 - exp\left\{-\frac{1}{c}\left[g(x) - g(\alpha)\right]\right\}$ tenemos las siguientes cuatro igualdades:

$$F\left(y_r\right) - F(t) = exp\left\{-\frac{1}{c}\left[g(t) - g(\alpha)\right]\right\} - exp\left\{-\frac{1}{c}\left[g\left(y_r\right) - g(\alpha)\right]\right\} \quad (3.2)$$

$$1 - F\left(y_r\right) = exp\left\{-\frac{1}{c}\left[g\left(y_r\right) - g(\alpha)\right]\right\} \quad (3.3)$$

$$dF\left(y_r\right) = F'\left(y_r\right) dy_r = \frac{1}{c} exp\left\{-\frac{1}{c}\left[g\left(y_r\right) - g(\alpha)\right]\right\} g'\left(y_r\right) dy_r \quad (3.4)$$

Sustituyendo las ecuaciones (3.2), (3.3), (3.4) en la ecuación (3.1) tenemos que la esperanza condicional es igual a:

$$= A_1 \int_t^\beta g(y_r) \frac{[F(y_r) - F(t)]^{r-k-1} exp\left\{-\frac{n-r}{c}[g(y_r) - g(\alpha)]\right\}}{[1 - F(t)]^{n-k}} *$$

$$\frac{1}{c} exp\left\{-\frac{1}{c}[g(y_r) - g(\alpha)]\right\} g'(y_r)\, dy_r$$

Lo cual si es simplificado y reacomodado

$$= \frac{A_1}{[1 - F(t)]^{n-k}} \int_t^\beta \frac{g(y_r)\, g'(y_r)}{c} [F(y_r) - F(t)]^{r-k-1} *$$

$$exp\left\{-\frac{n-r+1}{c}[g(y_r) - g(\alpha)]\right\} dy_r \qquad (3.5)$$

Integrando por partes en (3.5)

$$u = g(y_r)[F(y_r) - F(t)]^{r-k-1} \qquad v = exp\left\{-\frac{n-r+1}{c}[g(y_r) - g(\alpha)]\right\}$$

$$\Rightarrow du = g(y_r)(r-k-1)[F(y_r) - F(t)]^{r-k-2} exp\left\{-\frac{1}{c}[g(y_r) - g(\alpha)]\right\} \frac{1}{c} g'(y_r) +$$

$$[F(y_r) - F(t)]^{r-k-1} g'(y_r)\, dy_r$$

$$\Rightarrow dv = exp\left\{-\frac{n-r+1}{c}[g(y_r) - g(\alpha)]\right\} \left[-\frac{n-r+1}{c} g'(y_r)\right] dy_r$$

Entonces

$$(3.5) = \frac{-A_1}{[1 - F(t)]^{n-k}(n-r+1)} \int_t^\beta u\,dv = \frac{-A_1}{[1 - F(t)]^{n-k}(n-r+1)} \left[-\int_t^\beta v\,du + uv\Big|_t^\beta\right]$$

Sustituyendo $u$, $v$ y evaluando [1] :

$$g(y_r)[F(y_r) - F(t)]^{r-k-1} exp\left\{-\frac{n-r+1}{c}[g(y_r) - g(\alpha)]\right\}\Bigg|_t^\beta = 0$$

[1] Note que por las condiciones que cumple la función $g$, lo que en realidad estamos haciendo al evaluar en $\beta$ es tomar el límite cuando $y_r$ tiende a $\beta$ esto es:

$$g(y_r)[F(y_r) - F(t)]^{r-k-1} exp\left\{-\frac{n-r+1}{c}[g(y_r) - g(\alpha)]\right\}\Bigg|_t^\beta =$$

$$\lim_{y_r \to \infty} g(y_r)[F(y_r) - F(t)]^{r-k-1} exp\left\{-\frac{n-r+1}{c}[g(y_r) - g(\alpha)]\right\} = 0$$

Lo cual es válido por el hecho de que $\lim_{y_r \to \infty} g(y_r) exp\left\{-\frac{1}{c} g(y_r)\right\} = 0$

$$\int_t^\beta vdu = \int_t^\beta g(y_r)(r-k-1)\left[F(y_r)-F(t)\right]^{r-k-2} exp\left\{-\frac{1}{c}\left[g(y_r)-g(\alpha)\right]\right\}$$

$$\frac{1}{c}g^{'}(y_r)\, exp\left\{-\frac{n-r+1}{c}\left[g(y_r)-g(\alpha)\right]\right\}dy_r \qquad +$$

$$\int_t^\beta exp\left\{-\frac{n-r+1}{c}\left[g(y_r)-g(\alpha)\right]\right\}\left[F(y_r)-F(t)\right]^{r-k-1} g^{'}(y_r)\, dy_r$$

De donde después de sustituir el valor de $A_1$ obtenemos que (3.5) tiene la siguiente forma:

$$= \frac{A_2}{\left[1-F(t)\right]^{n-k}}\int_t^\beta g^{'}(y_r)\left[F(y_r)-F(t)\right]^{r-k-1} exp\left\{-\frac{n-r+1}{c}\left[g(y_r)-g(\alpha)\right]\right\}dy_r$$

$$+\frac{(n-k)!}{(r-k-2)!(n-r+1)!}\int_t^\beta g(y_r)\left\{\frac{\left[F(y_r)-F(t)\right]^{r-k-2}\left[1-F(y_r)\right]^{n-r+1}}{\left[1-F(t)\right]^{n-k}}\right\}dF(y_r) \quad (3.6)$$

donde

$$A_2 = \frac{(n-k)!}{(r-k-1)!\,(n-r+1)!}$$

Ahora note que el segundo sumando de la expresión anterior es igual a $E\left[g\left(Y_{r-1:n}\right)\mid Y_{k:n}=t\right]$ (Ver ecuación (3.1)) de donde obtenemos que:

$$E\left[g\left(Y_{r:n}\right)\mid Y_{k:n}=t\right] = \frac{A_2}{\left[1-F(t)\right]^{n-k}}\int_t^\beta g^{'}(y_r)\left[F(y_r)-F(t)\right]^{r-k-1} *$$

$$exp\left\{-\frac{n-r+1}{c}\left[g(y_r)-g(\alpha)\right]\right\}dy_r + E\left[g\left(Y_{r-1:n}\right)\mid Y_{k:n}=t\right] \quad (3.7)$$

Resolvemos la integral involucrada en (3.7) por el método de integración por partes:

Sea

$u = \left[F(y_r)-F(t)\right]^{r-k-1}$ y $v = exp\left\{-\frac{n-r+1}{c}\left[g(y_r)-g(\alpha)\right]\right\}$

$\Rightarrow \quad du = \frac{1}{c}(r-k-1)\left[F(y_r)-F(t)\right]^{r-k-2} exp\left\{-\frac{1}{c}\left[g(y_r)-g(\alpha)\right]\right\} g^{'}(y_r)\, dy_r$

$\Rightarrow \quad dv = -\frac{n-r+1}{c} exp\left\{-\frac{n-r+1}{c}\left[g(y_r)-g(\alpha)\right]\right\} g^{'}(y_r)\, dy_r$

Por lo que la integral en (3.7) es igual a:

$$= -\frac{c}{n-r+1}\int_t^\beta udv = -\frac{c}{n-r+1}\left[-\int_t^\beta vdu + uv\Big|_t^\beta\right]$$

De donde tras sustituir y evaluar[2]

$$[F(y_r) - F(t)]^{r-k-1} exp\left\{-\frac{n-r+1}{c}\left[g(y_r) - g(\alpha)\right]\right\}\Bigg|_t^{\beta} = 0$$

$$\int_t^{\beta} vdu = \frac{r-k-1}{c}\int_t^{\beta} g'(y_r)\left[F(y_r) - F(t)\right]^{r-k-2} exp\left\{-\frac{n-r+2}{c}\left[g(y_r) - g(\alpha)\right]\right\} dy_r$$

Entonces al resolver la integral por partes hemos llegado a que:

$$\int_t^{\beta} g'(y_r)\left[F(y_r) - F(t)\right]^{r-k-1} exp\left\{-\frac{n-r+1}{c}\left[g(y_r) - g(\alpha)\right]\right\} dy_r = \frac{r-k-1}{n-r+1} *$$

$$\int_t^{\beta} g'(y_r)\left[F(y_r) - F(t)\right]^{(r-1)-k-1} exp\left\{-\frac{n-(r-1)+1}{c}\left[g(y_r) - g(\alpha)\right]\right\} dy_r \tag{3.8}$$

Ahora note que de esta última igualdad la integral que aparece en el lado izquierdo es igual a la integral del lado derecho haciendo la sustitución de $r$ por $r-1$, con la diferencia que aparece el factor $\frac{r-k-1}{n-r+1}$ multiplicando a la integral del lado derecho.

Así pues de la igualdad (3.8) se puede observar que si llevamos a cabo $(r-k-1)$ veces el mismo proceso de integración antes descrito llegamos a que la integral que aparece en la expresión (3.7) es:

$$= \frac{(r-k-1)(r-k-2)\dots 1}{(n-r+1)(n-r+2)\dots(n-r+(r-k-1))}\int_t^{\beta} g'(y_r) *$$

$$exp\left\{-\frac{n-r+r-k}{c}\left[g(y_r) - g(\alpha)\right]\right\} dy_r$$

$$= \frac{(r-k-1)!\,(n-r)!}{(n-k-1)!}\int_t^{\beta} g'(y_r)\, exp\left\{-\frac{n-k}{c}\left[g(y_r) - g(\alpha)\right]\right\} dy_r \tag{3.9}$$

Por lo que sustituyendo este último resultado en la ecuación (3.7) llegamos a que la esperanza condicional $E\left[g(Y_{r:n}) \mid Y_{k:n} = t\right]$ bajo la hipótesis de suficiencia se puede expresar como:

$$= \frac{n-k}{(n-r+1)\left[1 - F(t)\right]^{n-k}}\int_t^{\beta} g'(y_r)\, exp\left\{-\frac{n-k}{c}\left[g(y_r) - g(\alpha)\right]\right\} dy_r +$$

$$E\left[g(Y_{r-1:n}) \mid Y_{k:n} = t\right] \tag{3.10}$$

[2] Recuerde que

$$\lim_{y_{(r)}\to\beta^-} exp\left\{-\frac{n-r+1}{c}\left[g(y_r) - g(\alpha)\right]\right\} = 0$$

Ahora resolveremos la integral involucrada en (3.10). Para ello procederemos por el método de cambio de variable:

Sea

$$u = -\frac{n-k}{c}\left[g\left(y_r\right) - g\left(\alpha\right)\right] \Rightarrow du = -\frac{n-k}{c}g'\left(y_r\right)$$

Entonces la integral de la expresión (3.10) toma la siguiente forma:

$$= -\frac{c}{n-k}\int_{-\frac{n-k}{c}[g(t)-g(\alpha)]}^{-\infty} exp\left\{u\right\}du = -\frac{c}{n-k}\left\{0 - exp\left\{-\frac{n-k}{c}\left[g\left(t\right) - g\left(\alpha\right)\right]\right\}\right\}$$

$$= \frac{c}{n-k}exp\left\{-\frac{1}{c}\left[g\left(t\right) - g\left(\alpha\right)\right]\right\}^{n-k} = \frac{c}{n-k}\left[1 - F\left(t\right)\right]^{n-k} \tag{3.11}$$

De donde finalmente tras sustituir este último resultado en la ecuación (3.10) obtenemos que bajo la hipótesis de suficiencia:

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = \frac{c}{n-r+1} + E\left[g\left(Y_{r-1:n}\right)\middle| Y_{k:n} = t\right] \tag{3.12}$$

Ahora usando recurrentemente lo anterior:

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = \frac{c}{n-r+1} + \frac{c}{n-r+2}\cdots\frac{c}{n-k} + E\left[g\left(Y_{k:n}\right)\middle| Y_{k:n} = t\right]$$

Y como $E\left[g\left(Y_{k:n}\right)\middle| Y_{k:n} = t\right] = g\left(t\right)$ entonces se tiene:

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g\left(t\right) + c\sum_{j=k}^{r-1}\frac{1}{n-j}$$

Con lo cual termina la prueba de suficiencia.

- $\Rightarrow$:]

Para comenzar con la otra parte de la prueba partiremos de la ecuación (3.1) la cual es válida por definición de esperanza condicional para el caso absolutamente continuo.

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = A_1\int_t^{\beta} g\left(y_r\right)\left\{\frac{\left[F\left(y_r\right) - F\left(t\right)\right]^{r-k-1}\left[1 - F\left(y_r\right)\right]^{n-r}}{\left[1 - F\left(t\right)\right]^{n-k}}\right\}dF\left(y_r\right)$$

Pero por hipótesis:

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g\left(t\right) + c\sum_{j=k}^{r-1}\frac{1}{n-j}$$

Por lo tanto:

$$\left[g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1-F(t)]^{n-k}=A_1\underbrace{\int_t^{\beta}g(y_r)\left[F(y_r)-F(t)\right]^{r-k-1}\left[1-F(y_r)\right]^{n-r}dF(y_r)}_{I} \tag{3.13}$$

Ahora nos centraremos en resolver la integral del lado derecho de la ecuación anterior por medio del método de integración por partes.

