2ej
16

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"LA FUNCION GENERATRIZ"

T E S I S

QUE PARA OBTENER EL TITULO DE:

M A T E M A T I C O

P R E S E N T A :

MENDOZA ALQUICIRA JUAN FRANCISCO

MEXICO, D. F. 1987.

INTRODUCCION

El germen del concepto de Función Generatriz se encuentra en el trabajo del Matemático Francés Abraham de Moivre (1667-1754); esto lo podemos ver en su obra The Doctrine of Chances, en donde encuentra la probabilidad de obtener una suma X al lanzar n dados cada uno con el mismo número de caras usando una técnica que al desarrollarse daría origen a dicho concepto.

Un uso importante de la Función Generatriz consiste en su aplicación a la solución del problema de encontrar la distribución de probabilidad de una Variable Aleatoria que es suma de Variables Aleatorias Independientes.

Otro problema importante está en los Teoremas de Continuidad, a partir de estos Teoremas las Funciones Generatrices pueden ser usadas para el estudio de Distribuciones Límite.

Es nuestro propósito hacer un estudio sobre las Funciones Generatrices en el presente trabajo, cuya organización es la siguiente:

En el capítulo I, sólo se proporcionan las definiciones principales de lo que se entenderá por una Función Generatriz, y veremos algunos ejemplos particulares sobre estas funciones.

En el capítulo II se estudian algunas propiedades que satisfacen estas Funciones Generatrices, además se demuestran unos Teoremas relativos a estas Funciones, así también algunos teoremas auxiliares son enunciados y demostrados; ya que son necesarios para poder establecer los teoremas principales sobre las Funciones Generatrices.

En el capítulo III veremos algunos conceptos importantes sobre convergencia de Funciones de Distribución, los cuales nos proporcionan, propiedades que son útiles para el estudio de la Función Generatriz.

El capítulo IV está compuesto de algunos resultados sobre la Transformada de Laplace.
Los Teoremas de Continuidad se presentan en el capítulo V.
Finalmente se constituye un apéndice, con todos aquellos resultados adicionales que fueron usados en la demostración de algunos Teoremas contenidos en los capítulos anteriores.

CONTENIDO

CAPITULO I

Funciones Generatrices.

Este capítulo aborda sólo las definiciones de lo que se va a entender por una función generatriz de una distribución de probabilidad, así como para una sucesión de números $\{a_i\}$.

Se tratan los casos discreto y continuo, y se dan algunos ejemplos de funciones generatrices.

Empezamos con el caso discreto.

Definición I. Sea $a_0, a_1, \ldots$ una sucesión de números reales, si $A(s) = a_0 + a_1 s + a_2 s^2 + \cdots = \sum_{n=0}^{\infty} a_n s^n$ converge en algún intervalo $-s_0 < s < s_0$, con $s_0 > 0$ entonces $A(s)$ para $-s_0 < s < s_0$ se llamará la función generatriz de la sucesión $\{a_n\}$.

Observación I.

Si $\{a_n\}$ está acotada, es decir existe $K > 0$ tal que $|a_n| \le K$ para toda n, entonces $\left|\sum_{n=0}^{\infty} a_n s^n\right| \le \sum_{n=0}^{\infty} |a_n| |s|^n \le K \sum_{n=0}^{\infty} |s|^n = K/1-|s|$

cuando $|s| < 1$.

Por lo tanto, en este caso la función generatriz está definida por lo menos en el intervalo $-1 < s < 1$.

A continuación veamos dos ejemplos de funciones generatrices.

I. Sea $a_J = 1$ para toda J, entonces la función generatriz de la sucesión $\{a_J\}$ es;

$$A(s) = \sum_{J=0}^{\infty} a_J s^J = \sum_{J=0}^{\infty} s^J = 1/1-s \quad \text{para} \quad |s| < 1.$$

2. *Sea la siguiente sucesión* $a_J = 1/J!$ *su función generatriz es:*

$$A(s) = \sum_{J=0}^{\infty} a_J s^J = \sum_{J=0}^{\infty} s^J/J! = e^s \quad \text{para Todo } s.$$

Función Generatriz de una Variable Aleatoria.

Definición 2. Sea X una Variable Aleatoria que toma valores enteros no negativos y sea $a_j = P[X = j]$ $j=0,1,2,3,\ldots$ Entonces diremos que la función generatriz de la sucesión $\{a_j\}$ es la función generatriz de la Variable Aleatoria X, también en algunas ocasiones se dice que es la función generatriz de la distribución de probabilidad $\{a_j\}$ de la Variable Aleatoria X.

Recordemos que $\sum_j a_j = 1$, entonces vemos que para $|s|\leq 1$ tenemos

$$\left|A(s)\right| = \left|\sum_{n=0}^{\infty} a_n s^n\right| \leq \sum_{n=0}^{\infty} |a_n|\,|s|^n \leq \sum_{n=0}^{\infty} a_n = 1$$

por tanto, la función generatriz de una Variable Aleatoria X que toma valores enteros no negativos, existe al menos para $|s| \leq 1$.

El intervalo $[-1, 1]$ puede ser el único intervalo de convergencia de la serie. En efecto, consideremos la siguiente Variable Aleatoria X, con probabilidades dadas por $P[X = 2^k] = 1/2^k$, $k = 1, 2, \ldots$

Entonces $A(s) = \sum_{k=1}^{\infty} \frac{s^{(2^k)}}{2^k}$

esta serie diverge cuando $|s| > 1$ pues en este caso tenemos

$$\lim_{x \to \infty} |s|^x / x = \infty$$

A continuación veamos algunos ejemplos de funciones generatrices de Variables Aleatorias que sólo toman valores enteros no negativos.

I. Sea X la Variable Aleatoria que toma los valores $1,2,3,4,5,6$, con probabilidades identicas $P[X=1] = \cdots = P[X=6] = 1/6$

Entonces la Función Generatriz de X es:

$$A(s) = \sum_{j=1}^{6} P[X=j]\,s^j = a_1 s + a_2 s^2 + \cdots + a_6 s^6$$

En donde $a_j = P[X = j]$

Entonces

$$A(s) = \frac{s + s^2 + s^3 + s^4 + s^5 + s^6}{6}$$

2. Sea X la variable aleatoria con distribución Bernoulli, sabemos que X toma sólo dos valores 0 y 1 con probabilidades

$P[X=0]=q \quad , \quad P[X=1]=p$ en donde p+q=1.

Entonces la función generatriz de esta variable aleatoria es:

$$A(s) = qs^0 + ps$$
$$= (q + ps)$$

3. Sea X la variable aleatoria con distribución Binomial con parámetros n y p es decir

$P[X=j]=\binom{n}{j} p^j q^{n-j}$ j= 0, 1 2 ,... ,n por tanto la función generatriz de esta distribución es:

$$A(s)=\sum_{j=0}^{n}\binom{n}{j}p^j q^{n-j}s^j = \sum_{j=0}^{n}\binom{n}{j}(ps)^j q^{n-j} = (q+ps)^n$$

4. Sea X la variable aleatoria con distribución Poisson, con parámetro λ , es decir:

$P[X=j] = \frac{\lambda^j e^{-\lambda}}{j}$ j= 0, 1, 2,...

Entonces su función generatriz está dada por:

$$A(s)=\sum_{j=0}^{\infty}\frac{\lambda^j e^{-\lambda}}{j!}s^j = \sum_{j=0}^{\infty}\frac{(\lambda s)^j e^{-\lambda}}{j!} = e^{-\lambda}\sum_{j=0}^{\infty}\frac{(\lambda s)^j}{j!} = e^{-\lambda+\lambda s}$$

5. Sea X la variable aleatoria con distribución geométrica, es decir

$P[X=j]=q^j p$ j= 0,1,2,... en donde p+q=1

la función generatriz de esta variable es:

$$A(s)=\sum_{j=0}^{\infty} q^j p s^j = p\sum_{j=0}^{\infty}(qs)^j = p/1-qs$$

siempre que

$$|qs|<1 .$$

6. *Sea X la variable aleatoria que toma los valores 0, 1 y 2, con probabilidades* $1/4$, $1/2$ y $1/4$ *respectivamente. Entonces*

$$A(s) = \sum_{J=0}^{2} P[X=J]\, s^J = 1/4 + 1/2\, s + 1/4\, s^2$$
$$= 1/4\,(1+s)^2$$

7. *Sea X la variable aleatoria que toma los valores 1, 2, 3, ... con probabilidades* $1/2,\ 1/4,\ 1/8,\ \ldots,\ 1/2^k,\ \ldots$

Entonces vemos que

$$P[X=K] = 1/2^K \quad , \quad K = 1, 2, 3, \ldots$$

Por tanto $\quad A(s) = \sum_{J=1}^{\infty} s^J/2^J$

$$= \sum_{J=1}^{\infty} (s/2)^J = \frac{1}{1-s/2} - 1$$

$$= s/2-s \qquad \text{si} \quad \left|\frac{s}{2}\right| < 1$$

$$\text{o bien} \quad |s| < 2$$

Denotaremos al valor esperado de una variable aleatoria X como $E(X)$ en adelante.

Ahora bien ya vimos que si X es una variable aleatoria que toma sólo valores enteros no negativos, la función generatriz de X está dada por:

$$A(s)=\sum_{j=0}^{\infty} a_J S^J \text{ , donde } a_J = P[X=J] \quad , J=0,1,2,\ldots$$

además se vió que $A(s)$ existe al menos en $-1 \leq s \leq 1$

Notemos lo siguiente:

$$A(s)=\sum_{j=0}^{\infty} a_J S^J = \sum_{J=0}^{\infty} P[X=J]\, S^J$$

$$= E\left(S^X\right)$$

Por tanto notamos que $A(s)=E\left(S^X\right)$ es decir, vemos que $A(s)$, la función generatriz de la variable aleatoria X se puede expresar también como $E(S^X)$, es decir como el valor esperado de la variable aleatoria S^X , lo anterior lo formulamos como una definición.

<u>Definición 4.</u> Si X es cualquier variable aleatoria se define a $\phi_X(s)=E\left(S^X\right)$ como la función generatriz de Momentos Factoriales de X.

En el caso general conviene definir a $\phi_X(s)$ únicamente para $S \geq 0$ pues de otra manera la función $\phi_X(s)$ podría no ser real aún en casos muy simples; por otra parte veremos más adelante que la función $\phi_X(s)$ para $S \geq 0$ nos da la información que requerimos.

Podemos observar que en el caso general la función $\phi_X(s)$ *puede estar bien definida únicamente para* $s = 1$ *, siendo igual a 1 en ese punto e infinita en los otros. Por ejemplo consideremos la variable aleatoria X cuyas probabilidades están dadas por las relaciones siguientes:*

$$P[X=n] = P[X=-n] = 6/\pi^2 n^2 \qquad n = 1, 2, \ldots$$

Entonces $\phi_X(s) = \frac{6}{\pi^2} \sum_{n=-\infty}^{\infty} s^n/n^2$

Pero cuando $s > 1$ $\sum_{n=0}^{\infty} s^n/n^2 = \infty$

y cuando $s < 1$ $\sum_{n=-\infty}^{0} s^n/n^2 = \infty$

por lo tanto $\phi_X(s)$ *es finita únicamente para* $s = 1$.

A continuación se define la función generatriz de momentos para una variable aleatoria.