Para ello hacemos:

$$u=g(y_r)\left[F(y_r)-F(t)\right]^{r-k-1}\qquad \text{y} \qquad v=\left[1-F(y_r)\right]^{n-r+1}$$

$$\Rightarrow du=g(y_r)(r-k-1)\left[F(y_r)-F(t)\right]^{r-k-2}dF(y_r)+\left[F(y_r)-F(t)\right]^{r-k-1}g'(y_r)\,dy_r$$

$$\Rightarrow dv=-(n-r+1)\left[1-F(y_r)\right]^{n-r}dF(y_r)$$

Con las sustituciones anteriores la integral I es igual a:

$$I=-\frac{1}{n-r+1}\int_t^{\beta}udv=-\frac{1}{n-r+1}\left[-\int_t^{\beta}vdu+uv\Big|_t^{\beta}\right]$$

Pero como[3]

$$g(y_r)\left[F(y_r)-F(t)\right]^{r-k-1}\left[1-F(y_r)\right]^{n-r+1}\Big|_t^{\beta}=0$$

Entonces:

$$I=\frac{r-k-1}{n-r+1}\int_t^{\beta}\left[1-F(y_r)\right]^{n-r+1}g(y_r)\left[F(y_r)-F(t)\right]^{r-k-2}dF(y_r)+$$

$$\frac{1}{n-r+1}\int_t^{\beta}\left[1-F(y_r)\right]^{n-r+1}\left[F(y_r)-F(t)\right]^{r-k-1}g'(y_r)\,dy_r$$

Por lo tanto, sustituyendo el resultado de la igualdad anterior en la ecuación (3.13) tenemos que:

$$\left[g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1-F(t)]^{n-k}=$$

$$\frac{(n-k)!}{(r-k-2)!\,(n-r+1)!}\int_t^{\beta}g(y_r)\left[F(y_r)-F(t)\right]^{r-k-2}\left[1-F(y_r)\right]^{n-r+1}dF(y_r)+$$

[3] En esta parte utilizamos el hecho de que

$$\lim_{x\to\beta}g(x)[1-F(x)]^n=0$$

Lo cual fue probado en el teorema (2.7) del capítulo 2.

$$\frac{(n-k)!}{(r-k-1)!\,(n-r+1)!}\int_t^{\beta} g'(y_r)\,[F(y_r)-F(t)]^{r-k-1}\,[1-F(y_r)]^{n-r+1}\,dy_r$$

Ahora note que el primer sumando de la expresión anterior es igual al lado derecho de la ecuación (3.13) pero haciendo la sustitución[4] de $r$ por $r-1$. Entonces sustituyendo el valor de la integral del primer sumando y usando la ecuación (3.13) hemos probado que:

$$\left[g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1-F(t)]^{n-k}=\left[g(t)+c\sum_{j=k}^{r-2}\frac{1}{n-j}\right][1-F(t)]^{n-k}+$$

$$\frac{(n-k)!}{(r-k-1)!\,(n-r+1)!}\int_t^{\beta} g'(y_r)\,[F(y_r)-F(t)]^{r-k-1}\,[1-F(y_r)]^{n-r+1}\,dy_r$$

$\Rightarrow$

$$\left[g(t)+c\sum_{j=k}^{r-2}\frac{1}{n-j}+\frac{c}{n-r+1}\right][1-F(t)]^{n-k}=\left[g(t)+c\sum_{j=k}^{r-2}\frac{1}{n-j}\right][1-F(t)]^{n-k}+$$

$$\frac{(n-k)!}{(r-k-1)!\,(n-r+1)!}\int_t^{\beta} g'(y_r)\,[F(y_r)-F(t)]^{r-k-1}\,[1-F(y_r)]^{n-r+1}\,dy_r$$

Por lo tanto tras distribuir y eliminar los sumandos que se repiten en ambos lados de la igualdad anterior obtenemos que la ecuación (3.13) es equivalente a:

$$\frac{c}{n-r+1}[1-F(t)]^{n-k}=D\int_t^{\beta} g'(y_r)\,[F(y_r)-F(t)]^{r-k-1}\,[1-F(y_r)]^{n-r+1}\,dy_r \qquad (3.14)$$

En donde

$$D=\frac{(n-k)!}{(r-k-1)!\,(n-r+1)!}$$

Nuestro siguiente paso será derivar respecto a $t$ ambos lados de la igualdad de la ecuación (3.14) obteniendo con ello lo siguiente:

[4] Recuerde que la variable $y_r$ es muda para las expresiones con integrales por lo que

$$\int_t^{\beta} H(y_r)\,dy_r=\int_t^{\beta} H(y_{r-1})\,dy_{r-1}$$

$$\frac{d}{dt}\left\{\frac{c\left[1-F\left(t\right)\right]^{n-k}}{n-r+1}\right\} = D\int_t^{\beta}\frac{d}{dt}g'\left(y_r\right)\left[F\left(y_r\right)-F\left(t\right)\right]^{r-k-1}\left[1-F\left(y_r\right)\right]^{n-r+1}dy_r{}^{5}$$

$$\Rightarrow \frac{c}{n-r+1}\left(n-k\right)\left[1-F\left(t\right)\right]^{n-k-1}\frac{d}{dt}\left(1-F\left(t\right)\right) = D\left(r-k-1\right)\frac{d}{dt}\left(1-F\left(t\right)\right)*$$

$$\int_t^{\beta} g'\left(y_{r:n}\right)\left[F\left(y_r\right)-F\left(t\right)\right]^{r-k-2}\left[1-F\left(y_r\right)\right]^{n-r+1}dy_r$$

Sustituyendo el valor de $D$ y eliminando términos que se anulan:

$$\Rightarrow \frac{c}{n-r+1}\left(n-k\right)\left[1-F\left(t\right)\right]^{n-k-1} = \frac{(n-k)!}{(r-k-2)!\,(n-r+1)!}*$$

$$\int_t^{\beta} g'\left(y_r\right)\left[F\left(y_r\right)-F\left(t\right)\right]^{r-k-2}\left[1-F\left(y_r\right)\right]^{n-r+1}dy_r$$

Entonces derivando $(r-k-1)$ veces mas respeto de $t$ llegamos a que la ecuación (3.13) es equivalente a lo siguiente:

$$\Rightarrow \frac{c\,(n-k)!\left[1-F\left(t\right)\right]^{n-r+1}}{(n-r+1)\,(n-r+1)!} = \frac{(n-k)!}{(n-r+1)!}\int_t^{\beta} g'\left(y_r\right)\left[1-F\left(y_r\right)\right]^{n-r+1}dy_r \qquad (3.15)$$

[5] Observe que en ésta última igualdad, la derivada está siendo introducida dentro del signo de integración lo cual es algo que no siempre podemos hacer. Para justificar este paso se citará un teorema del libro *"Advanced Calculus"* de Hans Sagan[10] página 426 y cuyo enunciado se escribe a continuación sin demostración:

**Teorema** *Sea $H:[a,b]\times[c,d]\to\mathbb{R}$, si $\gamma_1\gamma_2:[c,d]\to\mathbb{R}\in C^1[c,d]$, $a\leq\gamma_1(t)\leq\gamma_2(t)\leq b$ para todo $t\in[c,d]$, y si para cada $t$ fijo en $[c,d]$ $H$, vista como función de $x$, es "continua" en $[a,b]$, entonces la función $G:[c,d]\to\mathbb{R}$, definida como:*

$$G(t)=\int_{\gamma_1(t)}^{\gamma_2(t)}H(x,t)dx$$

*es diferenciable en $[c,d]$ y*

$$\frac{d}{dt}G(t)=\int_{\gamma_1(t)}^{\gamma_2(t)}\frac{\partial}{\partial t}H(x,t)dx+H(\gamma_2(t),t)\gamma_2'(t)-H(\gamma_1(t),t)\gamma_1'(t)$$

Lo que pide el teorema para poder introducir el símbolo de derivada dentro del de integral es que la función $H(x,t)$ sea continua como función de x para cada t fijo, lo cual en nuestro caso se tiene por la continuidad de F así como la de $g'(x)$ que forman parte de nuestras hipótesis.
Además debe darse cuenta que en nuestro caso

$$H(\gamma_2(t),t)\gamma_2'(t)-H(\gamma_1(t),t)\gamma_1'(t)=0$$

Por el hecho de que $H(t,t)=0$, y que $\gamma_2(t)$ es constante.

$$\Rightarrow \frac{c}{n-r+1}\left[1-F(t)\right]^{n-r+1} = \int_t^{\beta} g'(y_r)\left[1-F(y_{r:n})\right]^{n-r+1} dy_r \tag{3.16}$$

Derivando la ecuación anterior y usando el Teorema Fundamental del Cálculo

$$c\left[1-F(t)\right]^{n-r}\frac{d}{dt}(1-F(t)) \underbrace{=}_{T.F.C} -g'(t)\left[1-F(t)\right]^{n-r+1}$$

$$\Rightarrow \frac{d(1-F(t))}{1-F(t)} = -\frac{1}{c}g'(t)\,dt$$

$$\Rightarrow dLn(1-F(t)) = -\frac{1}{c}g'(t)\,dt$$

La igualdad anterior no es más que una ecuación diferencial para $F(x)$ con condición inicial $F(\alpha)=0$ la cual se resuelve integrando respecto a $t$ en el intervalo $(\alpha, x)$

$$\Rightarrow \int_\alpha^x dLn(1-F(t)) = \int_\alpha^x -\frac{1}{c}g'(t)\,dt$$

Recordando la hipótesis de que $F(\alpha)=0$

$$\Rightarrow Ln(1-F(t))|_\alpha^x = -\frac{1}{c}\left[g(x)-g(\alpha)\right]$$

$$\Rightarrow Ln(1-F(x)) = -\frac{1}{c}\left[g(x)-g(\alpha)\right]$$

$$\Rightarrow F(x) = 1-exp\left\{-\frac{1}{c}\left[g(x)-g(\alpha)\right]\right\}$$

Con lo cual terminaría la prueba del teorema. □

**Corolario 3.1.** *Suponiendo las mismas condiciones que el teorema (3.1) se tiene que:*

$$E\left[g(Y_{k+1:n})\,|\,Y_{k:n}=t\right] = g(t)+c\frac{1}{n-k} \Leftrightarrow F(x) = 1-exp\left\{-\frac{1}{c}\left[g(x)-g(\alpha)\right]\right\} \tag{3.17}$$

*Demostración.* La demostración se resuleve aplicando el teorema (3.1) con $r=k+1$ □

El resultado anterior fue probado por L.Y. Ouyang [6] en 1994 y posteriormente publicado en 1995. La demostración que ahí se expone es más sencilla ya que la esperanza condicional es tomada de estadísticos de orden adyacentes facilita mucho los cálculos al momento de evaluar las integrales involucradas.

**Teorema 3.2.** *Con las mismas condiciones que el teorema (3.1) se tiene que:*

$$E\left[\frac{1}{n-m+1}\sum_{r=m}^{n} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g\left(t\right)+\frac{c}{n-m+1}\sum_{r=m}^{n}\sum_{j=k}^{r-1}\frac{1}{n-j}\Leftrightarrow$$

$$F\left(x\right)=1-exp\left\{-\frac{1}{c}\left[g\left(x\right)-g\left(\alpha\right)\right]\right\} \tag{3.18}$$

*donde* $1\leq k<m\leq n$

*Demostración.* Nuevamente comenzaremos con la parte de la suficiencia.

- $\Leftarrow$:]

Esta parte de la prueba es sencilla ya que únicamente hacemos uso de las propiedades de linealidad del operador esperanza condicional así como del teorema (3.1) demostrado anteriormente.

$$E\left[\frac{1}{n-m+1}\sum_{r=m}^{n} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=\frac{1}{n-m+1}\sum_{r=m}^{n}E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]$$

$$\underbrace{=}_{\text{Teo 3.1}}\frac{1}{n-m+1}\sum_{r=m}^{n}\left\{g\left(t\right)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right\}=g\left(t\right)+\frac{c}{n-m+1}\sum_{r=m}^{n}\sum_{j=k}^{r-1}\frac{1}{n-j}$$

Con lo cual termina la primera parte de la prueba.