Definición 5 *Sea X una variable aleatoria, y* $F(x)$ *su función de distribución, entonces la función generatriz de momentos está definida como:*

$M(s) = E(e^{sx})$ *para toda* s *para la cual* $M(s)$ *existe.*

Ahora bien vemos que: $\int_0^\infty e^{sx}\,dF(x)$ *es finita para* $s \leq 0$ *y además vemos que si* $\int_0^\infty e^{sx}\,dF(x) < \infty$ *para un* $s > 0$, *entonces*

$\int_0^\infty e^{s_0 x}\,dF(x) < \infty$ *para todo* $s_0 < s$.

De igual forma vemos que $\int_{-\infty}^0 e^{sx}\,dF(x)$ *es finita para* $s \geq 0$.

y si $\int_{-\infty}^0 e^{sx}\,dF(x) < \infty$ *para un* $s < 0$, *entonces* $\int_{-\infty}^0 e^{s_0 x}\,dF(x) < \infty$ *para todo* $s < s_0$, *ya que* $\int_{-\infty}^0 e^{sx}\,dF(x)$ *es finita para todo* $s \geq 0$.

Entonces esto muestra que $M(s)$ *está definida sobre algún intervalo que contiene al 0.*

Además notamos que si X toma sólo valores no negativos este intervalo contendrá a $(-\infty, 0]$ *y si X toma sólo valores negativos el intervalo contendrá a* $[0, \infty)$.

Ahora bien el intervalo puede consistir unicamente del punto 0, veamos a continuación algunos ejemplos.

Ejemplo 1. Sea P una función de probabilidades definida como:

$$P[X=n] = P[X=-n] = \frac{6}{\pi^2 n^2} \quad , \; n = 1,2,3,\ldots$$

Entonces $M(s) = \frac{6}{\pi^2} \sum_{n=1}^{\infty} 1/n^2 \, e^{ns} + \frac{6}{\pi^2} \sum_{n=1}^{\infty} 1/n^2 \, e^{-ns}$

observamos lo siguiente:

$\lim_{n \to \infty} \frac{e^{ns}}{n^2} = \infty$ cuando $S > 0$, por lo tanto la serie $\sum \frac{e^{ns}}{n^2}$ diverge para $S > 0$.

Para $S < 0$, $\lim_{n \to \infty} \frac{e^{-ns}}{n^2} = \infty$ por lo tanto la serie $\sum_{n=1}^{\infty} \frac{e^{-ns}}{n^2}$ diverge.

Entonces para $S \neq 0$, alguna de las series no converge y por lo tanto $M(s) = \infty$.

Ejemplo 2. Sea $f(x) = 1/\pi(1+x^2)$ la función de densidad de la Variable Aleatoria X. Entonces vemos que $M(s) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{sx}}{1+x^2} \, dx$;

si $S > 0$ $\lim_{x \to \infty} \frac{e^{sx}}{1+x^2} = \infty$ y por tanto $\int_{0}^{\infty} \frac{e^{sx}}{1+x^2} \, dx = \infty$.

Igualmente si $S < 0$ $\lim_{x \to -\infty} e^{sx}/_{1+x^2} = \infty$ y por tanto $\int_{-\infty}^{0} \frac{e^{sx}}{1+x^2} \, dx = \infty$

entonces $M(s) = \infty$ para todo $S \neq 0$.

Ejemplo 3. Sea $f(x)=c\,e^{-|x|^{\alpha}}$ la función de densidad de la Variable Alectoria X donde $0<\alpha<1$, $x\in\mathbb{R}$.

Entonces la función generatriz de momentos está dada por:

$M(s)=c\int_{-\infty}^{\infty}e^{sx}e^{-|x|^{\alpha}}\,dx$. y observamos que si $s>0$

$\lim\limits_{x\to\infty} sx/x^{\alpha} = \infty$ y si $s<0$ $\lim\limits_{x\to-\infty} sx/|x|^{\alpha}=\infty$

por tanto $M(s)=\infty$ para todo $s\neq 0$.

Ejemplo 4.

Sea $f(x)=\begin{cases}1/2\,e^{-|\sqrt{x}|}, & x>0\\ 0, & x\leq 0\end{cases}$ la función de densidad de la variable aleatoria X.

En este caso $M(s)=1/2\int_{0}^{\infty}e^{sx}e^{-|\sqrt{x}|}\,dx$

si $s>0$ $\lim\limits_{x\to\infty} sx/|\sqrt{x}| = \infty$

y si además $s\leq 0$, $M(s)\leq \frac{1}{2}\int_{0}^{\infty}e^{-\sqrt{x}}\,dx = 1$

Entonces $M(s)<\infty$ para $s\leq 0$

$M(s)=\infty$ para $s>0$

A continuación veamos ciertas relaciones entre la función generatriz de momentos $M(s)$ y la función generatriz de momentos factoriales $\phi_X(s)$.

Ya que $M(s) = E(e^{sX})$ $s \in \mathbb{R}$

$\phi_X(s) = E(s^X)$ $s \geq 0$.

Además si $s > 0$, $s^X = e^{X \log s}$, entonces observamos que

$M(\log s) = E(e^{X \log s}) = E(s^X) = \phi_X(s)$,

y además para todo $s \in \mathbb{R}$ tenemos que

$\phi_X(e^s) = E(e^{sX}) = M(s)$ y por tanto tenemos las siguientes relaciones:

$\phi_X(s) = M(\log s)$ para $s > 0$

$M(s) = \phi_X(e^s)$ para todo s real.

En donde claramente estas últimas relaciones son ciertas para todos aquellos valores de s donde $\phi_X(s)$ y $M(s)$ están definidas.

A continuación veamos ciertas relaciones entre la función generatriz de momentos $M(s)$ y la función generatriz de momentos factoriales $\phi_X(s)$.

Ya que $M(s) = E(e^{sX}) \quad s \in \mathbb{R}$

$\phi_X(s) = E(s^X) \quad s \geq 0$.

Además si $s > 0$, $s^X = e^{X \log s}$, entonces observamos que

$$M(\log s) = E(e^{X \log s}) = E(s^X) = \phi_X(s),$$

y además para todo $s \in \mathbb{R}$ tenemos que

$$\phi_X(e^s) = E(e^{sX}) = M(s)$$

y por tanto tenemos las siguientes relaciones:

$\phi_X(s) = M(\log s)$ para $s > 0$

$M(s) = \phi_X(e^s)$ para todo s real.

En donde claramente estas últimas relaciones son ciertas para todos aquellos valores de s donde $\phi_X(s)$ y $M(s)$ están definidas.

Para terminar este capítulo, veamos dos últimas definiciones.

Definición 6. Sea X una variable aleatoria, $M(S)$ su función generatriz de momentos, entonces la función generatriz Acumulante de X está dada como:

$$\begin{aligned} G(s) &= \log M(s) \\ &= \log E(e^{sX}) \end{aligned}$$

para toda s, para la cual $G(s)$ exista.

Como caso particular, observamos que si X es una variable aleatoria que toma un número finito de valores $x_1, x_2, \ldots, x_n$ con probabilidades $p_1, p_2, \ldots, p_n$; es decir X es una variable aleatoria simple, entonces en este caso $G(s)$ siempre existe, ya que vemos que $M(S)$ es la suma de un número finito de cantidades positivas; por lo tanto $M(s) < \infty$ para todo s.

En el capítulo siguiente se estudiarán las propiedades y relaciones entre $G(s)$ y $M(s)$.

Definición 7. Sea X una Variable Aleatoria que toma únicamente valores no negativos y sea $F(x)$ su función de distribución, entonces la siguiente función $\varphi(s) = \int_0^\infty e^{-sx} dF(x)$ se llamará la Transformada de Laplace de la distribución $F(x)$.

Observamos que $\varphi(s)$ es finita $\forall\, s \geq 0$, puesto que;

$$\varphi(s) = \int_0^\infty e^{-sx} dF(x) \leq 1$$

CAPITULO II

PROPIEDADES DE LA FUNCION GENERATRIZ PARA DIFERENTES TIPOS DE VARIABLES ALEATORIAS.

En el capítulo I ya se han visto las definiciones principales de lo que es una Función Generatriz de una distribución de probabilidad, y para una sucesión $\{a_n\}$.

En el presente capítulo nos dedicamos a demostrar algunas propiedades y teoremas que cumplen estas funciones.

En la sección II.I, donde se estudian las funciones generatrices de variables aleatorias discretas que sólo toman valores enteros no negativos, empezamos viendo unos Lemas y Proposiciones, que son propiedades relativas a la función generatriz de variables aleatorias para este caso particular.

El Teorema I, nos muestra que la distribución de probabilidad de una variable aleatoria está determinada de forma única por su función generatriz, El teorema 2 es muy conocido e importante, ya que nos da una forma de hallar la distribución de probabilidad de una variable aleatoria que es suma de variables aleatorias independientes.

El teorema 3 (teorema de Continuidad), es de gran importancia, ya que éste puede ser usado para el estudio de distribuciones Límite;

A continuación el teorema 4 establece condiciones equivalentes para caracterizar a las funciones generatrices de probabilidad para el caso presente.

En la sección II.2 se estudia la Función Generatriz de Momentos de Variables Aleatorias que sólo toman un número finito de valores El Teorema 5, muestra que la Función Generatriz de Momentos tiene una representación en serie de potencias, en la cual los coeficientes de esta serie son los momentos de la Variable Aleatoria, y el Corolario 2 afirma que para calcular los momentos de la Variable Aleatoria, basta calcular la k-ésima derivada de la Función Generatriz de Momentos y evaluarla en el cero, y así obtener el k-ésimo momento de X.

En el Teorema 6, en analogía con el Teorema I de la sección II.I se prueba que la Función Generatriz de Momentos determina a la distribución de forma única.
A continuación se estudia la Función Generatriz Acumulante, y se demuestra que ésta y la Función Generatriz de Momentos son ambas convexas.
El Teorema 7 nos proporciona una forma de calcular la Esperanza y la Varianza de la Variable Aleatoria, conociendo su Función Generatriz Acumulante.
Para finalizar esta sección el Teorema 8 nos da una manera de obtener la Función Generatriz Acumulante de una distribución de probabilidad de una Variable Aleatoria que es suma de Variables Aleatorias Independientes.

En la sección II.2 se estudia la Función Generatriz de Momentos en el caso general.
El teorema 9 nos dice que si la Función Generatriz de Momentos es finita en una vecindad del cero, entonces existen los momentos de todos los ordenes y además la función generatriz tiene una representación en serie de potencias, donde igualmente los coeficientes en esta serie son los momentos respectivos de la Variable Aleatoria X, y como un corolario a este Teorema veremos que la K-ésima derivada de esta función calculada en el cero nos da el K-ésimo momento de la Variable Aleatoria.

En el Teorema IO se dan condiciones necesarias y suficientes para que la función generatriz de momentos exista en una vecindad del cero.
Y para finalizar este capítulo, el Teorema II nos dice que si la función generatriz de momentos existe en una vecindad del cero, entonces ésta determina a la distribución.

II.I ESTUDIO DE LA FUNCION GENERATRIZ DE VARIABLES ALEATORIAS DISCRETAS QUE SOLO TOMAN VALORES ENTEROS NO NEGATIVOS.