- $\Rightarrow$:]

Supongamos que:

$$E\left[\frac{1}{n-m+1}\sum_{r=m}^{n} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g\left(t\right)+\frac{c}{n-m+1}\sum_{r=m}^{n}\sum_{j=k}^{r-1}\frac{1}{n-j}$$

$$\Rightarrow\sum_{r=m}^{n}\underbrace{E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]}_{(3.1)}=\sum_{r=m}^{n}\left\{g\left(t\right)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right\} \tag{3.19}$$

Note ahora que usando la ecuación (3.1) del teorema anterior se tiene que:

$$\sum_{r=m}^{n}A_1\int_t^{\beta}g\left(y_r\right)\left\{\frac{\left[F\left(y_r\right)-F\left(t\right)\right]^{r-k-1}\left[1-F\left(y_r\right)\right]^{n-r}}{\left[1-F\left(t\right)\right]^{n-k}}\right\}dF\left(y_r\right)=\sum_{r=m}^{n}\left\{g\left(t\right)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right\}$$

$$\Rightarrow \sum_{r=m}^{n}\left\{g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right\}[1-F(t)]^{n-k}=\sum_{r=m}^{n}\frac{(n-k)!}{(r-k-1)!\,(n-r)!}*$$

$$\int_{t}^{\beta} g(y_r)\,[F(y_r)-F(t)]^{r-k-1}\,[1-F(y_r)]^{n-r}\,dF(y_r) \tag{3.20}$$

Observe que la ecuación anterior es igual a la ecuación (3.13) salvo el símbolo de suma que aparece en ambos lados. Y como la ecuación (3.13) es equivalente[6] a (3.15), entonces la anterior igualdad es equivalente a la ecuación (3.15) poniendo los respectivos símbolos de sumas en ambos lados quedando de la siguiente forma:

$$\sum_{r=m}^{n}\frac{c\,(n-k)!\,[1-F(t)]^{n-r+1}}{(n-r+1)\,(n-r+1)!}=\sum_{r=m}^{n}\frac{(n-k)!}{(n-r+1)!}\int_{t}^{\beta} g'(y_r)\,[1-F(y_r)]^{n-r+1}\,dy_r$$

Derivando respecto a $t$

$$\sum_{r=m}^{n}\frac{c\,(n-k)!\,[1-F(t)]^{n-r}}{(n-r+1)!}\frac{d}{dt}[1-F(t)]\underbrace{=}_{T.F.C}-\sum_{r=m}^{n}\frac{(n-k)!}{(n-r+1)!}g'(t)\,[1-F(t)]^{n-r+1}$$

$$\Rightarrow \sum_{r=m}^{n}\left\{\frac{c\,(n-k)!\,[1-F(t)]^{n-r}}{(n-r+1)!}\frac{d}{dt}[1-F(t)]+\frac{(n-k)!}{(n-r+1)!}g'(t)\,[1-F(t)]^{n-r+1}\right\}=0$$

$$\Rightarrow \sum_{r=m}^{n}\frac{(n-k)!\,[1-F(t)]^{n-r}}{(n-r+1)!}\left\{c\frac{d}{dt}[1-F(t)]+g'(t)\,[1-F(t)]\right\}=0$$

Como $[1-F(t)]>0 \quad \forall t\in(\alpha,\beta)$ , entonces $\sum_{r=m}^{n}\frac{(n-k)![1-F(t)]^{n-r}}{(n-r+1)!}>0 \quad \forall t\in(\alpha,\beta)$, de donde para que la igualdad anterior sea cierta es necesario que:

$$c\frac{d}{dt}[1-F(t)]+g'(t)\,[1-F(t)]=0\Leftrightarrow\frac{d\,[1-F(t)]}{1-F(t)}=-\frac{g'(t)}{c}dt$$

Lo cual implica, como ya se había probado, que:

$$F(x)=1-exp\left\{-\frac{1}{c}[g(x)-g(\alpha)]\right\}$$

Con lo que termina la prueba. □

Observe que el teorema anterior pide que $1\leq k<m\leq n$ por lo que no se aplica cuando $k=m$, así que para contemplar este caso particular presentaremos el siguiente corolario para la solución este problema.

[6] Para ver detalles, revisar la demostración del teorema (3.1)

**Corolario 3.2.** *Bajo la mismas condiciones que el teorema (3.1) se tiene que:*

$$E\left[\frac{1}{n-k+1}\sum_{r=k}^{n} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g\left(t\right)+\frac{n-k}{n-k+1}c \Leftrightarrow$$

$$F\left(x\right)=1-exp\left\{-\frac{1}{c}\left[g\left(x\right)-g\left(\alpha\right)\right]\right\} \tag{3.21}$$

*Demostración.* La prueba se realizará en ambos sentidos al mismo tiempo partiendo del teorema anterior haciendo $m=k+1$:

$$F\left(x\right)=1-exp\left\{-\frac{1}{c}\left[g\left(x\right)-g\left(\alpha\right)\right]\right\}$$

$$\Leftrightarrow E\left[\frac{1}{n-k}\sum_{r=k+1}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right]=g(t)+\frac{c}{n-k}\sum_{r=k+1}^{n}\sum_{j=k}^{r-1}\frac{1}{n-j}$$

Pero como[7]

$$\sum_{r=k+1}^{n}\sum_{j=k}^{r-1}\frac{1}{n-j}=n-k$$

Por lo tanto lo anterior pasa si solo si:

$$E\left[\frac{1}{n-k}\sum_{r=k+1}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right]=g(t)+c$$

$$\Leftrightarrow \frac{1}{n-k+1}\, E\left[\sum_{r=k+1}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right]=\frac{n-k}{n-k+1}g(t)+\frac{n-k}{n-k+1}c$$

[7] Esta igualdad puede verse de manera más fácil ordenando los sumandos de la siguiente forma:

$$\begin{array}{ccccccc}
\frac{1}{n-k} & & & & & & \\
\frac{1}{n-k} & + & \frac{1}{n-(k+1)} & & & & \\
\frac{1}{n-k} & + & \frac{1}{n-(k+1)} & + & \frac{1}{n-(k+2)} & & \\
\vdots & & \vdots & & \vdots & \ddots & \\
\frac{1}{n-k} & + & \frac{1}{n-(k+1)} & + & \frac{1}{n-(k+2)} & + \ \dots & 1 \\
\downarrow & & \downarrow & & \downarrow & \dots & \downarrow \\
1 & & 1 & & 1 & & 1
\end{array}$$

El cual es un arreglo de $(n-k)\times(n-k)$. De donde se deduce el resultado de la ecuación al sumar el último renglón formado por unos.

$$\Leftrightarrow \frac{1}{n-k+1}\, E\left[\sum_{r=k+1}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right] = \frac{n-k}{n-k+1}g(t) + \frac{n-k}{n-k+1}c$$

$$\Leftrightarrow \frac{g(t)}{n-k+1} + \frac{1}{n-k+1}\, E\left[\sum_{r=k+1}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right] = \underbrace{\frac{g(t)}{n-k+1} + \frac{(n-k)g(t)}{n-k+1}}_{g(t)} + \frac{(n-k)c}{n-k+1}$$

$$\Leftrightarrow \frac{g(t)}{n-k+1} + \frac{1}{n-k+1}\, E\left[\sum_{r=k+1}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right] = g(t) + \frac{n-k}{n-k+1}c$$

Recordando el hecho de que $E\left[g(Y_{k:n})\right| Y_{k:n}=t] = g(t)$

$$\Leftrightarrow E\left[\frac{1}{n-k+1}\sum_{r=k}^{n} g(Y_{r:n})\middle| Y_{k:n}=t\right] = g(t) + \frac{n-k}{n-k+1}c$$

Por lo que siguiendo todas las implicaciones hemos probado el corolario

□

En seguida se presenta otro teorema que de igual forma caracterizará a las funciones de la forma $F(x) = 1 - exp\left\{-\frac{1}{c}\left[g(x) - g(\alpha)\right]\right\}$ pero que sólo involucra al primer estadístico de orden en el cálculo de la esperanza condicional

**Teorema 3.3.** *Con las misma condiciones que el teorema (3.1) se tiene que:*

$$E\left[g(Y_{1:n})\right| Y_{1:n} > t] = g(t) + \frac{c}{n} \Leftrightarrow F(x) = 1 - exp\left\{-\frac{1}{c}\left[g(x) - g(\alpha)\right]\right\} \tag{3.22}$$

*Demostración.* Para empezar primero recordemos la densidad del primer estadístico de orden la cual ya fue presentada en el primer capítulo de este trabajo y puede ser consultada en la ecuación (1.4):

$$f_{Y_{1:n}}(y) = \frac{d}{dy}F_{Y_{1:n}}(y) = n\left[1 - F(y)\right]^{n-1} f(y)$$

Ahora recordemos que la esperanza condicional dado un evento $B$, la cual puede ser encontrada en el segundo capítulo en la ecuación (2.10), esta dada por:

$$E\left[X\right| B] = \frac{E[XI_B]}{P(B)} = \frac{\int_B xf(x)\,dx}{P(B)}$$

Entonces:

$$E\left[g\left(Y_{1:n}\right)\middle| Y_{1:n} > t\right] = \frac{\displaystyle\int_t^{\beta} g\left(y_1\right) n\left[1 - F\left(y_1\right)\right]^{n-1} dF\left(y_1\right)}{\displaystyle\int_t^{\beta} n\left[1 - F\left(y_1\right)\right]^{n-1} dF\left(y_1\right)}$$

Resolviendo la integral del denominador:

$$\int_t^{\beta} n\left[1 - F\left(y_1\right)\right]^{n-1} dF\left(y_1\right) = -n\frac{\left[1 - F\left(y_1\right)\right]^{n}}{n}\Bigg|_t^{\beta} = \left[1 - F\left(t\right)\right]^{n}$$

Por lo que sustituyendo lo anterior hemos llegado a que por definición se tiene:

$$E\left[g\left(Y_{1:n}\right)\middle| Y_{1:n} > t\right] = \frac{\displaystyle\int_t^{\beta} g\left(y_1\right) n\left[1 - F\left(y_1\right)\right]^{n-1} dF\left(y_1\right)}{\left[1 - F\left(t\right)\right]^{n}} \tag{3.23}$$

Teniendo lo anterior, comenzaremos la prueba con la parte de la suficiencia del teorema

- $\Leftarrow$:]

Supondremos que $F\left(x\right) = 1 - exp\left\{-\frac{1}{c}\left[g\left(x\right) - g\left(\alpha\right)\right]\right\}$ que al ser sustituido en (3.23) nos permite obtener:

$$E\left[g\left(Y_{1:n}\right)\middle| Y_{1:n} > t\right] = \frac{n}{\left[1 - F\left(t\right)\right]^{n}} \int_t^{\beta} g\left(y_1\right) exp\left\{-\frac{n-1}{c}\left[g\left(y_1\right) - g\left(\alpha\right)\right]\right\} *$$
$$exp\left\{-\frac{1}{c}\left[g\left(y_1\right) - g\left(\alpha\right)\right]\right\} \frac{1}{c} g'\left(y_1\right) dy_1$$

Entonces

$$E\left[g\left(Y_{1:n}\right)\middle| Y_{1:n} > t\right] = \frac{n}{\left[1 - F\left(t\right)\right]^{n}} \underbrace{\int_t^{\beta} g\left(y_1\right) exp\left\{-\frac{n}{c}\left[g\left(y_1\right) - g\left(\alpha\right)\right]\right\} \frac{1}{c} g'\left(y_1\right) dy_1}_{IV}$$

Resolvemos (IV) por el método de integración por partes

Sea

$u = g\left(y_1\right)$ y $v = exp\left\{-\frac{n}{c}\left[g\left(y_1\right) - g\left(\alpha\right)\right]\right\}$

$\Rightarrow du = g'\left(y_1\right) dy_1$

$\Rightarrow dv = -\frac{n}{c} exp\left\{-\frac{n}{c}\left[g\left(y_1\right) - g\left(\alpha\right)\right]\right\} g'\left(y_1\right) dy_1$

Entonces al resolver la integral la esperanza condicional buscada es igual a[8] :

$$= -\frac{1}{[1-F(t)]^n}\int_t^\beta u dv = -\frac{1}{[1-F(t)]^n}\left[-\int_t^\beta v du + uv\Big|_t^{\beta}\right]$$

$$= \frac{-1}{[1-F(t)]^n} *$$

$$\left[-\int_t^\beta g'(y_1)\, exp\left\{-\frac{n}{c}\left[g(y_1)-g(\alpha)\right]\right\}dy_1 + g(y_1)\, exp\left\{-\frac{n}{c}\left[g(y_1)-g(\alpha)\right]\right\}\Big|_t^{\beta}\right]$$

$$= \frac{1}{[1-F(t)]^n} g(t)\left[1-F(t)\right]^n + \frac{1}{[1-F(t)]^n}\underbrace{\int_t^\beta exp\left\{-\frac{n}{c}\left[g(y_1)-g(\alpha)\right]\right\} g'(y_1)\,dy_1}_{IV}$$

La integral (IV) se resuelve por el método de cambio de variable haciendo $u = -\frac{n}{c}\left[g(y_{1:n}) - g(\alpha)\right]$.

$$= g(t) - \frac{c}{n\left[1-F(t)\right]^n} exp\left\{-\frac{n}{c}\left[g(y_1)-g(\alpha)\right]\right\}\Big|_t^{\beta}$$

$$= g(t) + \frac{c}{n\left[1-F(t)\right]^n}\left[1-F(t)\right]^n = g(t) + \frac{c}{n}$$

Por lo que siguiendo todas la igualdades anteriores hemos probado que:

$$E\left[g(Y_{1:n}) | Y_{1:n} > t\right] = g(t) + \frac{c}{n}$$

Con lo que termina la primera parte de la prueba.