Iniciamos este capítulo observando lo siguiente:
Ya hemos definido $\Phi_X(S) = E(S^X)$, donde X es una variable aleatoria y hemos visto que $\Phi_X(S)$ está definida al menos para $|S| \leq 1$ cuando X es una variable aleatoria que toma valores enteros no negativos.

En el mismo caso $\Phi_X(S)$ es una serie de potencias, entonces es posible derivar dicha función en cualquier intervalo $|S| < S_0$ en el que está definida y poder así encontrar sus derivadas sucesivamente como sigue:

Ya que $\Phi_X(S) = \sum_{J=0}^{\infty} a_J S^J$

$$= a_0 S^0 + a_1 S + a_2 S^2 + \cdots \text{ para } \quad |S| < S_0$$

donde $a_J = P[X=J]$, $J = 0, 1, 2, \ldots$

entonces sobre el intervalo $|S| < S_0$ podemos derivar $\Phi_X(S)$ término a término y obtener:

$$\Phi_X'(S) = a_1 + 2a_2 S + 3a_3 S^2 + \cdots = \sum_{J=1}^{\infty} J a_J S^{J-1}$$

$$= E\left(X S^{X-1}\right)$$

De igual forma vemos que:

$$\Phi_X''(S) = \sum_{J=2}^{\infty} J(J-1)\, a_J S^{J-2} = E\left[X(X-1) S^{X-2}\right]$$

y así sucesivamente vemos que:

$$\Phi_X^{(K)}(S) = E\left[X(X-1)\cdots(X-K+1) S^{X-K}\right]$$

de lo cual vemos lo siguiente:

En caso de que $S_0 > 1$ *, es decir el radio de convergencia sea mayor que* 1 *, tenemos que:*

$$\xi_1 = \phi_X^{(1)}(1) = E(X)$$

$$\xi_2 = \phi_X^{(2)}(1) = E(X(X-1))$$

$$\vdots \qquad\qquad \vdots$$

$$\xi_k = \phi_X^{(k)}(1) = E\left[X(X-1)\cdots(X-k+1)\right]$$

a las cantidades ξ_k *se les llama los momentos factoriales de X (debido a su estructura semejante a la de los factoriales).*
De aquí el nombre de función generatriz de momentos factoriales a $\phi_X(s)$.
Hemos supuesto que $S_0 > 1$ *para establecer las relaciones anteriores. El siguiente Lema nos permitirá establecer una relación más general.*

Lema I. Sea $f(x) = \sum_{n=0}^{\infty} b_n X^n$ una serie de potencias convergente en $|x| < 1$ y tal que $b_n \geq 0 \ \forall n$, entonces:

$$\lim_{x \to 1^-} f(x) = \sum_{n=0}^{\infty} b_n$$

Demostración.

Si $\sum_{n=0}^{\infty} b_n < \infty$ el resultado está dado por el Teorema de Abel, el cual se enuncia y demuestra en el apéndice.

Supongamos ahora que $\sum_{n=0}^{\infty} b_n = \infty$ y observemos que la función $f(x)$ definida en $0 < x < 1$ es creciente.

Supongamos entonces que:

$$\lim_{x \to 1^-} f(x) = A < \infty$$

Esto implica en particular que:

$$\sum_{n=0}^{\infty} b_n X^n \leq A \qquad \forall \quad 0 < x < 1$$

Y siendo positivos todos los términos de esta serie, se debe cumplir también que:

$$\sum_{n=0}^{N} b_n X^n \leq A \qquad \forall N, \forall \ 0 < x < 1$$

Por lo tanto, tomando límites en cada una de esta última relación cuando $x \to 1^-$ se debe tener:

$$\sum_{n=0}^{N} b_n \leq A \qquad \forall N$$

$$\therefore \sum_{n=0}^{\infty} b_n \leq A < \infty$$

lo cual contradice la hipótesis. Por lo tanto, si $\sum_{n=0}^{\infty} b_n = \infty$ entonces $\lim_{x \to 1^-} f(x) = \infty$, lo cual termina la demostración del Lema.

<u>Corolario I.</u> Sea X una variable aleatoria que toma solo valores enteros no negativos, sea ϕ_X su función generatriz de momentos factoriales y sea $p_n = P[X=n]$, entonces:

$$\lim_{s \to 1^-} \phi_X^{(k)}(s) = \sum_{n=0}^{\infty} n(n-1)(n-2)\cdots(n-k+1)\, p_n$$

$$= E\left[X(X-1)(X-2)\cdots(X-k+1)\right]$$

donde $\phi_X^{(k)}$ denota la k-ésima derivada de la función ϕ_X.

Sea X una variable aleatoria que toma los valores 0,1,2,3,... con probabilidades $P[X=j]=p_j \qquad j=0,1,2,\ldots$

Consideremos la sucesión $P[X>j]=q_j \quad j=0,1,2,\ldots$ y observemos que $q_k = p_{k+1}+p_{k+2}+\cdots \qquad k\geq 0$

Sean $P(s)$ y $Q(s)$ las funciones generatrices de las sucesiones $\{p_j\}$ y $\{q_j\}$ es decir $P(s)=\sum_{j=0}^{\infty} p_j s^j$ y $Q(s)=\sum_{j=0}^{\infty} q_j s^j$

sabemos que $P(s)$ está definida al menos para $|s|\leq 1$, ahora bien, los coeficientes de $Q(s)$ satisfacen que $q_j\leq 1$ para todo $j\geq 0$, entonces la serie $Q(s)$ converge por lo tanto en el intervalo abierto $|s|<1$.

De esta manera $P(s)$ y $Q(s)$ existen al menos para $|s|\leq 1$, $|s|<1$ respectivamente.

A continuación veamos algunos resultados que establecen unas relaciones entre las funciones antes definidas.

Proposición I.

Para $-1<s<1$, $Q(s) = \dfrac{1-P(s)}{1-s}$

Demostración . Veamos lo siguiente

$$(1-s)\ Q(s) = (1-s)\,(q_0 + q_1 s + q_2 s^2 + q_3 s^3 + \cdots\ )$$
$$= (q_0 + q_1 s + q_2 s^2 + \cdots\ ) - (q_0 s + q_1 s^2 + q_2 s^3 + \cdots\ )$$
$$= q_0 + (q_1 - q_0)s + (q_2 - q_1)s^2 + (q_3 - q_2)s^3 + \cdots + (q_n - q_{n-1})s^n + \cdots$$

pero $q_n - q_{n-1} = (p_{n+1} + p_{n+2} + \cdots) - (p_n + p_{n+1} + \cdots\ ) = -p_n$ para $n \geq 1$

y también $q_0 = p_1 + p_2 + p_3 + \cdots\ = 1 - p_0$

Por tanto:

$$(1-s)\ Q(s) = (1-p_0)s^0 - p_1 s - p_2 s^2 - p_3 s^3 - \cdots - p_n s^n - \cdots$$

es decir $(1-s)\ Q(s) = 1 - \sum_{j=0}^{\infty} p_j s^j$

$$= 1 - P(s)$$

lo cual se quería demostrar.

Proposición 2.

$$\lim_{s\to 1^-} P'(s) = \lim_{s\to 1^-} Q(s)$$

$$\lim_{s\to 1^-} P''(s) = 2\lim_{s\to 1^-} Q'(s)$$

Demostración. Obsérvese primero que todos los límites existen pues P', P'', Q y Q' son crecientes en $(0,1)$.

Sea $s\in(0,1)$, entonces la función $P(s)$ es continua en el intervalo $[s,1]$ y derivable en $(s,1)$, por lo tanto, aplicando el teorema del valor medio, sabemos que existe $s^*\in(s,1)$ tal que: $1-P(s)=(1-s)P'(s^*)$

Pero $1-P(s)=(1-s)\,Q(s)$ $\quad\forall\ s\in(0,1)$

por lo tanto, para cada $s\in(0,1)$ existe $s^*\in(s,1)$ tal que $P'(s^*)=Q(s)$, de lo cual se concluye la primera parte de la demostración.

Para demostrar la segunda parte, consideremos dos casos:

Caso I: $\sum_{n=0}^{\infty} k\,p_k < \infty$

por el Lema 1:

$$\lim_{s\to 1^-} P'(s) = \sum_{k=0}^{\infty} k\,p_k$$

$$\lim_{s\to 1^-} Q(s) = \sum_{k=0}^{\infty} q_k$$

Y entonces por la primera parte de la proposición:

$$\sum_{k=0}^{\infty} q_k = \sum_{k=0}^{\infty} k\,p_k < \infty$$

En particular, $P'(s)$ y $Q(s)$ son continuas en $[0,1]$ y
$$P'(1) = Q(1)$$
Ahora, derivando la relación:
$$1 - P(s) = (1-s)\ Q(s) \qquad s \in (0,1)$$

obtenemos:
$$-P'(s) = (1-s)\ Q'(s) - Q(s)$$

Y de aquí:
$$P'(1) - P'(s) = (1-s)\ Q'(s) + P'(1) - Q(s)$$
$$= (1-s)\ Q'(s) + Q(1) - Q(s) \qquad s \in (0,1)$$

Sea entonces $s \in (0,1)$, Por el teorema del valor medio existen $s^*, s^{**} \in (s,1)$ tales que:
$$P'(1) - P'(s) = (1-s)\ P''(s^*)$$
$$Q(1) - Q(s) = (1-s)\ Q'(s^{**})$$

Por lo tanto:
$$(1-s)\ P''(s^*) = (1-s)\ Q'(s) + (1-s)\ Q'(s^{**})$$
$$\therefore\ P''(s^*) = Q'(s) + Q'(s^{**})$$

Entonces tomando límites cuando $s \to 1^-$ se obtiene el resultado.

Caso II : $\sum_{k=0}^{\infty} k p_k = \infty$

En este caso:
$$\sum_{k=0}^{\infty} k(k-1)\ p_k \geq \sum_{k=2}^{\infty} k p_k = \infty$$
$$\sum_{k=0}^{\infty} k q_k \geq \sum_{k=0}^{\infty} k p_k = \infty.$$

Pero por el Lema **1**:

$$\lim_{s \to 1^-} P''(s) = \sum_{k=0}^{\infty} k(k-1)\, p_k$$

$$\lim_{s \to 1^-} Q'(s) = \sum_{k=0}^{\infty} k\, p_k$$

Por lo tanto se tiene el resultado.