- $\Rightarrow$:]

Para probar esta parte , ahora supondremos que:

$$E\left[g(Y_{1:n}) | Y_{1:n} > t\right] = g(t) + \frac{c}{n}$$

Lo cual al ser sustituido en (3.23):

$$n\int_t^\beta g(y_1)\left[1-F(y_1)\right]^{n-1} dF(y_1) = \left[1-F(t)\right]^n\left[g(t) + \frac{c}{n}\right]$$

[8] observe que:

$$\lim_{y_1\to\beta} g(y_1) exp\left\{-\frac{n}{c}[g(y_1)-g(\alpha)]\right\} = 0$$

$$\Rightarrow \int_t^\beta g(y_1)\underbrace{n[1-F(y_1)]^{n-1}\,dF(y_1)}_{-d[1-F(y_1)]^n} = [1-F(t)]^n\left[g(t)+\frac{c}{n}\right]$$

$$\Rightarrow \underbrace{-\int_t^\beta g(y_1)\,d[1-F(y_1)]^n}_{V} = [1-F(t)]^n\left[g(t)+\frac{c}{n}\right] \tag{3.24}$$

Resolvemos (V) por partes:

Sea

$u = g(y_1)$ y $v = [1-F(y_1)]^n$

$\Rightarrow du = g'(y_1)\,dy_1$

$\Rightarrow dv = d[1-F(y_1)]^n$

Entonces:

$$(V) = -\int_t^\beta u\,dv = uv\Big|_\beta^t + \int_t^\beta [1-F(y_1)]^n g'(y_1)\,dy_1$$

$$(V) = g(t)[1-F(t)]^n + \int_t^\beta [1-F(y_1)]^n g'(y_1)\,dy_1$$

Entonces si sustituimos (V) en (3.24) obtenemos:

$$(V) = g(t)[1-F(t)]^n + \int_t^\beta [1-F(y_1)]^n g'(y_1)\,dy_1 = [1-F(t)]^n\left[g(t)+\frac{c}{n}\right]$$

$$\Rightarrow \int_t^\beta [1-F(y_1)]^n g'(y_1)\,dy_1 = \frac{c}{n}[1-F(t)]^n$$

Derivando respecto a $t$ y usando el Teorema Fundamental del Cálculo:

$$-[1-F(t)]^n g'(t) = \frac{c}{n}n[1-F(t)]^{n-1}\frac{d}{dt}[1-F(t)]$$

$$\Rightarrow \frac{d[1-F(t)]}{[1-F(t)]} = -\frac{1}{c}g'(t)\,dt$$

De donde se concluye que:

$$F(x) = 1 - exp\left\{-\frac{1}{c}[g(x)-g(\alpha)]\right\}$$

Con lo que termina la prueba del teorema. □

**Corolario 3.3.** *Si $F$ es de la forma:*

$$F(x) = 1 - exp\left\{-\frac{1}{c}\left[g(x) - g(\alpha)\right]\right\} \qquad x \in (\alpha, \beta)$$

*Entonces:*

$$E[g\left(Y_{1:n}\right)] = g(\alpha) + \frac{c}{n}$$

*Demostración.* Usando el teorema anterior tenemos:

$$E[g\left(Y_{1:n}\right) \mid Y_{1:n} > t] = g(t) + \frac{c}{n} \qquad \forall t \in (\alpha, \beta)$$

Recordando el resultado del teorema (2.5) probado en el capítulo anterior tenemos:

$$\lim_{t \to \alpha} E[g\left(Y_{1:n}\right) \mid Y_{1:n} > t] = E[g\left(Y_{1:n}\right)]$$

Entonces:

$$E[g\left(Y_{1:n}\right)] = \lim_{t \to \alpha} g(t) + \frac{c}{n} = g(\alpha) + \frac{c}{n}$$

□

**Resultado**

Sea g una función que cumple con las condiciones del teorema (3.1) y supongamos que F es dada por la expresión (3.33), entonces

$$E[g(Y_{r:n})] = E[g(Y_{k:n})] + c\sum_{j=k}^{r-1}\frac{1}{n-j} \qquad con \qquad r > k$$

*Demostración.* Aplicando el teorema (3.1) a la función de distribución F obtenemos que:

$$E[g(Y_{r:n}) |\, Y_{k:n} = t] = g(t) + c\sum_{j=k}^{r-1}\frac{1}{n-j}$$

De donde será claro que:

$$E[g(Y_{r:n}) |\, Y_{k:n}] = g(Y_{k:n}) + c\sum_{j=k}^{r-1}\frac{1}{n-j}$$

Finalmente tomando esperanzas y recordando que [9] : $E\left[E[g(X) \mid Y]\right] = E[g(X)]$ se tiene el resultado

$$E\left[E[g(Y_{r:n}) |\, Y_{k:n}]\right] = E\left[g(Y_{k:n})\right] + c\sum_{j=k}^{r-1}\frac{1}{n-j}$$

[9] La prueba de este resultado es omitida en este trabajo pero puede ser consultada en la página 86 proposición 4.13 de las notas de Miguel Ángel García Álvarez [4]

$$\Rightarrow E\left[g(Y_{r:n})\right] = E[g(Y_{k:n})] + c\sum_{j=k}^{r-1}\frac{1}{n-j} \qquad con \qquad r > k$$

Así pues debe darse cuenta que únicamente debemos obtener la esperanza de algún estadístico de orden para poder calcular todos los demás. □

En los teoremas y corolarios anteriores se supone que $r > k$, sin embargo, todos estos pueden ser nuevamente enunciados para el caso en que $r < k$. Estos teoremas se presentan a continuación con sus respectivas pruebas y variantes las cuales siguen la misma lógica de los teoremas anteriormente demostrados por lo que las demostraciones serán muy parecidas al seguir los mismos procedimientos de integración y razonamiento.

A continuación se presente el teorema análogo al (3.1) suponiendo ahora que $r < k$

**Teorema 3.4.** *Sea $c \in \mathbb{R} - \{0\}$ y consideremos $\alpha \geq -\infty$ y $\beta \leq \infty$ los extremos izquierdo y derecho de una función de distribución $F(\cdot)$ de una variable aletaroria absolutamente continua $X$, con $0 < F(x) < 1$ $\forall x \in (\alpha, \beta)$, y sea $g\colon (\alpha, \beta] \rightarrow \mathbb{R}$ una función continua sobre el intervalo $(\alpha, \beta]$ con primera derivada continua sobre $(\alpha, \beta)$ y tal que $g(x) \rightarrow^{+}_{-} \infty$ cuando $x \rightarrow \alpha^{+}$ si $c > 0$ o $c < 0$ respectivamente.*

*Si $E[g(X)] < \infty$, entonces:*

$$E\left[g\left(Y_{r:n}\right)|\, Y_{k:n} = t\right] = g(t) + c\sum_{j=r}^{k-1}\frac{1}{j} \Leftrightarrow F(x) = exp\left\{\frac{1}{c}\left[g(\beta) - g(x)\right]\right\}$$

*con $1 \leq r < k \leq n$ y $\forall t \in (\alpha, \beta)$*

Debe de notarse que ahora la función de distribución ha cambiado respecto a los teoremas anteriores en donde se pedía que $k$ fuera menor que $r$.

*Demostración.* La prueba será similar a la demostración del teorema (3.1), por lo que primero daremos una expresión para la esperanza condicional involucrada en el teorema.

Usando el resultado del teorema (2.4) que nos da la densidad condicional de $Y_{r:n} \mid Y_{k:n} = t$ con $r < k$ y por definición de esperanza condicional para el caso absolutamente continuo dado en (2.13) tenemos:

$$E\left[g\left(Y_{r:n}\right)|\, Y_{k:n} = t\right] = A_3\int_{\alpha}^{t} g(y_r)\, H(y_r)^{r-1}\left[1 - H(y_r)\right]^{k-r-1} dH(y_r)$$

Donde:

$$A_3 = \frac{(k-1)!}{(r-1)!\,(k-1-r)!}$$

$$H(x) = \begin{cases} \dfrac{F(x)}{F(t)} & x \leq t \\ \\ 1 & x > t \end{cases}$$

Por lo que al sustituir y simplificar:

$$E\left[g\left(Y_{r:n}\right) | Y_{k:n} = t\right] = A_3 \int_{\alpha}^{t} g\left(y_r\right) \frac{F\left(y_r\right)^{r-1}\left[F(t) - F\left(y_r\right)\right]^{k-r-1}}{F(t)^{k-1}} dF\left(y_r\right) \tag{3.25}$$

Comenzaremos la prueba con la parte de suficiencia.

- $\Leftarrow$:]

Como suponemos que

$$F(x) = exp\left\{\frac{1}{c}\left[g(\beta) - g(x)\right]\right\}$$

$$\Rightarrow dF\left(y_r\right) = -\frac{1}{c}g'\left(y_r\right) exp\left\{\frac{1}{c}\left[g(\beta) - g\left(y_r\right)\right]\right\} dy_r$$

Lo cual al ser sustituido en (3.25) se obtiene que la esperanza condicional $E\left[g\left(Y_{r:n}\right) | Y_{k:n} = t\right]$ es igual a:

$$= \frac{A_3}{F(t)^{k-1}} \underbrace{\int_{\alpha}^{t} -\frac{1}{c} g(y_r)\left[F(t) - F(y_r)\right]^{k-r-1} exp\left\{\frac{r}{c}\left[g(\beta) - g\left(y_r\right)\right]\right\} g'\left(y_r\right) dy_r}_{VI}$$

Resolvemos (VI) por partes:

Sea

$u = g\left(y_r\right)\left[F(t) - F\left(y_r\right)\right]^{k-r-1}$ y $v = exp\left\{\frac{r}{c}\left[g(\beta) - g\left(y_r\right)\right]\right\}$

Entonces

$du = -g\left(y_r\right)(k-r-1)\left[F(t) - F\left(y_r\right)\right]^{k-r-2} dF\left(y_r\right) + \left[F(t) - F\left(y_r\right)\right]^{k-r-1} g'\left(y_r\right) dy_r$

$dv = -\frac{r}{c} exp\left\{\frac{r}{c}\left[g(\beta) - g\left(y_r\right)\right]\right\} g'\left(y_r\right) dy_r$

Así que resolviendo la integral:

$$= \frac{A_3}{F(t)^{k-1} r} \int_{\alpha}^{t} u dv = \frac{A_3}{F(t)^{k-1} r}\left[-\int_{\alpha}^{t} v du + uv\Big|_{\alpha}^{t}\right]$$

Y como:

$$uv|_{\alpha}^{t} = g(y_r)\left[F(t) - F(y_r)\right]^{k-r-1} exp\left\{\frac{r}{c}\left[g(\beta) - g(y_r)\right]\right\}\Big|_{\alpha}^{t} = 0$$

Entonces:

$$E\left[g(Y_{r:n})|\,Y_{k:n} = t\right] = \underbrace{A_4 \int_{\alpha}^{t} g(y_r)\,\frac{F(y_r)^{(r+1)-1}\left[F(t) - F(y_r)\right]^{k-(r+1)-1}}{F(t)^{k-1}} dF(y_r)}_{= E\left[g(Y_{r+1})|\,Y_k = t\right]} -$$

$$\frac{(k-1)!}{(r)!\,(k-r-1)!}\int_{\alpha}^{t} \frac{g'(y_r)\left[F(t) - F(y_r)\right]^{k-r-1} exp\left\{\frac{r}{c}\left[g(\beta) - g(y_r)\right]\right\}}{F(t)^{k-1}} dy_r$$

Donde:

$$A_4 = \frac{(k-1)!}{(r)!\,(k-r-2)!}$$

Por lo que se obtiene que:

$$E\left[g(Y_{r:n})|\,Y_{k:n} = t\right] = E\left[g(Y_{r+1:n})|\,Y_{k:n} = t\right] -$$

$$\frac{(k-1)!}{(r)!\,(k-r-1)!F(t)^{k-1}}\underbrace{\int_{\alpha}^{t} g'(y_r)\left[F(t) - F(y_r)\right]^{k-r-1} exp\left\{\frac{r}{c}\left[g(\beta) - g(y_r)\right]\right\} dy_r}_{VII} \quad (3.26)$$

Resolvemos (VII) por el método de integración por partes: Sea

$u = \left[F(t) - F(y_r)\right]^{k-r-1}$ y $v = exp\left\{\frac{r}{c}\left[g(\beta) - g(y_r)\right]\right\} dy_r$

Entonces

$du = -(k-r-1)\left[F(t) - F(y_r)\right]^{k-r-2} dF(y_r)$

$dv = -\frac{r}{c} exp\left\{\frac{r}{c}\left[g(\beta) - g(y_r)\right]\right\} g'(y_r)\,dy_r$

Con lo que se llega a que:

$$(VII) = \frac{k-r-1}{r}\underbrace{\int_{\alpha}^{t} g'(y_r)\left[F(t) - F(y_r)\right]^{k-(r+1)-1} exp\left\{\frac{r+1}{c}\left[g(\beta) - g(y_r)\right]\right\} dy_r}_{VIII}$$

Note que como $y_r$ es variable muda dentro de la integral, entonces (VIII) es igual que (VII) sustituyendo $r$ por $r+1$. Por lo que podemos usar la ecuación anterior de manera iterativa logrando con ello integrar $(k-r-2)$ veces por partes, obteniendo que:

$$(VII) = \frac{(k-r-1)!}{r(r+1)\ldots(r+(k-r)-2)}\int_{\alpha}^{t} g'(y_r)\,exp\left\{\frac{r+(k-r)-1}{c}\left[g(\beta) - g(y_r)\right]\right\} dy_r$$

$$(VII) = \frac{(k-r-1)!\,(r-1)!}{(k-2)!}\int_{\alpha}^{t} g'(y_r)\, exp\left\{\frac{k-1}{c}\left[g(\beta)-g(y_r)\right]\right\} dy_r$$

Sustituyendo (VII) en (3.26) obtenemos:

$$E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = E\left[g(Y_{r+1:n})|\, Y_{k:n}=t\right] -$$

$$\frac{(k-1)!}{(r)!\,(k-r-1)!F(t)^{k-1}}\frac{(k-r-1)!\,(r-1)!}{(k-2)!}\int_{\alpha}^{t} g'(y_r)\, exp\left\{\frac{k-1}{c}\left[g(\beta)-g(y_r)\right]\right\} dy_r$$

Lo cual tras simplificar y resolver la integral involucrada.

$$E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = E\left[g(Y_{r+1:n})|\, Y_{k:n}=t\right] + \frac{c}{rF(t)^{k-1}} exp\left\{\frac{k-1}{c}\left[g(\beta)-g(y_r)\right]\right\}\Bigg|_{\alpha}^{t}$$

$$\Rightarrow E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = E\left[g(Y_{r+1:n})|\, Y_{k:n}=t\right] + \frac{c}{rF(t)^{k-1}}F(t)^{k-1}$$

Por lo tanto:

$$E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = E\left[g(Y_{r+1:n})|\, Y_{k:n}=t\right] + \frac{c}{r} \tag{3.27}$$

Entonces usando la ecuación (3.27) recursivamente obtenemos:

$$E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = \frac{c}{r} + \frac{c}{r+1} + \ldots + \frac{c}{k-1} + \underbrace{E\left[g(Y_{k:n})|\, Y_{k:n}=t\right]}_{g(t)}$$

$$\Rightarrow \qquad E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = g(t) + c\sum_{j=r}^{k-1}\frac{1}{j}$$

Con lo que termina la primera parte del teorema.

- $\Leftarrow$:]

Utilizando la hipótesis de que

$$E\left[g(Y_{r:n})|\, Y_{k:n}=t\right] = g(t) + c\sum_{j=r}^{k-1}\frac{1}{j}$$

Lo cual al ser sustituido en (3.25) obtenemos:

$$\underbrace{A_3\int_{\alpha}^{t} g(y_r)\, F(y_r)^{r-1}\left[F(t)-F(y_r)\right]^{k-r-1} dF(y_r)}_{IX} = \left[g(t) + c\sum_{j=r}^{k-1}\frac{1}{j}\right] F(t)^{k-1} \tag{3.28}$$

Resolviendo (IX) por partes:

Sea

$u = g(y_r)\left[F(t) - F(y_r)\right]^{k-r-1} \qquad \text{y} \qquad v = F(y_r)^r$

Entonces[10]

$du = -g(y_r)(k-r-1)\left[F(t) - F(y_r)\right]^{k-r-2} dF(y_r) + \left[F(t) - F(y_r)\right]^{k-r-1} g'(y_r)\, dy_r$

$dv = rF(y_r)^{r-1} dF(y_r)$

$$\Rightarrow (IX) = \frac{k-r-1}{r}\int_\alpha^t g(y_r)\left[F(t) - F(y_r)\right]^{k-r-2} F(y_r)^r\, dF(y_r) - \frac{1}{r}\int_\alpha^t F(y_r)^r \left[F(t) - F(y_r)\right]^{k-r-1} g'(y_r)\, dy_r$$

Sustituyendo (IX) en (3.28)

$$\underbrace{\frac{(k-1)!}{r!(k-r-2)!}\int_\alpha^t g(y_r)\, F(y_r)^{(r+1)-1}\left[F(t) - F(y_r)\right]^{k-(r+1)-1} dF(y_r)}_{X} - \frac{(k-1)!}{r!(k-r-1)!}*$$

$$\int_\alpha^t F(y_r)^r\left[F(t) - F(y_r)\right]^{k-r-1} g'(y_r)\, dy_r = \left[g(t) + c\sum_{j=r}^{k-1}\frac{1}{j}\right] F(t)^{k-1}$$

Observe que la expresión (X) es igual al lado izquierdo de la ecuación (3.28) sustituyendo $r$ por $r+1$. Por lo tanto usando la igualdad de (3.28) podemos concluir que:

$$\left[g(t) + c\sum_{j=r+1}^{k-1}\frac{1}{j}\right] F(t)^{k-1} - \frac{(k-1)!}{r!(k-r-1)!}\int_\alpha^t F(y_r)^r\left[F(t) - F(y_r)\right]^{k-r-1} g'(y_r)\, dy_r =$$

$$\left[g(t) + c\sum_{j=r}^{k-1}\frac{1}{j}\right] F(t)^{k-1}$$

Por lo que tras eliminar términos semejantes obtenemos:

$$A_5\int_\alpha^t F(y_r)^r\left[F(t) - F(y_r)\right]^{k-r-1} g'(y_r)\, dy_r = \frac{c}{r}F(t)^{k-1} \tag{3.29}$$

[10] Note que en esta parte hacemos uso del resultado que nos garantiza que $\lim_{x\to\alpha} g(x)F(x)^n = 0$ el cual es consecuencia del teorema (2.9) del capítulo 2.

Donde

$$A_5 = -\frac{(k-1)!}{r!(k-r-1)!}$$

Derivamos la expresión (3.29) respecto a t (Recuerde en esta parte que bajo las hipótesis del teorema la derivada respecto a t puede introducirse dentro del símbolo de integración).

$$\Rightarrow A_5 \int_{\alpha}^{t} \frac{d}{dt} F(y_r)^r \left[F(t) - F(y_r)\right]^{k-r-1} g'(y_r)\, dy_r = \frac{d}{dt}\frac{c}{r}F(t)^{k-1}$$

De donde se obtiene que:

$$A_5(k-r-1) \int_{\alpha}^{t} F(y_r)^r \left[F(t) - F(y_r)\right]^{k-r-2} g'(y_r)\, dy_r = \frac{c}{r}(k-1)F(t)^{k-2}$$

Derivando $(k-r-2)$ veces más:

$$A_5(k-r-1)! \int_{\alpha}^{t} F(y_r)^r g'(y_r)\, dy_r = \frac{c}{r}\underbrace{(k-1)(k-2)\ldots(k-(k-r-1))}_{r+1} F(t)^{k-(k-r)}$$

Por lo que hemos probado que la ecuación (3.28) es equivalente a lo siguiente:

$$A_5(k-r-1)! \int_{\alpha}^{t} F(y_r)^r g'(y_r)\, dy_r = \frac{c}{r}\frac{(k-1)!}{r!}F(t)^r \qquad (3.30)$$

$$\Rightarrow A_5 \frac{(k-r-1)!r!}{(k-1)!} \int_{\alpha}^{t} F(y_r)^r g'(y_r)\, dy_r = \frac{c}{r}F(t)^r$$

Pero al sustituir el valor de $A_5$ se obtiene que:

$$A_5 \frac{(k-r-1)!r!}{(k-1)!} = -\frac{(k-1)!}{r!(k-r-1)!}\frac{(k-r-1)!r!}{(k-1)!} = -1$$

Por lo tanto:

$$-\int_{\alpha}^{t} F(y_r)^r g'(y_r)\, dy_r = \frac{c}{r}F(t)^r \qquad (3.31)$$

Derivando nuevamente a (3.31) respecto a $t$ y usando el T.F.C.

$$\Rightarrow -F(t)^r g'(t) = \frac{c}{r} r F(t)^{r-1}\frac{d}{dt}F(t)$$

$$\Rightarrow -\frac{1}{c}g'(t)dt = \frac{dF(t)}{F(t)} = dLn(F(t))$$

Lo anterior es una ecuación diferencial para $F$ con valor inicial $F(\beta) = 1$, la cual es resuelta integrando ambos lados de la ecuación respecto a $t$ en el intervalo $(x, \beta]$ dando como resultado:

$$F(x) = exp\left\{\frac{1}{c}\left[g(\beta) - g(x)\right]\right\}$$

Lo cual es lo que se quería probar.

$\square$

**Teorema 3.5.** *Bajo las mismas condiciones del teorema (3.4) se tiene que:*

$$E\left[\frac{1}{m}\sum_{r=1}^{m} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{c}{m}\sum_{r=1}^{m}\sum_{j=r}^{k-1}\frac{1}{j} \Leftrightarrow F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$$

*Donde* $1 \leq m < k \leq n$

*Demostración.* Comenzaremos la prueba con la parte de suficiencia.

- $\Leftarrow$:]

Utilizando el teorema (3.4) así como las propiedades de linealidad del operador esperanza condicional:

$$E\left[\frac{1}{m}\sum_{r=1}^{m} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=\frac{1}{m}\sum_{r=1}^{m} E\left[g\left(Y_{r:n}\right)\right| Y_{k:n}=t] \underbrace{=}_{3.4} \frac{1}{m}\sum_{r=1}^{m}\left\{g(t)+c\sum_{j=r}^{k-1}\frac{1}{j}\right\}$$

De donde tras distribuir la suma se tiene el resultado:

$$E\left[\frac{1}{m}\sum_{r=1}^{m} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{c}{m}\sum_{r=1}^{m}\sum_{j=r}^{k-1}\frac{1}{j}$$

- $\Rightarrow$:]

Por hipótesis:

$$E\left[\frac{1}{m}\sum_{r=1}^{m} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{c}{m}\sum_{r=1}^{m}\sum_{j=r}^{k-1}\frac{1}{j}$$

$$\Rightarrow \frac{1}{m}\sum_{r=1}^{m} \underbrace{E\left[g\left(Y_{r:n}\right)\right| Y_{k:n}=t]}_{XI} = \frac{1}{m}\sum_{r=1}^{m}\left\{g(t)+c\sum_{j=r}^{k-1}\frac{1}{j}\right\}$$

Note que (XI) está dado por la ecuación (3.25)

$$\sum_{r=1}^{m} A_3\int_{\alpha}^{t} g\left(y_r\right)\frac{F\left(y_r\right)^{r-1}\left[F(t)-F\left(y_r\right)\right]^{k-r-1}}{F(t)^{k-1}}dF\left(y_r\right)=\sum_{r=1}^{m}\left\{g(t)+c\sum_{j=r}^{k-1}\frac{1}{j}\right\}$$

Entonces:

$$\sum_{r=1}^{m} A_3 \int_{\alpha}^{t} g(y_r) F(y_r)^{r-1} [F(t) - F(y_r)]^{k-r-1} dF(y_r) = \sum_{r=1}^{m} \left[ g(t) + c \sum_{j=r}^{k-1} \frac{1}{j} \right] F(t)^{k-1}$$

Observe que la ecuación anterior es igual a la ecuación dada por (3.28) del teorema (3.4) si a esta última le agregáramos el símbolo de $\sum_{r=1}^{m}$ en ambos lados. Y como la ecuación (3.28) es equivalente a la ecuación (3.30) entonces:

$$\sum_{r=1}^{m} A_5 (k-r-1)! \int_{\alpha}^{t} F(y_r)^r g'(y_r) \, dy_r = \sum_{r=1}^{m} \frac{c}{r} \frac{(k-1)!}{r!} F(t)^r$$

Lo que implica tras sustituir el valor de $A_5$ que:

$$\sum_{r=1}^{m} -\frac{(k-1)!}{r!} \int_{\alpha}^{t} F(y_r)^r g'(y_r) \, dy_r = \sum_{r=1}^{m} \frac{c}{r} \frac{(k-1)!}{r!} F(t)^r$$