Corolario 2. La esperanza $E(X)$ satisface las relaciones:

$$E(X) = \lim_{s \to 1^-} P'(s) = \lim_{s \to 1^-} Q(s)$$ o bien;

$$E(X) = \sum_{j=1}^{\infty} j\, p_j = \sum_{k=0}^{\infty} q_k$$

Corolario 3. Se tiene que la Varianza satisface:

$$Var(X) = \lim_{s \to 1^-} P''(s) + \lim_{s \to 1^-} P'(s) - \left[\lim_{s \to 1^-} P'(s) \right]^2$$

$$= 2 \lim_{s \to 1^-} Q'(s) + \lim_{s \to 1^-} Q(s) - \left[\lim_{s \to 1^-} Q(s) \right]^2$$

__Teorema I__. *La distribución de probabilidad de una variable aleatoria con valores enteros no negativos está determinada de manera única por su función generatriz.*

__Demostración__. *Sea X una variable aleatoria con la distribución de probabilidad* $\{a_k, k \geq 0\}$ *, donde* $a_k = P[X=k]$ *con* $k = 0, 1, 2, \ldots$ *Como hemos visto* $A(s) = \sum_{n=0}^{\infty} a_n s^n$ *es convergente en el intervalo* $|s| \leq 1$ *, entonces es posible derivar término a término esta serie de potencias en una vecindad del cero:*

$$A'(s) = \sum_{n=1}^{\infty} n a_n s^{n-1} = a_1 + 2a_2 s + 3a_3 s^2 + \cdots$$

$$A''(s) = \sum_{n=2}^{\infty} n(n-1) s^{n-2} = 2a_2 + 6a_3 s + 12 a_4 s^2 + \cdots$$

$$\vdots$$

$$A^{(k)}(s) = \sum_{n=k}^{\infty} n(n-1)(n-2)\cdots(n-k+1)\, a_n s^{n-k} = \sum_{n=k}^{\infty} k! \binom{n}{k} a_n s^{n-k}$$

Por tanto

$$\frac{A^{(k)}(s)}{k!} = \sum_{n=k}^{\infty} \binom{n}{k} a_n s^{n-k} = \binom{k}{k} a_k s^0 + \binom{k+1}{k} a_{k+1} s + \binom{k+2}{k} a_{k+2} s^2 + \cdots$$

entonces si hacemos $s=0$ *vemos que*

$$\frac{A^{(k)}(0)}{k!} = a_k \qquad \forall\, k.$$

Lo anterior demuestra que podemos recobrar todos los a_k *a partir de* $A(s)$.

Por tanto, no solo la distribución de probabilidad determina a la función generatriz, sino que también la función generatriz determina de manera única a la distribución.

Notamos que la anterior demostración nos da una fórmula explicita para calcular las probabilidades conociendo la función generatriz, es por ello que a veces a esta función generatriz, que hemos llamado función generatriz de momentos factoriales, se le conoce también en este caso particular como función generatriz de probabilidades.

A continuación veamos un Teorema que satisface la Función Generatriz en general.

Teorema 2. Si las variables aleatorias $X_1, X_2, \ldots, X_n$ son independientes y tienen a $\phi_{X_1}, \phi_{X_2}, \ldots, \phi_{X_n}$ como funciones generatrices, entonces la función generatriz de la suma $X_1 + X_2 + \cdots + X_n$ está dada por el producto $\phi_{X_1} \phi_{X_2} \cdots \phi_{X_n}$

Demostración. Si X_1 y X_2 son variables aleatorias independientes, entonces S^{X_1} y S^{X_2} también lo son, entonces:

$$E\left[S^{X_1+X_2}\right] = E\left[S^{X_1} S^{X_2}\right] = E\left[S^{X_1}\right] E\left[S^{X_2}\right]$$

es decir la función generatriz de la variable $X_1 + X_2$ esel producto de la función generatriz de la variable X_1 por la función generatriz de la variable X_2.

Ahora bien la conclusión es inmediata, ya que si suponemos que $X_1, X_2, \ldots, X_{n-1}$ cumplen la condición del Teorema: es decir la función generatriz de la suma $X_1 + X_2 + \cdots + X_{n-1}$ es igual a $\phi_{X_1} \phi_{X_2} \cdots \phi_{X_{n-1}}$ entonces es inmediato que la función generatriz de la suma $(X_1 + X_2 + \cdots + X_{n-1}) + X_n$ está dada por: $(\phi_{X_1} \phi_{X_2} \cdots \phi_{X_{n-1}}) \phi_{X_n}$ ya que $(X_1 + X_2 + \cdots + X_{n-1})$ y X_n son independientes y esto termina la demostración del Teorema.

Como observación vemos lo siguiente:

El Teorema 1 afirma que la función generatriz determina de manera única a la distribución de probabilidad correspondiente a una variable aleatoria que sólo toma valores enteros no negativos; propiedad que veremos más adelante también la satisface la función generatriz de momentos, y la Transformada de Laplace respectivamente.

Pero en el caso presente vemos que el Teorema 2 anteriormente demostrado, nos da un método para estudiar sumas de Variables Aleatorias Independientes, que sólo toman valores enteros no negativos por medio de Funciones Generatrices.

En algunos casos el producto de las Funciones Generatrices toma una forma simple y entonces podemos deducir su distribución de probabilidad observando su serie de potencias.

En seguida damos un ejemplo sencillo de lo que hemos afirmado anteriormente.

Ejemplo. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes, con distribución Bernoulli, es decir $P[X_i=0]=q$, $P[X_i=1]=p$ para toda i, por lo tanto la función generatriz de cada X_i es:

$$\phi_{X_i}(s) = \sum_{j=0}^{1} P[X_i=j]s^j = qs^0+ps = q+ps$$

entonces la función generatriz de S_n, donde $S_n = X_1+X_2+\cdots+X_n$ está dada por $[\phi(s)]^n = (q+ps)^n$

$$[\phi(s)]^n=(q+ps)^n=\sum_{k=0}^{n}\binom{n}{k}q^{n-k}(ps)^k=\sum_{k=0}^{n}\binom{n}{k}q^{n-k}p^k s^k$$

Por otro lado, por definición la Función Generatriz de S_n está dada por: $\sum_{k=0}^{\infty} P[S_n=k]s^k$

Por comparación de estas dos últimas expresiones vemos que:

$$P[S_n=k]=\binom{n}{k}p^k q^{n-k}\quad,\ 0\le k\le n\ ,\quad P[S_n=k]=0\ ,\ k>n$$

Vemos que S_n es una Binomial con parámetros n y p, es decir hemos demostrado que la suma de n variables aleatorias independientes $X_1, X_2, \ldots, X_n$, cada una distribuida como una Bernoulli se distribuye como una Binomial con parámetros n y p.

El siguiente es un teorema de Continuidad, el cual nos permite estudiar el límite de una sucesión de distribuciones; es decir dada una sucesión de distribuciones convergente, el límite lo podemos conocer, calculando el límite de la sucesión de Funciones Generatrices correspondiente, e inversamente; ya que el teorema establece condiciones necesarias y suficientes para existencia de tales límites.

<u>Teorema 3.</u> Sea $\{a_{k,n}\}$ una sucesión, donde para cada n fija $\{a_{k,n}\}$ es una distribución de probabilidad, es decir $0 \leq a_{k,n} \leq 1$ y $\sum_{k=0}^{\infty} a_{k,n} = 1$ y sea $A_n(s) = \sum a_{k,n} s^k$ para $|s| \leq 1$, entonces $\{a_{k,n}\}$ es convergente para cada k si, y sólo si $A_n(s)$ converge en algún intervalo $0 < s < R$ con $R > 0$, además si, $\lim_{n \to \infty} a_{k,n} = a_k$ para cada k y $A(s)$ es la función generatriz de $\{a_k\}$, entonces $A(s) = \lim_{n \to \infty} \sum_{k=0}^{\infty} a_{k,n} s^k$ uniformemente en todo intervalo $|s| \leq R$ con $R < 1$.

En particular cuando $\sum a_k = 1$, $A(s)$ es la Función Generatriz de la distribución de probabilidad límite $\{a_k\}$.

<u>Observación.</u> Es posible que $\sum_{k=0}^{\infty} a_k$ sea menor que 1, al finalizar la demostración, se da un ejemplo donde ésto sucede.

<u>Demostración.</u> $\Rightarrow]$ Supongamos que $a_k = \lim_{n \to \infty} a_{k,n}$; sea

$$A_n(s) = \sum_{k=0}^{\infty} a_{k,n} s^k \quad |s| \leq 1 \quad \text{y} \quad A(s) = \sum_{k=0}^{\infty} a_k s^k \quad |s| < 1$$

entonces $|a_{k,n} - a_k| \leq 1$ ya que $0 \leq a_{k,n} \leq 1$ para toda n y k, por lo tanto tenemos que para cualquier $R < 1$ y $|s| \leq R$

$$\left| A_n(s) - A(s) \right| = \left| \sum_{k=0}^{\infty} a_{k,n} s^k - \sum_{k=0}^{\infty} a_k s^k \right| = \left| \sum_{k=0}^{\infty} (a_{k,n} - a_k) s^k \right|$$

$$\leq \sum_{k=0}^{\infty} |a_{k,n} - a_k| \, |s|^k \leq \sum_{k=0}^{\infty} |a_{k,n} - a_k| \, R^k$$

$$= \sum_{k=0}^{r} |a_{k,n} - a_k| \, R^k + \sum_{k=r+1}^{\infty} |a_{k,n} - a_k| \, R^k$$

$$\leq \sum_{k=0}^{r} |a_{k,n} - a_k| + \sum_{k=r+1}^{\infty} R^k \leq \sum_{k=0}^{r} |a_{k,n} - a_k| + R^r \sum_{k=0}^{\infty} R^k$$

$$= \sum_{k=0}^{r} |a_{k,n} - a_k| + R^r/1-R$$

Ahora bien dado $\varepsilon > 0$ *, podemos escoger* r *lo suficientemente grande tal que* $R^r/1-R < \varepsilon$ *, entonces el miembro de la derecha será menor que* 2ε *, ya que para este* $\varepsilon > 0$ *existen* $N_i(\varepsilon) \in \mathbb{N}$ $i = 1, 2, \ldots, r$ *tales que* $|a_{i,n} - a_i| < \varepsilon/r$ *para todo* $n \geq N_i(\varepsilon)$ *y entonces es claro que* $\sum_{k=0}^{r} |a_{k,n} - a_k| < \varepsilon$ *para todo* $n \geq N$ *donde* $N = \sup \{ N_i(\varepsilon) \in \mathbb{N} \mid i = 1, 2, \ldots, r \}$.

Por tanto el miembro de la izquierda puede hacerse tan pequeño como se quiera, uniformemente en s *para* $|s| \leq R$ *, entonces vemos que* $\sum_{k=0}^{\infty} a_{k,n} s^k$ *converge uniformemente a* $A(s)$ *para* $|s| \leq R$.

$\Leftarrow]$ *Supongamos ahora que* $A_n(s)$ *converge en algún intervalo* $0 < s < R$ *con* $1 > R > 0$

Sea $A(s) = \lim_{n \to \infty} A_n(s) \qquad 0 < s < R$

Siendo $A_n(s)$ *una función no negativa y monótona creciente en* $0 < s < R$ *para cada n, el límite* $A(s)$ *también lo es. Entonces el límite cuando* $s \to 0^+$ *de* $A(s)$ *existe. Sea* $A(0)$ *dicho límite*

Ahora bien, para $0<s<R$, *tenemos:*

$$a_{0,n} \leq \sum_{k=0}^{\infty} a_{k,n} s^k = a_{0,n} + \sum_{k=1}^{\infty} a_{k,n} s^k$$

$$\leq a_{0,n} + \sum_{k=1}^{\infty} s^k = a_{0,n} + s/1-s$$

Es decir:

$$a_{0,n} \leq A_n(s) \leq a_{0,n} + s/1-s \qquad \forall \; 0<s<R$$

$$\therefore \quad \lim_{n\to\infty} \sup a_{0,n} \leq A(s) \qquad \forall \; 0<s<R$$

$$\lim_{n\to\infty} \inf a_{0,n} \geqslant A(s) - s/1-s \qquad \forall \; 0<s<R$$

Así que haciendo tender s *a* 0 *obtenemos:*

$$\lim_{n\to\infty} \sup a_{0,n} \leq A(0)$$

$$\lim_{n\to\infty} \inf a_{0,n} \geqslant A(0)$$

$$\therefore \quad \lim_{n\to\infty} \inf a_{0,n} = \lim_{n\to\infty} \sup a_{0,n} = A(0)$$

De manera que el límite cuando $n\to\infty$ *de* $(a_{0,n})$ *existe y es igual a* $A(0)$.

Sea ahora $$A_n^*(s) = \frac{A_n(s) - a_{0,n}}{s} = \sum_{k=1}^{\infty} a_{k,n} s^{k-1} \qquad 0<|s|\leq R$$

$$A^*(s) = \frac{A(s) - A(0)}{s} \qquad 0<|s|\leq R$$

Entonces $$\lim_{n\to\infty} A_n^*(s) = A^*(s) \qquad 0<|s|\leq R$$

Ahora bien, $A_n^*(s)$ *es una función no negativa y monótona en* $0 < s < R$ *; por lo tanto su límite* $A^*(s)$ *también lo es.*

Entonces el límite cuando $s \to 0^+$ *de* $A^*(s)$ *existe.*

Sea, $A^*(0)$ *dicho límite, es decir:*

$$A^*(0) = \lim_{s \to 0^+} A^*(s)$$

Ahora, para $0 < s < R$ *, tenemos:*

$$a_{1,n} \leq \sum_{k=1}^{\infty} a_{k,n} s^{k-1} = a_{1,n} + \sum_{k=2}^{\infty} a_{k,n} s^{k-1}$$

$$\leq a_{1,n} + \sum_{k=2}^{\infty} s^{k-1} = a_{1,n} + s/1-s$$

Es decir:

$$a_{1,n} \leq A_n^*(s) \leq a_{1,n} + s/1-s \qquad \forall\ 0 < s < R$$

como antes se demuestra entonces que el límite cuando $n \to \infty$ *de* $(a_{1,n})$ *existe y es igual a* $A^*(0)$.

Considerando ahora:

$$A_n^{**}(s) = \frac{A_n^*(s) - a_{1,n}}{s} = \sum_{k=2}^{\infty} a_{k,n} s^{k-2} \qquad 0 < s \leq R$$

$$A^{**}(s) = \frac{A^*(s) - A^*(0)}{s} \qquad 0 < s \leq R$$

y haciendo un razonamiento análogo al anterior se demuestra que el límite cuando $n \to \infty$ *de* $(a_{2,n})$ *existe.*

De manera que por un razonamiento de inducción se demuestra que el límite cuando $n \to \infty$ *de* $(a_{k,n})$ *existe para cualquier* K .

Ahora bien veamos que la suma $\sum_{k=0}^{\infty} a_k$ puede ser estrictamente menor que 1.

__Ejemplo.__ Sean $X_1, X_2, \ldots, X_n$ variables aleatorias, tales que: $P[X_1 = 2^k] = 1/2^k \qquad k = 1, 2, \ldots$

$$P[X_2 = 2^{k+1}] = 1/2^k \qquad k = 1, 2, \ldots$$

$$P[X_3 = 2^{k+2}] = 1/2^k \qquad k = 1, 2, \ldots$$

$$\vdots$$

$$P[X_n = 2^{k+(n-1)}] = 1/2^k \qquad k = 1, 2, \ldots$$

Entonces:

$$\lim_{n \to \infty} P[X_n = k] = a_k = 0 \text{ para toda } k, \quad \therefore \sum_{k=0}^{\infty} a_k = 0$$

A continuación veamos un teorema donde se caracteriza a las Funciones Generatrices de Probabilidad de una Variable Aleatoria que toma valores enteros no negativos.

Antes veamos el siguiente Lema.

<u>Lema 2.</u> Sea $\theta \in \mathbb{R}$ y $\{F_{n,\theta} \; n=1,2,\dots\}$ una familia de distribuciones de probabilidad de esperanza θ y varianza $\sigma_n^2(\theta)$

Supongamos que $\sigma_n^2(\theta) \to 0$ cuando $n \to \infty$, entonces:

$$E_{n,\theta}(f) = \int_{-\infty}^{\infty} f(x)\, dF_{n,\theta}(x) \longrightarrow f(\theta) \text{ cuando } n \to \infty$$

para toda función $f: \mathbb{R} \to \mathbb{R}$ continua en θ y acotada.

<u>Demostración.</u> En primer lugar vemos que se tiene lo siguiente:

$$\left| E_{n,\theta}(f) - f(\theta) \right| = \left| \int_{-\infty}^{\infty} f(x)\, dF_{n,\theta}(x) - f(\theta) \right| \leq \int_{-\infty}^{\infty} |f(x) - f(\theta)|\, dF_{n,\theta}(x)$$

Ahora bien, como f es continua en θ, para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x) - f(\theta)| < \varepsilon$ si $|x - \theta| \leq \delta$ y como f está acotada, para x fuera de la vecindad anterior existe $M > 0$ tal que $|f(x) - f(\theta)| < M$.

Por lo tanto

$$\int_{-\infty}^{\infty} |f(x) - f(\theta)|\, dF_{n,\theta}(x) = \int_{[-\delta,\delta]} |f(x) - f(\theta)|\, dF_{n,\theta}(x) + \int_{[-\delta,\delta]^c} |f(x) - f(\theta)|\, dF_{n,\theta}(x)$$

$$< \varepsilon + M\, P\left[|X_n - \theta| > \delta \right]$$

Pero por la desigualdad de Chebyshev, se tiene:

$$P\left[|X_n - \theta| > \delta \right] < \frac{\sigma_n^2(\theta)}{\delta^2}$$

Y como $\sigma_n^2(\theta) \to 0$ cuando $n \to \infty$, existe N tal que $\sigma_n^2(\theta) < \dfrac{\varepsilon\, \delta^2}{M}$ para todo $n \geq N$.

Vemos finalmente que si $n \geq N$ *tendremos:*

$$\left| E_{n,\theta}(f) - f(\theta) \right| < 2\varepsilon$$

lo cual demuestra el Lema.

Corolario 4. *Supongamos que para cada* θ *en un intervalo* I *(finito o infinito) se tiene una familia* $\{F_{n,\theta}\,,\ n = 1,2,\dots\}$ *como en el Lema3 y supongamos que* $\vartheta_n^2(\theta) \longrightarrow 0$ *cuando* $n \to \infty$ *uniformemente en dicho intervalo, entonces para toda función* $f: \mathbb{R} \longrightarrow \mathbb{R}$ *continua en* I *y acotada, se tiene:*

$E_{n,\theta}(f) \longrightarrow f(\theta)$ *cuando* $n \to \infty$ *, uniformemente en* I .

Demostración. *Es inmediato ya que si* $\vartheta_n^2(\theta) \longrightarrow 0$ *, uniformemente en* I *, la cota* $\dfrac{\varepsilon\, \delta^2}{M}$ *es independiente de* θ *y se tiene el resultado.*

Ya con el Lema 2 veamos lo siguiente:

Para cada $n \in \mathbb{N}$ y $\theta \in [0,1]$.

Sea X_n una variable aleatoria Binomial con parámetros n y θ

$S_n = \frac{X_n}{n}$ y $F_{n,\theta}$ la distribución de S_n .

Sabemos que la esperanza y la varianza de S_n tienen los siguientes valores:

$$E(S_n) = \theta$$
$$V(S_n) = \frac{\theta(1-\theta)}{n}$$

Ahora como $\frac{\theta(1-\theta)}{n} \leq \frac{1}{n}$, $V(S_n) \to 0$ uniformemente en θ cuando $n \to \infty$.

Por lo tanto, la familia de distribuciones $\{F_{n,\theta}\ n=1,2,\ldots\ \theta \in [0,1]\}$ satisfacen las hipótesis del corolario 4 al Lema 2, así que, para toda función $f: [0,1] \to \mathbb{R}$ continua, tenemos:

$$\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}\theta^k(1-\theta)^{n-k} = \int_{-\infty}^{\infty} f(x)\, dF_{n,\theta}(x) \longrightarrow f(\theta)$$

uniformemente en el intervalo $[0,1]$ cuando $n \to \infty$.

Este último resultado, no es otra cosa que el Teorema de Aproximación de Bernstein (Robert G. Bartle, The Elements oj Real Analysis Second Edition, Wiley International Edition 1967 p171) .

A esta última expresión del miembro izquierdo, se le conoce como el Polinomio de Bernstein, en seguida lo enunciamos como una definición.

<u>Definición1.</u> El Polinomio de Bernstein de grado n correpóndiente a la función $f: [0,1] \to \mathbb{R}$, lo denotamos como:

$$B_{n,f}(\theta) = \sum_{j=0}^{n} f(jh)\binom{n}{j}\theta^j(1-\theta)^{n-j}, \text{ donde } h = 1/n .$$

Una vez definido el Polinomio de Bernstein $B_{n,f}$ *de grado n, correspondiente a una función* $f:[0,1]\to\mathbb{R}$ *pasaremos a demostrar que existe otra forma de poderlo expresar, que nos será útil.*

<u>Definición 2.</u> Dada una sucesión de números $\{a_J\}$ *infinita, se define el operador diferencia* Δ *mediante* $\Delta a_i = a_{i+1} - a_i$*, éste a su vez produce una nueva sucesión* $\{\Delta a_i\}$ *, ahora bien al aplicar* Δ *por segunda vez se obtiene la sucesión con elementos:*

$$\begin{aligned}\Delta^2 a_i &= \Delta(\Delta a_i)\\ &= \Delta a_{i+1} - \Delta a_i\\ &= a_{i+2} - 2a_{i+1} + a_i\end{aligned}$$

de esta manera se puede definir la r-ésima potencia Δ^r *inductivamente por medio de* $\Delta^r = \Delta\,\Delta^{r-1}$ *, y se puede ver que:*

(A) $$\Delta^r a_i = \sum_{J=0}^{r} \binom{r}{J}(-1)^{r-J} a_{i+J}$$

En efecto, supongamos que la fórmula vale para $r=n$ *y toda* a_i*, entonces:*

$$\begin{aligned}\Delta^{n+1} a_i &= \Delta(\Delta^n a_i) = \Delta^n a_{i+1} - \Delta^n a_i\\ &= \sum_{J=0}^{n}\binom{n}{J}(-1)^{n-J}a_{i+J+1} - \sum_{J=0}^{n}\binom{n}{J}(-1)^{n-J}a_{i+J}\\ &= \sum_{J=0}^{n}\binom{n}{J}(-1)^{n-J}a_{i+J+1} + \sum_{J=0}^{n}\binom{n}{J}(-1)^{n+1-J}a_{i+J}\\ &= \sum_{J=0}^{n}\binom{n}{J}(-1)^{n-J}a_{i+J+1} + \binom{n}{0}(-1)^{n+1}a_i + \sum_{J=1}^{n}\binom{n}{J}(-1)^{n+1-J}a_{i+J}\end{aligned}$$

$$= \binom{n}{0}(-1)^{n+1}a_i + \sum_{J=0}^{n}\binom{n}{J}(-1)^{n-J}a_{i+J+1} + \sum_{J=0}^{n-1}\binom{n}{J+1}(-1)^{n-J}a_{i+J+1}$$

$$= \binom{n}{0}(-1)^{n+1}a_i + \binom{n}{n}(-1)^0 a_{i+n+1} + \sum_{J=0}^{n-1}\binom{n}{J}(-1)^{n-J}a_{i+J+1}$$
$$+ \sum_{J=0}^{n-1}\binom{n}{J+1}(-1)^{n-J}a_{i+J+1}$$

$$= \binom{n+1}{0}(-1)^{n+1}a_i + \binom{n+1}{n+1}(-1)^0 a_{i+n+1}$$
$$+ \sum_{J=0}^{n-1}\left[\binom{n}{J}+\binom{n}{J+1}\right](-1)^{n-J}a_{i+J+1}$$

$$= \binom{n+1}{0}(-1)^{n+1}a_i + \binom{n+1}{n+1}(-1)^0 a_{i+n+1} + \sum_{J=0}^{n-1}\binom{n+1}{J+1}(-1)^{n-J}a_{i+J+1}$$

$$= \binom{n+1}{0}(-1)^{n+1}a_i + \binom{n+1}{n+1}(-1)^0 a_{i+n+1} + \sum_{J=1}^{n}\binom{n+1}{J}(-1)^{n+1-J}a_{i+J}$$

$$= \sum_{J=0}^{n+1}\binom{n+1}{J}(-1)^{n+1-J}a_{i+J}$$

así que la fórmula (A) vale también para $r = n+1$.