Derivando respecto a **t** y usando el T.F.C.

$$\sum_{r=1}^{m} -\frac{(k-1)!}{r!} F(t)^r g'(t) = \sum_{r=1}^{m} \frac{c}{r} \frac{(k-1)!}{r!} r F(t)^{r-1} \frac{d}{dt} F(t)$$

$$\Rightarrow \sum_{r=1}^{m} \left\{ c \frac{(k-1)!}{r!} F(t)^{r-1} \frac{d}{dt} F(t) + \frac{(k-1)!}{r!} F(t)^r g'(t) \right\} = 0$$

$$\Rightarrow \sum_{r=1}^{m} \underbrace{\frac{(k-1)!}{r!} F(t)^{r-1}}_{>0} \left\{ c \left( \frac{d}{dt} F(t) \right) + F(t) g'(t) \right\} = 0$$

Como $\frac{(k-1)!}{r!} F(t)^{r-1} > 0 \quad \forall \quad t \in (\alpha, \beta)$. Entonces para que la igualdad anterior sea cierta es necesario que:

$$c \left( \frac{d}{dt} F(t) \right) + F(t) g'(t) = 0$$

Lo anterior no es más que nuevamente una ecuación diferencial para F con problema inicial $F(\beta) = 1$, el cual se resuelve fácilmente llegando a que:

$$F(x) = exp \left\{ \frac{1}{c} [g(\beta) - g(x)] \right\}$$

Lo cual termina la prueba. □

Observe que el teorema anterior pide que $1 \leq m < k \leq n$ por lo que no se aplica cuando $k = m$, así que haremos un resultado especial para este caso en el siguiente corolario, el cual enunciamos a continuación

**Corolario 3.4.** *Bajo las misma condiciones del teorema (3.4) se tiene que:*

$$E\left[\frac{1}{k}\sum_{r=1}^{k} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{k-1}{k}c \Leftrightarrow F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$$

*Demostración.* La prueba se hará en ambos sentidos simultáneamente:

Partiremos del teorema (3.5) tomando $m=k-1$

$$\underbrace{E\left[\frac{1}{k-1}\sum_{r=1}^{k-1} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{c}{k-1}\sum_{r=1}^{k-1}\sum_{j=r}^{k-1}\frac{1}{j}}_{XII} \Leftrightarrow F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$$

Como:

$$(XII) \Leftrightarrow \frac{k}{k-1}E\left[\frac{1}{k}\sum_{r=1}^{k-1} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{c}{k-1}\sum_{r=1}^{k-1}\sum_{j=r}^{k-1}\frac{1}{j}$$

$$(XII) \Leftrightarrow \frac{k}{k-1}E\left[\frac{1}{k}\sum_{r=1}^{k} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]-\frac{1}{k-1}g(t)=g(t)+\frac{c}{k-1}\sum_{r=1}^{k-1}\sum_{j=r}^{k-1}\frac{1}{j}$$

Como

$$\sum_{r=1}^{k-1}\sum_{j=r}^{k-1}\frac{1}{j}=k-1$$

Lo cual se puede ver acomodando la suma como en el siguiente arreglo:

$$\underbrace{\begin{array}{ccccccccc}
\frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} & + & \dots & + & \frac{1}{k-1}\\
\downarrow & + & \frac{1}{2} & + & \frac{1}{3} & + & \dots & + & \frac{1}{k-1}\\
1 & & \downarrow & + & \frac{1}{3} & + & \dots & + & \frac{1}{k-1}\\
 & & 1 & & \downarrow & \ddots & \dots & + & \frac{1}{k-1}\\
 & & & & 1 & & \ddots & & \vdots\\
 & & & & & & & & \frac{1}{k-1}\\
 & & & & & & & & \downarrow\\
 & & & & & & & & 1
\end{array}}_{k-1}$$

Sustituyendo el valor de la doble suma:

$$(XII) \Leftrightarrow \frac{k}{k-1}E\left[\frac{1}{k}\sum_{r=1}^{k} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{1}{k-1}g(t)+\frac{c}{k-1}(k-1)$$

$$(XII) \Leftrightarrow E\left[\frac{1}{k}\sum_{r=1}^{k} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{k-1}{k}c$$

Por lo que siguiendo las implicaciones se tiene el resultado:

$$E\left[\frac{1}{k}\sum_{r=1}^{k} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{k-1}{k}c \Leftrightarrow F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$$

□

**Teorema 3.6.** *Bajo las mismas condiciones del teorema 3.4 se tiene que:*

$$E\left[g\left(Y_{n:n}\right)\middle| Y_{n:n}<t\right]=g(t)+\frac{c}{n} \Leftrightarrow F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$$

*Demostración.* Antes de comenzar cualquier implicación, daremos una expresión que nos ayude para la Esperanza condicional involucrada.

Recordando la densidad del máximo estadístico de orden así como por definición de esperanza condicional dado un evento se tiene que:

$$E\left[g\left(Y_{n:n}\right)\middle| Y_{n:n}<t\right]=\frac{\displaystyle\int_{\alpha}^{t} g\left(y_n\right) n!F\left(y_n\right)^{n-1} dF\left(y_n\right)}{P\left[Y_n \leq t\right]}$$

Y como:

$$P\left[Y_{n:n}<t\right]=\int_{\alpha}^{t} n!F\left(y_n\right)^{n-1} dF\left(y_n\right)=(n-1)!F(t)^n$$

Entonces tenemos que la esperanza se expresa como:

$$E\left[g\left(Y_{n:n}\right)\middle| Y_{n:n}<t\right]=\frac{n\displaystyle\int_{\alpha}^{t} g\left(y_n\right) F\left(y_n\right)^{n-1} dF\left(y_n\right)}{F(t)^n} \tag{3.32}$$

- $\Leftarrow$:]

Por hipótesis se tiene que $F(x)=exp\left\{\frac{1}{c}\left[g(\beta)-g(x)\right]\right\}$ lo que si es sustituido en la ecuación (3.32) se tiene:

$$E\left[g\left(Y_{n:n}\right)\middle| Y_{n:n}<t\right]=\frac{n}{F(t)^n}\int_{\alpha}^{t} -\frac{1}{c} g\left(y_n\right) exp\left\{\frac{n}{c}\left[g(\beta)-g\left(y_n\right)\right]\right\} g^{'}\left(y_n\right) dy_n$$

Integrando por partes:

Sea

$u = g(y_n)$ y $v = exp\left\{\frac{n}{c}\left[g(\beta) - g(y_r)\right]\right\}$

$\Rightarrow du = g'(y_n)\, dy_n$

$\Rightarrow dv = -\frac{n}{c} exp\left\{\frac{n}{c}\left[g(\beta) - g(y_n)\right]\right\} g'(y_n)$

Entonces la esperanza condicional se expresa como:

$$\frac{1}{F(t)^n}\left[-\int_\alpha^t g'(y_n)\, exp\left\{\frac{n}{c}\left[g(\beta) - g(y_n)\right]\right\} dy_n + \underbrace{g(y_n)\, exp\left\{\frac{n}{c}\left[g(\beta) - g(y_n)\right]\right\}\Big|_\alpha^t}_{g(t)F(t)^n}\right]$$

Entonces:

$$E\left[g(Y_{n:n})\,|\,Y_{n:n} < t\right] = g(t) - \frac{1}{F(t)^n}\underbrace{\int_\alpha^t g'(y_n)\, exp\left\{\frac{n}{c}\left[g(\beta) - g(y_n)\right]\right\} dy_n}_{XIII}$$

Resolviendo (XIII) por el método de cambio de variable:

$$= g(t) + \frac{c}{nF(t)^n}\left[exp\left\{\frac{n}{c}\left[g(\beta) - g(y_n)\right]\right\}\Big|_\alpha^t\right] = g(t) + \frac{c}{n}$$

De donde es obtiene el resultado:

$$E\left[g(Y_{n:n})\,|\,Y_{n:n} < t\right] = g(t) + \frac{c}{n}$$

- $\Rightarrow$:]

Por hipótesis y por la ecuación (3.32) tenemos:

$$\frac{n\int_\alpha^t g(y_n)\, F(y_n)^{n-1}\, dF(y_n)}{F(t)^n} = g(t) + \frac{c}{n}$$

$$\Rightarrow \underbrace{\int_\alpha^t g(y_n)\, dF(y_n)^n}_{XIV} = \left[g(t) + \frac{c}{n}\right] F(t)^n$$

Resolvemos a (XIV) por el método de integración por partes:

Sea

$u = g\left(y_n\right)$ y $v = F\left(y_n\right)^n$

$\Rightarrow du = g'\left(y_n\right) dy_n$

$\Rightarrow dv = dF\left(y_n\right)^n$

Llevando a cabo las sustituciones:

$$\Rightarrow -\int_{\alpha}^{t} F\left(y_n\right)^n g'\left(y_n\right) dy_n + g\left(y_n\right) F\left(y_n\right)^n \Big|_{\alpha}^{t} = \left[g(t) + \frac{c}{n}\right] F(t)^n$$

Y como: (Ver teoremas (2.8),(2.9) del capítulo 2)

$$\lim_{y_n \to \alpha} g(y_n)F(y_n)^n = 0$$

$$\Rightarrow \int_{\alpha}^{t} F\left(y_n\right)^n g'\left(y_n\right) dy_n = -\frac{c}{n}F(t)^n$$

Derivando respecto a $t$ y usando el T.F.C.

$$F(t)^n g'(t) = -cF(t)^{n-1}\frac{d}{dt}F(t)$$

$$\Rightarrow -\frac{1}{c}g'(t)dt = \frac{dF(t)}{F(t)} = dLn(F(t))$$

Lo anterior es una ecuación diferencial para $F$ con valor inicial $F(\beta) = 1$, la cual se resuelve integrando ambos lados obteniendo con ello que:

$$F(x) = exp\left\{\frac{1}{c}\left[g(\beta) - g(x)\right]\right\}$$

Con lo cual termina la prueba. □

**Corolario 3.5.** *Si $F$ es de la forma:*

$$F(x) = exp\left\{\frac{1}{c}\left[g(\beta) - g(x)\right]\right\} \qquad x \in (\alpha, \beta)$$

*Entonces:*

$$E[g\left(Y_{n:n}\right)] = g(\beta) + \frac{c}{n}$$

*Demostración.* Usando el teorema anterior tenemos:

$$E[g(Y_{n:n}) \mid Y_{n:n} < t] = g(t) + \frac{c}{n} \qquad \forall t \in (\alpha, \beta)$$

Recordando el resultado del teorema (2.5) probado en el capítulo anterior tenemos:

$$\lim_{t \to \beta} E[g(Y_{n:n}) \mid Y_{n:n} < t] = E[g(Y_{n:n})]$$

Entonces:

$$E[g(Y_{n:n})] = \lim_{t \to \beta} g(t) + \frac{c}{n} = g(\beta) + \frac{c}{n}$$

□

## 3.2. Algunos ejemplos y aplicaciones

Es finalmente en esta sección en donde se resuelve de manera sencilla el problema principal que planteó el presente trabajo y el cual consistía en encontrar un método que construyera una función de distribución una vez que la esperanza condicional de los estadísticos de orden es conocida. Mas aún, se puede ver ahora que poniendo restricciones adicionales al dominio de la variable aleatoria, podremos afirmar que la distribución que estamos construyendo es la única que cumple con esa esperanza condicional.

En la parte anterior se plantearon teoremas que caracterizaban a dos tipos de distribución:

$$F(x) = 1 - exp\left\{-\frac{1}{c}\left[g(x) - g(\alpha)\right]\right\} \tag{3.33}$$

$$F(x) = exp\left\{\frac{1}{c}\left[g(\beta) - g(x)\right]\right\} \tag{3.34}$$

Así cuando tenemos una función de distribución de la forma (3.33) y le aplicamos los teoremas (3.1),(3.2), (3.3) y el corolario (3.2) sabríamos calcular inmediatamente las siguientes esperanzas condicionales:

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g(t) + c\sum_{j=k}^{r-1}\frac{1}{n-j} \qquad con \qquad r > k$$

$$E\left[\frac{1}{n-m+1}\sum_{r=m}^{n} g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g(t) + \frac{c}{n-m+1}\sum_{r=m}^{n}\sum_{j=k}^{r-1}\frac{1}{n-j} \qquad con \qquad r > k$$

$$E\left[\frac{1}{n-k+1}\sum_{r=k}^{n} g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g(t) + \frac{n-k}{n-k+1}c \qquad con \qquad r > k$$

$$E\left[g\left(Y_{1:n}\right)\middle| Y_{1:n} > t\right] = g(t) + \frac{c}{n}$$

Las igualdades anteriores son válidas para todo t en el intervalo $(\alpha, \beta)$

Por otro lado, si la función de distribución está definida como en (3.34) y si aplicamos los teoremas (3.4),(3.5),(3.6) y el corolario (3.4) sabríamos entonces que:

$$E\left[g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g(t) + c\sum_{j=r}^{k-1}\frac{1}{j} \qquad con \qquad r < k \qquad \forall t \in (\alpha, \beta)$$

$$E\left[\frac{1}{m}\sum_{r=1}^{m} g\left(Y_{r:n}\right)\middle| Y_{k:n} = t\right] = g(t) + \frac{c}{m}\sum_{r=1}^{m}\sum_{j=r}^{k-1}\frac{1}{j} \qquad con \qquad r < k \qquad \forall t \in (\alpha, \beta)$$

$$E\left[\frac{1}{k}\sum_{r=1}^{k} g\left(Y_{r:n}\right)\middle| Y_{k:n}=t\right]=g(t)+\frac{k-1}{k}c \qquad con \qquad r<k \qquad \forall t\in(\alpha,\beta)$$

$$E\left[g\left(Y_{n:n}\right)|\, Y_{n:n}<t\right]=g(t)+\frac{c}{n} \qquad \forall t\in(\alpha,\beta)$$

Sin embargo debe ser claro que aunque el teorema nos ayuda a encontrar la esperanza condicional de estadísticos de orden una vez que la función de distribución es conocida, la parte importante de todos estos teoremas y corolarios es la otra implicación en donde podremos decir de qué función de distribución provienen los estadísticos de orden una vez que es conocida la esperanza condicional.