Se define $\Delta^0 a_i = a_i$ *para que la igualdad (A) tenga sentido también para* $r = 0$.

Ahora derivemos una relación de reciprocidad que es valida para una pareja arbitraria de sucesiones $\{a_i\}$ *y* $\{c_i\}$, *que permite expresar las diferencias* $\Delta^r a_i$ *en función de* $\Delta^r c_i$ *e inversamente.*

<u>Proposición 3.</u> (fórmula de reciprocidad). Sean $\{a_i\}$ *y* $\{c_i\}$ *dos sucesiones y* $v \in \mathbb{Z}^+$, *entonces:*

$$(B)\quad \sum_{r=0}^{v} c_r \binom{v}{r} \Delta^r a_i = \sum_{J=0}^{v} a_{i+J} \binom{v}{J} (-1)^{v-J} \Delta^{v-J} c_J$$

<u>Demostración.</u>

$$\sum_{r=0}^{v} c_r \binom{v}{r} \Delta^r a_i = \sum_{r=0}^{v} c_r \binom{v}{r} \left[\sum_{J=0}^{r} \binom{r}{J} (-1)^{r-J} a_{i+J} \right]$$

$$= \sum_{r=0}^{v} \left(\sum_{J=0}^{r} \binom{v}{r} \binom{r}{J} (-1)^{r-J} c_r a_{i+J} \right)$$

$$= \binom{v}{0}\binom{0}{0}(-1)^0 c_0 a_{i+0} + \binom{v}{1}\binom{1}{0}(-1)^1 c_1 a_{i+0} +$$

$$\binom{v}{1}\binom{1}{1}(-1)^0 c_1 a_{i+1} + \binom{v}{2}\binom{2}{0}(-1)^2 c_2 a_{i+0} +$$

$$\binom{v}{2}\binom{2}{1}(-1)^1 c_2 a_{i+1} + \binom{v}{2}\binom{2}{2}(-1)^0 c_2 a_{i+2} + \cdots$$

$$+ \binom{v}{v}\binom{v}{0}(-1)^v c_v a_{i+0} + \binom{v}{v}\binom{v}{1}(-1)^{v-1} c_v a_{i+1} + \cdots$$

$$+ \binom{v}{v}\binom{v}{v-1}(-1) c_v a_{i+(v-1)} + \binom{v}{v}\binom{v}{v}(-1)^0 c_v a_{i+v}$$

$$= \sum_{J=0}^{v} \left(\sum_{r=J}^{v} \binom{v}{r} \binom{r}{J} (-1)^{r-J} C_r \right) a_{i+J}$$

Pero

$$\sum_{r=J}^{v} \binom{v}{r} \binom{r}{J} (-1)^{r-J} C_r = \sum_{K=0}^{v-J} \binom{v}{K+J} \binom{K+J}{J} (-1)^{K} C_{K+J}$$

$$= \sum_{K=0}^{v-J} \binom{v}{J} \binom{v-J}{K} (-1)^{v-J-K} (-1)^{2K+J-v} C_{K+J}$$

$$= \binom{v}{J} (-1)^{J-v} \sum_{K=0}^{v-J} \binom{v-J}{K} (-1)^{v-J-K} C_{K+J}$$

$$= \binom{v}{J} (-1)^{v-J} \Delta^{v-J} C_J$$

$\therefore$

$$\sum_{r=0}^{v} C_r \binom{v}{r} \Delta^r a_i = \sum_{J=0}^{v} \left[\binom{v}{J} (-1)^{v-J} \Delta^{v-J} C_J \right] a_{i+J}$$

que es lo que se quería demostrar.

En particular vemos que si $a_i = 1$ *para toda* i *entonces*
$\Delta^0 a_i = 1$ *y* $\Delta^K a_i = 0$ *para todo* $K \geq 1$ *de manera que (B) se reduce a:*

$$(C) \qquad c_0 = \sum_{J=0}^{\nu} \binom{\nu}{J} (-1)^{\nu-J} \Delta^{\nu-J} c_J$$

Aplicando (C) a la sucesión $\{c_{J+K}\}$ *, obtenemos:*

$$c_K = \sum_{J=0}^{\nu} \binom{\nu}{J} (-1)^{\nu-J} \Delta^{\nu-J} c_{J+K}$$

la cual es una fórmula de inversión que expresa la sucesión dada en función de sus diferencias.
Ahora bien, sea $0 < \theta < 1$ *fija y sea* $c_r = \theta^r$ *, entonces vemos lo siguiente:*

$$\Delta^K c_r = \sum_{J=0}^{K} \binom{K}{J} (-1)^{K-J} c_{r+J} = \sum_{J=0}^{K} \binom{K}{J} (-1)^{K-J} \theta^{r+J}$$

$$= \sum_{J=0}^{K} \binom{K}{J} (-1)^{J+K} \theta^{r+J} = \sum_{J=0}^{K} \binom{K}{J} (-1)^J \theta^J \theta^r (-1)^K$$

$$= \sum_{J=0}^{K} \binom{K}{J} (-\theta)^J \theta^r (-1)^K$$

$$= \theta^r (-1)^K \sum_{J=0}^{K} \binom{K}{J} (-\theta)^J = \theta^r (1-\theta)^K (-1)^K$$

por lo tanto (B) toma la forma siguiente:

$$(D) \qquad \sum_{r=0}^{\nu} \theta^r \binom{\nu}{r} \Delta^r a_i = \sum_{J=0}^{\nu} a_{i+J} \binom{\nu}{J} \theta^J (1-\theta)^{\nu-J}$$

ya que $\Delta^{\nu-J} C_J = \theta^J (1-\theta)^{\nu-J} (-1)^{\nu-J}$

Ahora bien, sea $f: \mathbb{R} \to \mathbb{R}$ *y dados* $x \in \mathbb{R}$ *y* $h > 0$ *definamos*

$$a_K = f(x+Kh) \qquad K = 0, 1, 2, \ldots$$

Entonces definamos las siguientes razones de diferencias para una función $f(x)$.

Definición 3. $\underset{h}{\Delta} f(x) = \dfrac{f(x+h) - f(x)}{h} = V_1(x)$

$$\underset{h}{\Delta}^2 f(x) = \underset{h}{\Delta} V_1(x) = \frac{V_1(x+h) - V_1(x)}{h} = V_2(x)$$

$$\underset{h}{\Delta}^3 f(x) = \underset{h}{\Delta} V_2(x) = \frac{V_2(x+h) - V_2(x)}{h} = V_3(x)$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\underset{h}{\Delta}^n f(x) = \underset{h}{\Delta} V_{n-1}(x) = \frac{V_{n-1}(x+h) - V_{n-1}(x)}{h} = V_n(x)$$

se ve inmediatamente $\underset{h}{\Delta}^n f(x) = \dfrac{\Delta^n a_0}{h^n}$ *, por lo tanto:*

$\underset{h}{\Delta}^r f(x)$ *tiene la siguiente expresión:*

$$\underset{h}{\Delta}^r f(x) = h^{-r} \sum_{J=0}^{r} \binom{r}{J} (-1)^{r-J} f(x+Jh)$$

Definamos $\underset{h}{\Delta}^0 f(x) = f(x)$ *, es decir* $\underset{h}{\Delta}^0 f(x) = \Delta^0 a_0$

Apliquemos (D) a la sucesión $a_K = f(x+Kh)$ *tomando* $i=0$, $\nu = n$ *y sustituyendo* $\Delta^r a_0$ *por su valor en términos de* $\underset{h}{\Delta}^r f(x)$.

Obtenemos:

$$\sum_{r=0}^{n} \theta^r \binom{n}{r} h^r \underset{h}{\Delta}^r f(x) = \sum_{J=0}^{n} f(x+Jh) \binom{n}{J} \theta^J (1-\theta)^{n-J}$$

En particular, tomando $x=0$ *y* $h = 1/n$ *obtenemos en el segundo miembro de esta igualdad el polinomio de Bernstein de grado n correspondiente a la función* f *restringida al intervalo* $[0,1]$ *; es decir se obtiene:*

$$(\text{II})\quad B_{n,f}(\theta) = \sum_{r=0}^{n} \binom{n}{r} (h\theta)^r \underset{h}{\Delta}^r f(0)$$

Teorema 4. Sea $f:[0,1]\to \mathbb{R}$ una función continua, las siguientes condiciones son equivalentes:

I) f es la función generadora de probabilidades restringida al intervalo $[0,1]$ de una variable aleatoria que toma sólo valores enteros no negativos.

II) $f(1)=1$ y $f^{(K)}(x)\geq 0$ $\forall$ $x\in(0,1)$ y todo $K=0,1,2,\ldots$

III) $f(1)=1$ y $\forall n\in \mathbb{N}$ y $K=0,1,2,\ldots,n$ $\quad \underset{h}{\Delta}^{K} f(0)\geq 0$ con $h=1/n$

Demostración.

I) $\Longrightarrow$ II)

Si f satisface (I) entonces f es de la forma siguiente:

$$f(x)=\sum_{J=0}^{\infty} P_J X^J \qquad 0\leq X\leq 1$$

donde $\quad 0\leq P_J\leq 1 \quad$ y $\quad \sum_{J=0}^{\infty} P_J = 1$

Entonces $f(1)=1$, $f(x)\geq 0$ $\forall$ $x\in[0,1]$ y además por tener una expresión como serie de potencias f es infinitamente diferenciable en $(0,1)$ y sus derivadas están dadas por:

$$f^{(K)}(x)=\sum_{J=K}^{\infty} J(J-1)\cdots(J-K+1)\, P_J\, X^{J-K} \qquad K=1,2,\ldots$$

por lo tanto se satisface (II)

II) $\Longrightarrow$ III)

Suponiendo (II) sea $n\in\mathbb{N}$ y $h=1/n$. Definamos inductivamente la siguiente familia de funciones

$$V_0:[0,1]\to\mathbb{R}$$

$$V_0(x)=\underset{h}{\Delta}^{0} f(x)=f(x)$$

$$V_k : [0, 1-kh] \longrightarrow \mathbb{R}$$

$$V_k(x) = \frac{V_{k-1}(x+h) - V_{k-1}(x)}{h} \qquad k = 1, 2, \ldots, n-1$$

Entonces V_k *es una función continua no negativa en* $[0, 1-kh]$, *infinitamente diferenciable en* $(0, 1-kh)$ *y todas sus derivadas son no negativas en el mismo intervalo.*

En efecto, $V_0 = f$ *y por hipótesis tiene estas propiedades.*
La demostración de estas propiedades para cualquier k se hace entonces por inducción:

supongamos que V_k *tiene las propiedades deseadas con* $k < n-1$ *entonces por definición:*

$$V_{k+1}(x) = \frac{V_k(x+h) - V_k(x)}{h}$$

Así que:

I) Siendo V_k *continua en* $[0, 1-kh]$ *e infinitamente diferenciable en* $(0, 1-kh)$, V_{k+1} *lo es también en* $[0, 1-(k+1)h]$ *y* $(0, 1-(k+1)h)$ *respectivamente.*

2) Siendo no negativas todas las derivadas de V_k, *tanto* V_k *como sus derivadas son funciones no decrecientes, por lo tanto*

$$V_{k+1}^{(m)}(x) = \frac{V_k^{(m)}(x+h) - V_k^{(m)}(x)}{h}$$

es no negativa para $x \in (0, 1-kh)$ *cualquiera que sea* $m = 0, 1, 2, \ldots$

Finalmente siendo V_k *continua en* $[0, 1-kh]$ *y no negativa en* $(0, 1-kh)$, V_k *es no negativa en el intervalo cerrado* $[0, 1-kh]$

Pero por definición:

$$\underset{h}{\Delta}^{k} f(0) = V_k(0)$$

$$\therefore \quad \underset{h}{\Delta}^{k} f(0) \geq 0$$

III) $\Longrightarrow$ *I)*

Supongamos (III) y consideremos el polinomio de Bernstein de grado n correspondiente a la función f *. Hemos visto que éste tiene las siguientes expresiones:*

$$B_{n,f}(\theta) = \sum_{r=0}^{n} \binom{n}{r} (h\theta)^r \Delta_h^r f(0) \quad \text{con} \quad h = 1/n$$

$$B_{n,f}(\theta) = \sum_{J=0}^{n} f(Jh) \binom{n}{J} \theta^J (1-\theta)^{n-J} \quad \text{con} \quad h = 1/n$$

De la primera expresión y usando la condición (III) se obtiene que los coeficientes de θ^r *son no negativos para cualquier* $r = 0,1,2,\ldots,n$

La suma de estos coeficientes está dada por:

$$\sum_{r=0}^{n} \binom{n}{r} h^r \Delta_h^r f(0) = B_{n,f}(1)$$

Pero, de la segunda expresión se obtiene:

$$B_{n,f}(1) = f(1) = 1 .$$

Por lo tanto, $B_{n,f} : [0,1] \longrightarrow \mathbb{R}$ *representa la función generadora de probabilidades de una variable aleatoria que toma los valores 0,1,2,3,...,n.*

Ahora bien, por el teorema de aproximación de Bernstein demostrado anteriormente, se tiene:

$B_{n,f}(\theta) \longrightarrow f(\theta)$ *cuando* $n \to \infty$ *, uniformemente en el intervalo* $[0,1]$.

Entonces, por el teorema 3 de este capítulo, f *es la función generatriz de una sucesión* $\{a_k,\ k=0,1,2,\ldots\}$ *con* $a_k \geq 0$.

Finalmente, como $f(1)=1$, f *es la función generatriz de una variable aleatoria que toma solo valores enteros no negativos, lo cual demuestra (I).*

II.2 FUNCIONES GENERATRICES DE VARIABLES ALEATORIAS QUE SOLO TOMAN UN NUMERO FINITO DE VALORES.

Ya hemos visto que si X es una Variable Aleatoria que toma un número finito de valores $X_1, X_2, \ldots, X_n$ *con probabilidades* $P_1, P_2, \ldots, P_n$ *su Función Generatriz de Momentos:*

$$M(s) = E(e^{sX}) = \sum_{J=1}^{n} P_J \, e^{sX_J}$$

está definida para todo número real s *, además vimos que* $M(s) > 0$ *para todo real* s .

En los siguientes Teoremas veremos algunas propiedades que satisface esta función $M(s)$.

El siguiente Teorema nos muestra a $M(s)$ *en su representación en serie de potencias, en donde vemos que los coeficientes en esta representación son los momentos respectivos de la Variable Aleatoria X.*

Teorema 5. Si $M(s)$ es la función generatriz de momentos de una variable aleatoria simple (es decir toma solo un número finito de valores), entonces:

$$M(s) = \sum_{k=0}^{\infty} \frac{s^k}{k!} E(X^k)$$

Demostración. Sean $x_1, x_2, \ldots, x_n$ los valores que toma X y $C = \max |x_i|$ entonces

$$S_m = \sum_{k=0}^{m} \frac{(sX)^k}{k!} \leq \sum_{k=0}^{m} \frac{(|s|\, C)^k}{k!} \leq \sum_{k=0}^{\infty} \frac{|s|^k C^k}{k!} = e^{|s| C}$$

$$\forall m \in \mathbb{N},\ s \in \mathbb{R}$$

Por lo tanto $\{S_m\}$ está uniformemente acotada

Aplicando entonces el teorema de convergencia dominada (ver apéndice) , se obtiene:

$$M(s) = E\left[e^{sX}\right] = E\left[\sum_{k=0}^{\infty} \frac{s^k X^k}{k!}\right] = \lim_{m \to \infty} E\left[\sum_{k=0}^{m} \frac{s^k X^k}{k!}\right]$$

$$= \lim_{m \to \infty} \sum_{k=0}^{m} \frac{s^k E[X^k]}{k!} = \sum_{k=0}^{\infty} \frac{s^k E[X^k]}{k!}$$

Como corolario veamos un resultado donde a partir de la Función Generatriz $M(s)$ podemos calcular los momentos de la variable aleatoria X.

Corolario 5. $$E(X^k) = M^{(k)}(0)$$

Demostración. Del Teorema anterior se sigue en particular que M *tiene un desarrollo en serie de potencias en cualquier vecindad de cero, por lo tanto* M *es infinitamente diferenciable en todo punto* s *y sus derivadas se obtienen derivando término a término la serie de potencias que la representa, es decir:*

$$M^{(m)}(s) = \sum_{k=m}^{\infty} \frac{k(k-1)\cdots(k-m+1)\, s^{k-m}}{k!} E(X^k)$$

En particular:

$$M^{(m)}(0) = E(X^m) \qquad m = 0, 1, 2, \ldots$$

Podemos ver también esta propiedad derivando término a término la serie $M(s) = \sum_{J=1}^{n} p_J e^{sx_J}$ es decir:

$$M'(s) = \sum_{J=1}^{n} p_J x_J e^{sx_J}$$

$$M''(s) = \sum_{J=1}^{n} p_J x_J^2 e^{sx_J}$$

y así sucesivamente:

$$M^{(k)}(s) = \sum_{J=1}^{n} p_J x_J^k e^{sx_J}$$

y entonces haciendo $s = 0$;

$$M^{(k)}(0) = \sum_{J=1}^{n} p_J x_J^k = E(X^k)$$

El siguiente Teorema nos asegura que esta función generatriz nos determina a la distribución de la variable aleatoria X de manera única; propiedad que vimos también la satisface la función generatriz de variables aleatorias que sólo toman valores enteros no negativos.

Teorema 6. *La Función* $M(s)$ *determina a la distribución de la Variable Aleatoria X de manera única.*

Demostración. *Sean* X *y* Y *dos variables aleatorias simples, que toman los valores* $x_1, x_2, \ldots, x_N$ *y* $y_1, y_2, \ldots, y_M$ *con probabilidades* $p_1, p_2, \ldots, p_N$ *y* $q_1, q_2, \ldots, q_M$ *respectivamente.*

Consideremos las siguientes sumas exponenciales finitas:

$$M(s) = E\left(e^{sX}\right) = \sum_{J} p_J e^{sx_J}$$

$$N(s) = E\left(e^{sY}\right) = \sum_{i} q_i e^{sy_i}$$

y supongamos que:

$M(s) = N(s)$ *para todo* s *, sean* $x_{J0} = \max\{x_J\}$

$y_{i0} = \max\{y_i\}$

veamos lo siguiente:

$$\frac{M(s)}{p_{J0} e^{sx_{J0}}} = \frac{p_1 e^{sx_1} + p_2 e^{sx_2} + \cdots + p_N e^{sx_N}}{p_{J0} e^{sx_{J0}}}$$

$$= \frac{p_1}{p_{J0}} e^{s(x_1 - x_{J0})} + \frac{p_2}{p_{J0}} e^{s(x_2 - x_{J0})} + \cdots + \frac{p_N}{p_{J0}} e^{s(x_N - x_{J0})}$$

ahora bien si $x_{J0} = \max\{x_J\}$ *entonces* $x_J < x_{J0}$ $\quad J \neq J_0$

y por lo tanto $\displaystyle \lim_{s \to \infty} \frac{M(s)}{p_{J0} e^{sx_{J0}}} = 1$

De igual forma se puede ver que $\displaystyle \lim_{s \to \infty} \frac{N(s)}{q_{i0} e^{sy_{i0}}} = 1$*, así que*

$$\lim_{s \to \infty} \frac{p_{J0} e^{sx_{J0}}}{q_{i0} e^{sy_{i0}}} = 1$$

por lo tanto $x_{J0} = y_{i0}$ *y* $p_{J0} = q_{i0}$.

Ahora bien el mismo argumento se aplica a:

$$\sum_{J \neq J_0} P_J e^{sX_J} = \sum_{\lambda \neq \lambda_0} q_\lambda e^{sy_\lambda}$$

y esto muestra inductivamente que para cada J existe λ tal que $X_J = y_\lambda$ y $P_J = q_\lambda$ y viceversa.

De esta manera la función $M(s) = \sum_J P_J e^{sX_J}$

determina de forma única a los X_J y P_J es decir, determina de manera única a la distribución de la Variable Aleatoria X.

Ahora bien, recordemos que en el capítulo I se ha definido a la Función Generatriz Acumulante de X como:

$G(s) = \log M(s) = \log E(e^{sX})$, *donde* $M(s)$ *es la Función Generatriz de Momentos de la Variable Aleatoria simple X.*

A continuación veremos que tanto $G(s)$ *como* $M(s)$ *son ambas convexas.*

Se tiene: $G'(s) = \dfrac{M'(s)}{M(s)}$

y también $G''(s) = \dfrac{M(s)M''(s) - M'(s)M'(s)}{M^2(s)} = \dfrac{MM'' - (M')^2}{M^2}$

pero sabemos que $M^{(k)}(s) = \sum_J p_J x_J^k e^{sx_J} = E(X^k e^{sX})$

entonces $M''(s) = E(X^2 e^{sX}) \geq 0$

Ahora bien

$$[M'(s)]^2 = [E(Xe^{sX})]^2 \leq E(e^{sX})\,E(X^2 e^{sX})$$
$$= M(s)\,M''(s)$$

esto último por la desigualdad de Schuarz, y entonces

$0 \leq MM'' - (M')^2$ *; por lo tanto* $G''(s) \geq 0$.

De esta forma la Función Generatriz de Momentos y la Función Generatriz Acumulante son ambas convexas.

En seguida veamos un Teorema que nos dice como poder calcular la Esperanza y la Varianza de una Variable Aleatoria X, una vez que se conoce su Función Generatriz Acumulante.