Para aplicar estos teoremas es necesario definir antes que nada cuál será la forma que tomará $g$ y bajo qué dominio estará definida de tal manera que cumpla las condiciones de los teoremas.

A continuación se darán unos ejemplos en donde se da la expresión para la función $g$ en el caso en que F es de la forma (3.33), para que posteriormente se trate el caso en donde F es de la forma (3.34).

### 3.2.1. Caracterización para las funciones de la forma $F(x)=1-exp\left\{-\frac{1}{c}[g(x)-g(\alpha)]\right\}$

Antes que nada es de destacar que *para cualquier función de distribución absolutamente continua* es posible encontrar una función $g$ que cumpla con todas las hipótesis del teorema y tal que al ser sustituida en (3.33) se obtenga la distribución deseada.

Para lograr lo anterior, sólo hay que notar que si $H(x)$ es la función de distribución con extremos $(\alpha,\beta)$ que deseamos obtener, entonces basta definir $g$ como [11] :

$$g(x)=-Ln(1-H(x)) \qquad x\in[\alpha,\beta) \tag{3.35}$$

[11] Observe que g definida de esta manera cumple con las condiciones del teorema:

1. $g:[\alpha,\beta)\rightarrow\mathbb{R}$ es continua en $[\alpha,\beta)$
2. $g:[\alpha,\beta)\rightarrow\mathbb{R}$ tiene primera derivada continua en $(\alpha,\beta)$
3. $\lim_{x\to\beta} g(x)=+\infty$
4. $E(g(X))$ es finita.

   La justificación de lo anterior es la siguiente:

   $$E[g(X)]=\int_\alpha^\beta g(x)H_X'(x)dx$$

   Pero por (3.35)

   $$E[g(X)]=\int_\alpha^\beta ln(1-H(x))H'(x))dx=\int_0^1 -ln(u)du=1$$

Para que al ser sustituida en (3.33) obtengamos:

$$F(x) = 1 - (1 - H(x))^{\frac{1}{c}} \qquad x \in (\alpha, \beta), c > 0$$

De donde si tomamos $c = 1$ obtenemos la distribución H(x).

Enseguida se aplicarán los teoremas a las funciones de distribución conocidas como la Uniforme, Pareto y Weibull haciendo las sustituciones pertinentes.

**Ejemplo 3.1 (Caracterización de la distribución Uniforme(0,1)).** Sea g definida como:

$$g(x) = -ln(1 - x^a) \qquad x \in [0, 1) \qquad a > 0$$

Entonces al sustituir en (3.33) obtenemos:

$$F(x) = 1 - (1 - x^a)^{1/c} \qquad x \in (0, 1) \qquad a > 0 \qquad c > 0 \tag{3.36}$$

De donde ahora ya podremos aplicar los teoremas (3.1),(3.2),(3.3) así como el corolario (3.2).

En particular si $a = 1$ y $c = 1$ en (3.36) entonces el teorema (3.1) caracteriza a la función de distribución Uniforme(0,1) de la siguiente manera:

$$E\left[-ln(1 - Y_{r:n}) | Y_{k:n} = t\right] = -ln(1 - t) + \sum_{j=k}^{r-1} \frac{1}{n - j} \Leftrightarrow X \sim Uniforme(0, 1) \qquad r > k$$

Lo anterior nos está diciendo que la distribución Uniforme(0,1) es la única distribución continua con extremos (0,1) que cumple con:

$$E\left[-ln(1 - Y_{r:n}) | Y_{k:n} = t\right] = -ln(1 - t) + \sum_{j=k}^{r-1} \frac{1}{n - j} \qquad \forall t \in (0, 1) \qquad r > k$$

**Ejemplo 3.2 (Caracterización de la distribución $Pareto(\alpha, b)$).** Sea g definida como:

$$g(x) = ln(x) \qquad x \in [\alpha, \infty) \qquad \alpha > 0$$

Entonces al sustituir en (3.33) obtenemos:

$$F(x) = 1 - \left(\frac{\alpha}{x}\right)^{1/c} \qquad x \in (\alpha, \infty) \qquad \alpha > 0 \qquad c > 0 \tag{3.37}$$

De donde ahora ya podremos aplicar nuevamente los teoremas (3.1),(3.2),(3.3) así como el corolario (3.2) a esta función de distribución.

Observe que si hacemos $b = 1/c$, entonces los teoremas estarán caracterizando a la función de distribución Pareto de parámetros $(\alpha, b)$. En particular el teorema (3.1) nos diría que:

$$E\left[ln(Y_{r:n}) | Y_{k:n} = t\right] = ln(t) + c\sum_{j=k}^{r-1} \frac{1}{n - j} \Leftrightarrow X \sim Pareto\left(\alpha, \frac{1}{c}\right) \qquad r > k$$

Esto quiere decir que si tenemos una esperanza condicional que pueda ser llevada a la forma:

$$E\left[ln(Y_{r:n})|\,Y_{k:n}=t\right]=ln(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\qquad \forall t\in(\alpha,\infty)\qquad r>k$$

Entonces la *única* distribución continua tal que $F(t)>0$ para todo $t\in(\alpha,\infty)$ que cumple con esta esperanza condicional es la distribución $Pareto(\alpha,\frac{1}{c})$

**Ejemplo 3.3 (Caracterización de la distribución $Weibull(\alpha,c,a)$).** Sea g definida como:

$$g(x)=(x-\alpha)^a\qquad x\in[\alpha,\infty)\qquad \alpha\geq 0\qquad a>0$$

Entonces al sustituir en (3.33) obtenemos:

$$F(x)=1-exp\left\{-\frac{1}{c}(x-\alpha)^a\right\}\qquad \alpha\geq 0\qquad a>0\qquad c>0 \tag{3.38}$$

Esta distribución es conocida como la distribución Weibull de parámetros $(\alpha,c,a)$, de donde tras aplicar el teorema (3.1) nos diria que la distribución Weibull cumple con lo siguiente:

$$E\left[(Y_{r:n}-\alpha)^a|\,Y_{k:n}=t\right]=(t-\alpha)^a+c\sum_{j=k}^{r-1}\frac{1}{n-j}\Leftrightarrow X\sim Weibull\,(\alpha,c,a)\qquad r>k$$

Esto quiere decir que si tenemos una esperanza condicional que pueda ser llevada a la forma:

$$E\left[(Y_{r:n}-\alpha)^a|\,Y_{k:n}=t\right]=(t-\alpha)^a+c\sum_{j=k}^{r-1}\frac{1}{n-j}\qquad \forall t\in(\alpha,\infty)\qquad r>k$$

Entonces la *única* distribución continua con extremos $(\alpha,\infty)$ que cumple con esta esperanza condicional para todo t es la distribución $Weibull(\alpha,c,a)$

Un caso particular de este ejemplo es cuando tomamos $a=0$ y $\alpha=0$ ya que con esto se obtiene la distribución exponencial de parámetro $c$.

La sustitución anterior nos lleva a la siguiente caracterización de la distribución exponencial.

$$E\left[Y_{r:n}|\,Y_{k:n}=t\right]=t+c\sum_{j=k}^{r-1}\frac{1}{n-j}\Leftrightarrow X\sim Exp\,(c)$$

En esta parte hay que hacer énfasis que en todos los ejemplos se está pidiendo que la esperanza condicional sea de esa forma para todo $t$ en el dominio de la distribución. Esto se hace de esta manera ya que es muy fácil construir ejemplos en donde la esperanza es de la forma requerida en un sólo punto $t_0$ pero que no sigue necesariamente la distribución que dice el teorema.

Para ver la necesidad de lo anterior consideremos los siguientes dos ejemplos:

**Ejemplo 3.4.** Consideremos a $Y_{1:5}, Y_{2:5}, \ldots, Y_{5:5}$ los estadísticos de orden de una m.a. de una distribución Uniforme(0,1).

Después de cálculos obtenemos:

$$E[Y_{2:5}|\,Y_{1:5} = 0.1] = \frac{4!}{(3!)(0.9^4)} \int_{0.1}^{1} y_2[1 - y_2]^3 dy_2 = 0.28$$

Sin embargo, notemos que si partimos de la esperanza condicional que hemos encontrado y la llevamos a la forma que aparece en el teorema (3.1) obtendríamos lo siguiente:

$$E[Y_{2:5}|\,Y_{1:5} = 0.1] = 2.8 \Leftrightarrow E[Y_{2:5}|\,Y_{1:5} = 0.1] = 0.1 + 0.18$$

Tomando $\alpha = 0$ y $\beta = \infty$ así como g(t)=t

$$\Leftrightarrow E[g\,(Y_{2:5})|\,Y_{1:5} = 0.1] = g(0.1) + 0.18$$

haciendo $c \sum_{j=1}^{1} \frac{1}{5-j}$ =0.18 lo que implicaría que c=0.72, obtenemos

$$E[g\,(Y_{2:5})|\,Y_{1:5} = 0.1] = g(0.1) + c \sum_{j=1}^{1} \frac{1}{5-j}$$

De donde **si suponemos que la esperanza toma esta forma para todo t** en $(\alpha, \beta)$ y aplicando el teorema (3.1) concluiríamos que $X \sim Exp(0.72)$

A simple vista parecería una contradicción ya que partimos del hecho de que la muestra es obtenida de una distribución uniforme y hemos concluido que X sigue una distribución exponencial, sin embargo esto no es una contradicción, ya que el teorema nos dice que X sigue una distribución exponencial si suponemos que el dominio de la v.a. X de donde es obtenida la muestra aleatoria es $(\alpha, \beta)$ y, además, la esperanza condicional toma la forma en que aparece en el teorema (3.1) para todo t.

Es obvio pues que la distribución uniforme no tiene los mismos extremos de dominio que la distribución exponencial y que además la esperanza condicional no toma la forma que queremos para todo t en $(0, 1)$ pues haciendo unos cálculos más obtenemos que para una distribución uniforme(0,1) la esperanza condicional toma la siguiente forma:

$$E[Y_{2:5}|\,Y_{1:5} = t] = \frac{4}{5}t + \frac{2}{5} \qquad \forall t \in (0, 1)$$

Expresión que **no puede** ser llevada a la forma:

$$E[g\,(Y_{2:5})|\,Y_{1:5} = t] = g(t) + K$$

Lo cual justifica el por qué el teorema no puede ser aplicado en este caso.

**Ejemplo 3.5.** Consideremos ahora $Y_{1:5}, Y_{2:5}, ..., Y_{5:5}$ los estadísticos de orden de una muestra aleatoria proveniente de una población con distribución Weibull(0,1,2).

Entonces tomando en cuenta que $F(x) = 1 - exp\{-x^2\}$ con x positiva y con ayuda de un programa de computadora para aproximar la integral involucrada llegamos a que:

$$E[Y_{2:5}|\, Y_{1:5} = 1] = \frac{4}{exp(-4)} \int_1^{\infty} 2y^2 exp(-4y^2)dy = 1.113169250$$

Nuevamente notemos que si partimos de la esperanza condicional que hemos encontrado y la llevamos a la forma que aparece en el teorema (3.1) obtendríamos lo siguiente:

$$E[Y_{2:5}|\, Y_{1:5} = 1] = 1.113169250 \Leftrightarrow E[Y_{2:5}|\, Y_{1:5} = 1] = 1 + 0.113169250$$

Tomando $\alpha = 0$ y $\beta = \infty$ así como g(t)=t

$$\Leftrightarrow E[g\,(Y_{2:5})|\, Y_{1:5} = 1] = g(1) + 0.113169250$$

haciendo $c \sum_{j=1}^{1} \frac{1}{5-j}$ =0.113169250 lo que implicaría que c=0.452677, obtenemos

$$E[g\,(Y_{2:5})|\, Y_{1:5} = 1] = g(1) + c \sum_{j=1}^{1} \frac{1}{5-j}$$

De donde **si suponemos que la esperanza toma esta forma para todo t** en $(\alpha, \beta)$ y aplicando el teorema (3.1) concluiríamos que $X \sim Exp(0.452677)$

Nuevamente se observa la necesidad de que la esperanza tenga la forma requerida para todo t en el dominio, ya que como puede verse en el ejemplo anterior, a pesar de que ambas distribuciones *(Exponencial y Weibull)* están definidas en el mismo dominio, no se cumplió el teorema debido a que la forma de la esperanza sólo la estamos restringiendo a un sólo valor de $t$.

Debe de ser claro ahora que para garantizar la unicidad de la función de distribución que construimos es necesario que la forma de la esperanza condicional se mantenga para *"todo"* valor de t en el dominio de la función de distribución.

Así pues, si sólo disponemos de la esperanza condicional para una sólo valor de $t$ lo único que podremos hacer es construir una función de distribución que cumpla con esa esperanza condicional, pero no podremos garantizar que ésta será la *única* que lo cumple.

### 3.2.2. Caracterización para las funciones de las forma $F(x) = exp\left\{\frac{1}{c}[g(\beta) - g(x)]\right\}$

Todo lo hecho en la sección anterior puede ser llevado al caso en que $r < k$. Es decir para el caso en que se caracteriza (con ayuda de los teoremas (3.4),(3.5) y (3.6)) a las funciones de distribución de que son de la forma (3.34).

Primero destacaremos que de igual forma que en la sección anterior, cualquier distribución absolutamente continua puede ser obtenida en (3.34). Para ello nuevamente consideremos que tenemos una función de distribución absolutamente continua $H(x)$ con soporte en $(\alpha, \beta)$ y definiremos [12] $g(x) = ln(H(x))$. De donde tras sustituir en (3.34) llegamos a que:

$$F(x) = H(x)^{-\frac{1}{c}} \qquad c < 0, \qquad x \in (\alpha, \beta)$$

De donde es claro que si $c = -1$ se obtiene la distribución H(x).

**Ejemplo 3.6 (Caracterización de la distribución Uniforme(0,1)).** Sea g definida como:

$$g(x) = ln(x) \qquad x \in (0, 1]$$

Entonces al sustituir en (3.34) obtenemos:

$$F(x) = x^{-\frac{1}{c}} \qquad x \in (0, 1) \qquad c < 0$$

Observe que cuando $c = -1$ obtenemos la distribución Uniforme(0,1)

Así pues aplicando el teorema (3.4), concluimos que la distribución Uniforme(0,1) tiene la siguiente caracterización:

$$E\left[ln\left(Y_{r:n}\right)|\, Y_{k:n} = t\right] = ln\left(t\right) - \sum_{j=r}^{k-1} \frac{1}{j} \Leftrightarrow X \sim Uniforme(0, 1) \qquad r < k$$

Lo anterior nos está diciendo que la distribución Uniforme(0,1) es la única distribución absolutamente continua en el intervalo (0,1) que cumple con:

$$E\left[ln\left(Y_{r:n}\right)|\, Y_{k:n} = t\right] = ln\left(t\right) - \sum_{j=r}^{k-1} \frac{1}{j} \qquad \forall t \in (0, 1) \qquad r < k$$

**Ejemplo 3.7 (Caracterización de la distribución *Pareto*$(\alpha, b)$).** Sea g definida como:

$$g(x) = ln\left(1 - \left(\frac{\alpha}{x}\right)^b\right) \qquad x \in (\alpha, \infty], \qquad \alpha > 0, \qquad b > 0$$

Entonces al sustituir en (3.34) obtenemos:

$$F(x) = \left(1 - \left(\frac{\alpha}{x}\right)^b\right)^{-1/c} \qquad x \in (\alpha, \infty) \qquad c < 0 \qquad b > 0, \qquad \alpha > 0$$

Observe que cuando $c = -1$ obtenemos la distribución *Pareto*$(\alpha, b)$

[12] Observe que g así definida cumple con las condiciones impuestas por el teorema (3.4)

Así pues aplicando el teorema (3.4), concluimos que la distribución *Pareto*$(\alpha, b)$ tiene la siguiente caracterización:

$$E\left[ln\left(1-\left(\frac{\alpha}{Y_{r:n}}\right)^b\right)\middle| Y_{k:n}=t\right]=ln\left(1-\left(\frac{\alpha}{t}\right)^b\right)-\sum_{j=r}^{k-1}\frac{1}{j}\Leftrightarrow X\sim Pareto(\alpha,b)$$

Lo anterior nos está diciendo que la distribución *Pareto*$(\alpha, b)$ es la única distribución absolutamente continua en el intervalo $(\alpha,\infty)$ que cumple con:

$$E\left[ln\left(1-\left(\frac{\alpha}{Y_{r:n}}\right)^b\right)\middle| Y_{k:n}=t\right]=ln\left(1-\left(\frac{\alpha}{t}\right)^b\right)-\sum_{j=r}^{k-1}\frac{1}{j}\qquad \forall t\in(\alpha,\infty)\qquad r<k$$

**Ejemplo 3.8.** Sea g definida como:

$$g(x)=x\qquad x\in(-\infty,0]$$

Entonces tras hacer las respectivas sustituciones y utilizando nuevamente el teorema (3.4) obtendríamos la siguiente caracterización:

$$E\left[Y_{r:n}\middle| Y_{k:n}=t\right]=t+c\sum_{j=r}^{k-1}\frac{1}{j}\Leftrightarrow F(x)=exp\left\{-\frac{1}{c}x\right\},\qquad c<0\qquad x\in(-\infty,0)$$

Cabe destacar que las caracterizaciones aquí expuestas sólo son válidas cuando la esperanza condicional tienen la forma $E\left[g(Y_{r:n})\middle| Y_{k:n}=t\right]=g(t)+c\sum_{j=r}^{k-1}\frac{1}{j}$ para todo t en el dominio de la distribución.

## 3.3. Ejercicios

Con ayuda de estos teoremas seremos capaces de resolver ejercicios muy variados como los que se mencionan a continuación.

1. Sea $Y_{1:3}, Y_{2:3}, Y_{3:3}$ los estadísticos de orden de una muestra aleatoria proveniente de una distribución exponencial de parámetro 1. Encuentre las siguientes esperanzas condicionales:
$$E\left[Y_{2:3}|\, Y_{1:3}=1\right], E\left[Y_{3:3}|\, Y_{1:3}=1\right], E\left[Y_{1:3}|\, Y_{1:3}>1\right]$$

2. Sea $Y_{1:3}, Y_{2:3}, Y_{3:3}$ los estadísticos de orden de una muestra aleatoria proveniente de una distribución Uniforme(0,1). Encuentre las siguientes esperanzas condicionales:
$$E\left[ln(Y_{1:3})|\, Y_{3:3}=0.9\right], E\left[ln(Y_{2:3})|\, Y_{3:3}=0.9\right], E\left[ln(Y_{3:3})|\, Y_{3:3}<0.9\right]$$

3. Sea $Y_{1:5}, Y_{2:5}, \ldots, Y_{5:5}$ los estadísticos de orden de una muestra aleatoria proveniente de una distribución absolutamente continua F con dominio en $(1, \infty)$ . Construya "la" función de distribución F dado que se sabe que
$$E\left[ln\left(Y_{1:5}\right)|\, Y_{1:5}>t\right]=ln(t)+2 \qquad \forall t \in (1, \infty)$$

4. Sea $Y_{1:5}, Y_{2:5}, \ldots, Y_{5:5}$ los estadísticos de orden de una muestra aleatoria proveniente de una distribución absolutamente continua F con dominio en $(-\infty, 0)$ . Construya "la" función de distribución F dado que se sabe que
$$E\left[Y_{1:5}|\, Y_{3:5}=t\right]=t-2 \qquad \forall t \in (-\infty, 0)$$

5. Sea $Y_{1:5}, Y_{2:5}, \ldots, Y_{5:5}$ los estadísticos de orden de cierta distribución con dominio en los reales positivos. Encuentra "una" posible forma de la distribución F si sabemos que $E[Y_{5:5}\mid Y_{1:5}=1]=5$

6. Sea $Y_{1:10}, Y_{2:10}, \ldots, Y_{10:10}$ los estadísticos de orden de una muestra aleatoria proveniente de una distribución exponencial de parámetro 1. Encuentre el primer momento de todos los estadísticos de orden.

7. Sea g una función que cumple con las condiciones del teorema (3.4) y supongamos que F es dada por la expresión (3.34). Demuestre que
$$E\left[g\left(Y_{r:n}\right)\right]=E[g\left(Y_{k:n}\right)]+c\sum_{j=r}^{k-1}\frac{1}{j} \qquad con \qquad r<k$$

# Conclusión

Los teoremas expuestos en el capítulo 3 de este trabajo nos dan una caracterización de las distribuciones por medio de la esperanza condicional de los estadísticos de orden. Así pues es por lo anterior que gracias a estos teoremas somos capaces de saber de cuál distribución vino una muestra aleatoria una vez que la esperanza condicional de sus estadísticos de orden es conocida y llevada a la forma en que aparece en alguno de los teoremas.

Inversamente, los teoremas nos ayudan a encontrar la esperanza condicional de alguna función del estadístico de orden una vez que conocemos la función de distribución de donde provino la muestra aleatoria.

Los teoremas (3.1),(3.2),(3.3) identifican de manera única a las funciones de distribución de la forma $F(x) = 1 - exp(-\frac{1}{c}[g(x) - g(\alpha)])$ lo cual ocurre (bajo la notación de estos teoremas) cuando $r$ es mayor que $k$, mientras que los teoremas (3.4),(3.5),(3.6) cumplen con el mismo propósito para las funciones de distribución de la forma $F(x) = exp(\frac{1}{c}[g(\beta) - g(x)])$ es decir cuando $r$ es menor que $k$. Además se probó que cualquier distribución puede ser llevada a cualquiera de estas dos formas haciendo la asignación pertinente a la función $g$, sin embargo esto ocasiona que a veces la función $g$ tome una expresión "complicada" como para poder ser aplicada en algún posible caso práctico.

Un resultado importante que pudimos obtener es que gracias a estos últimos teoremas, fuimos capaces de calcular la esperanza de funciones de cualquier estadístico de orden una vez que la función de distribución es llevada a alguna de las dos formas que se mencionan en los teoremas.

Por otro lado, para aplicar correctamente los resultados fue necesario pedir que la esperanza condicional que aparece en todos los teoremas y corolarios tenga esa forma *para todo t* en el dominio, ya que si sólo se da esa forma para un punto, entonces los teoremas no garantizan la unicidad de la función construida.

Debe ser claro que no por tener la esperanza condicional para un sólo punto seremos capaces de construir una función de distribución única que cumpla con eso, ya que como se pudo ver en los ejemplos del capítulo 3 esto no puede ocurrir. No obstante debe de notarse que estos teoremas nos ayudan a construir una **posible función distribución** de donde provine la muestra aleatoria una vez que se tiene la esperanza condicional para un punto $t_0$ de la forma:

$$E[g(Y_{r:n}) \mid Y_{k:n} = t_0]$$

Donde es necesario que la función $g$ cumpla con las condiciones de los teoremas mencionados.

# Bibliografía


[1] H.A. David. *Order Statics.* John Wiley New York, segunda edition, 1981.

[2] Hogg V.Robert et al. *Introduction to Mathematical Satatistics.* Macmillan, cuarta edition, 1978.

[3] Mood M. Alexander et al. *Introduction to the Theory of Satatistics.* McGraw-Hill, tercera edition, 1974.

[4] M. Ángel García Álvares. *Introducción a la Teoria de la Probabilidad Vol. I y II.* UNAM.

[5] Naray C. Giri. *Introduction to Probability and Statistics Part II.* Marcel Dekker, New York, primera edition.

[6] L.Y. Ouyang. Characterizations through the conditional distribution of function of one order statistic relative to an adjacent one. *Sankhyã Ser A*, pages 500–503, 1995.

[7] Ash B. Robert. *Probability and Measure Theory.* Harcourt, segunda edition, 2000.

[8] Bartle G. Robert. *The Elements of Real Analysis.* Jhon Wiler and Sons, segunda edition.

[9] Walter Rudin. *Principies of Mathematical Analysis.* McGraw-Hill, tercer edition.

[10] Hans Sagan. *Advanced Calculus of Real-Valued Function of Real variable and Vector-Valued Function of Vector Variable.* Houghton Mifflin Company, Boston, primera edition, 1981.

[11] Jong-Wuu Wu and L.Y. Ouyang. On characterizing distributions by conditional expectations of functions of order statistics. *Metrika*, pages 135–147, 1996.