Teorema 7. *La Función Generatriz Acumulante* $G(s)$ *satisface*

$G'(0) = E(X)$ *y además* $G''(0) = Var(X)$.

Demostración. *Como sabemos* $G(s) = \log M(s)$, *entonces vemos que*

$G'(s) = M'(s)/M(s)$ *y* $G''(s) = \dfrac{M(s)M''(s) - M'(s)M'(s)}{M^2(s)}$

observamos que $M(0) = E(1) = 1$.

Por tanto $G'(0) = \dfrac{M'(0)}{M(0)} = M'(0) = E(X)$

y

$$G''(0) = \frac{M(0)M''(0) - (M'(0))^2}{M^2(0)} = M''(0) - [M'(0)]^2$$

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

Ahora veamos la siguiente propiedad de $G(s)$.

Si hacemos $m_k = E(X^k)$ *, entonces observamos lo siguiente:*

$$M(s)=\sum_{k=0}^{\infty}\frac{s^k}{k!}E(X^k)=\frac{s^0}{0!}E(X^0)+\sum_{k=1}^{\infty}\frac{s^k}{k!}E(X^k)$$

$$=1+\sum_{k=1}^{\infty}\frac{s^k}{k!}E(X^k) .$$

Ahora bien ya que $M(s)\to 1$ *cuando* $s\to 0$ *, entonces la última expresión es válida para* $-s_0<s<s_0$ *con* $s_0>0$ *, es decir en alguna vecindad del* 0 .

Usando la expresión del Logaritmo en serie de Taylor;

$$\log(1+\tau)=\tau-\frac{1}{2}\tau^2+\frac{1}{3}\tau^3-\frac{1}{4}\tau^4+\cdots=\sum_{k=1}^{\infty}\frac{(-1)^{k+1}\tau^k}{k}$$

para $-1<\tau<1$.

Obtenemos lo siguiente:

$$G(s)=\log M(s)=\log\left[1+\sum_{k=1}^{\infty}\frac{s^k}{k!}E(X^k)\right]$$

$$=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\left(\sum_{k=1}^{\infty}\frac{s^k}{k!}E(X^k)\right)^n$$

cuando la anterior se expande, se puede efectuar un rearreglo con respecto a las potencias S^K, y se puede poner de la forma:

$$\mathcal{C}(S) = \sum_{\lambda=1}^{\infty} \frac{c_\lambda}{\lambda} S^\lambda$$

que también es válida para $-S_0 < S < S_0$

los c_λ son llamados los acumulantes de X, ahora si igualamos coeficientes de las expresiones:

$$\mathcal{C}(S) = \sum_{\lambda=1}^{\infty} \frac{c_\lambda}{\lambda} S^\lambda$$

y $$\mathcal{C}(S) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \left(\sum_{K=1}^{\infty} \frac{m_K}{K!} S^K \right)^n = \sum_{K=1}^{\infty} \frac{m_K}{K!} S^K - \frac{1}{2} \left(\sum_{K=1}^{\infty} \frac{m_K}{K!} S^K \right)^2 + \frac{1}{3} \left(\sum_{K=1}^{\infty} \frac{m_K}{K!} S^K \right)^3 - \cdots$$

observamos: $\frac{c_1}{1!} = \frac{m_1}{1!} \Rightarrow c_1 = m_1$ igualmente

$$\frac{c_2}{2!} S^2 = \frac{m_2}{2!} S^2 - \frac{1}{2} (m_1^2 S^2)$$

$$= \tfrac{1}{2} S^2 (m_2 - m_1^2) \Rightarrow c_2 = m_2 - m_1^2$$

que concuerda con

$$\mathcal{C}'(0) = E(X)$$

$$\mathcal{C}''(0) = Var(X)$$

Siguiendo este procedimiento se observa que cada c_λ está expresado como un polinomio en $m_1, m_2, \ldots, m_\lambda, \ldots$.

A continuación un último Teorema sobre esta Función Generatriz, que nos proporciona una forma de calcular la Función Generatriz Acumulante de una Variable Aleatoria Y que es suma de un número finito de Variables Aleatorias Independientes.

<u>Teorema 8.</u> Si $X_1, X_2, \ldots, X_n$ son Variables Aleatorias Independientes, entonces la Función Generatriz Acumulante de $Y = X_1 + \cdots + X_n$ es la suma de las Funciones Generatrices Acumulantes de cada una de las X_i.

<u>Demostración</u>. Por el Teorema 8 de este capítulo vemos que $Y = X_1 + X_2 + \cdots + X_n$ tiene como función generatriz de momentos a $M(s) = M_1(s) M_2(s) \cdots M_n(s)$, donde cada $M_i(s)$ es la Función Generatriz de X_i.
Por lo tanto tomando logaritmos de ambos lados de la anterior igualdad vemos que: $\log M(s) = \log M_1(s) + \log M_2(s) + \cdots + \log M_n(s)$ lo cual queriamos demostrar.

II.3 LA FUNCION GENERATRIZ EN EL CASO GENERAL.

En la definición 5 del capítulo I, se ha definido la Función Generatriz de Momentos $M(s)=E(e^{sX})=\int_{-\infty}^{\infty} e^{sx}dF(x)$ y además hemos visto que está definida sobre algún intervalo que contiene al 0 (pudiendo estar formado por el punto 0 exclusivamente).

En esta sección estudiaremos algunas de sus propiedades.

El primer Teorema nos muestra que si $M(s)$; la Función Generatriz de Momentos de la Variable Aleatoria X existe en una vecindad del 0 entonces los momentos de la Variable Aleatoria X existen, y además $M(s)$ tiene una representación en serie de potencias, en donde los momentos son los coeficientes de esta serie.

<u>*Teorema 9.*</u> *Supongamos que* $M(S)$ *es finita en un intervalo* $[-S_o, S_o]$, $S_o > 0$ *; entonces* $|\alpha_K| < \infty$ *para todo K=1, 2, ... y además* $M(S) = \sum_{K=0}^{\infty} \frac{\alpha_K}{K!} S^K$ *para* $S \in [-S_o, S_o]$ *, en donde*

$$\alpha_K = \int_{-\infty}^{\infty} x^K dF(x)$$

<u>*Demostración.*</u> *Por hipótesis, ya que* $M(S) = \int_{-\infty}^{\infty} e^{sx} dF(x) < \infty$ *sobre el intervalo* $[-S_o, S_o]$ *, las funciones* e^{sx} *y* e^{-sx} *son integrables con respecto a* F *para* $S \in [-S_o, S_o]$ *y por tanto de la desigualdad* $e^{|sx|} \leq e^{sx} + e^{-sx}$ *se obtiene que* $e^{|sx|}$ *es también integrable para* $S \in [-S_o, S_o]$ *es decir* $\int e^{|sx|} dF(x)$ *para* $S \in [-S_o, S_o]$

Ahora bien cuando $S > 0$

$$\lim_{x \to \infty} \frac{e^{S|x|}}{|x|^K} = \infty \quad y \quad \lim_{x \to -\infty} \frac{e^{S|x|}}{|x|^K} = \infty$$

es decir, $|x|^K < c\, e^{S|x|}$ *en una vecindad de* ∞ *y en una vecindad de* $-\infty$.

Por lo tanto la función $|x|^K$ *es integrable.*

Est implica que $|\alpha_K| < \infty$ *, ahora vemos que*

$$e^{sx} = \lim_{n \to \infty} \sum_{K=0}^{n} \frac{s^K x^K}{K!} \quad y \text{ además} \quad \left| \sum_{K=0}^{n} \frac{s^K x^K}{K!} \right| \leq \sum_{K=0}^{n} \frac{|sx|^K}{K!} < e^{|sx|}$$

y $e^{|sx|}$ *es integrable con respecto a* $F(x)$ *para* $|S| \leq S_o$.

Por lo tanto, por el Teorema de la Convergencia Dominada (ver apéndice); tenemos que

$$M(S) = \int_{-\infty}^{\infty} e^{sx} dF(x) = \lim_{n \to \infty} \sum_{K=0}^{n} \frac{S^K}{K!} \int_{-\infty}^{\infty} x^K dF(x)$$

es decir $$M(s)=\sum_{k=0}^{\infty}\frac{\alpha_k}{k!}s^k$$

De este último Teorema se desprende un Corolario, el cual nos da una forma de poder calcular los momentos por medio de derivación secesiva de la Función Generatriz dentro del intervalo de convergencia.

__Corolario 6.__ Suponiendo las mismas hipótesis que en el Teorema anterior, la k-ésima derivada $M^{(k)}(s)$ *de* $M(s)$ *satisface la relación*

$M^{(k)}(0)=\alpha_k$ *, y también es posible hallar la derivada de* $M(s)$ *en cualquier punto* $s\in(-s_0,s_0)$ *, especificamente tenemos lo siguiente:* $$M^{k}(s)=\int_{-\infty}^{\infty}x^k e^{sx}\,dF(x)$$

__Demostración.__ Se probará primero que la función $|x|^k e^{sx}$ *es integrable para* $s\in(-s_0,s_0)$.

Si K=0, obtenemos la función e^{sx} *, la cual es integrable para* $s\in[-s_0,s_0]$.

Supongamos que $|x|^n e^{sx}$ *es integrable para* $s\in(-s_0,s_0)$ *y demostrémoslo para* $n+1$.

Sea $s\in(-s_0,s_0)$ *y tomemos* a *y* b *tales que* $-s_0<a<s<b<s_0$;

sea c *una constante tal que;*

$$|x|\le c\,e^{(b-s)x}\qquad x\ge 0\qquad y$$

$$|x|\le c\,e^{(a-s)x}\qquad x\le 0$$

Entonces vemos

$$|x|^{n+1}e^{sx}\le\begin{cases}c|x|^n e^{bx} & \text{si } x\ge 0\\ c|x|^n e^{ax} & \text{si } x\le 0\end{cases}$$

Pero por hipótesis las funciones $|x|^n e^{bx}$ *y* $|x|^n e^{ax}$ *son integrables, por tanto* $|x|^{n+1}e^{sx}$ *también lo es.*

Pero notemos que $\left| \frac{\partial}{\partial \tau} e^{\tau x} \right|_{\tau = s} = |x| e^{sx}$

y como $|x| e^{sx}$ es integrable, por el Teorema 5 del apéndice vemos que son intercambiables la integral y la derivada, entonces obtenemos

$$M'(s) = \int_{-\infty}^{\infty} \frac{\partial}{\partial \tau} e^{\tau x} \Big|_{\tau = s} dF(x)$$

$$= \int_{-\infty}^{\infty} x e^{sx} dF(x) \qquad \text{para} \quad \tau \in (a, b)$$

y derivando sucesivamente:

$$M^{(K)}(s) = \int_{-\infty}^{\infty} x^K e^{sx} dF(x) \qquad K = I?2, 3, \ldots$$

En particular $M^{(K)}(0) = \int_{-\infty}^{\infty} x^K dF(x) = \alpha_K$

y como observación, vemos que

$$M^{(2)}(s) = \int_{-\infty}^{\infty} x^2 e^{sx} dF(x) \geq 0 \quad , \text{ ya que } \quad x^2 e^{sx} \geq 0$$

es decir $M(s)$ es una función convexa.

El siguiente Teorema da condiciones necesarias y suficientes para que la Función Generatriz $M(S)$ *sea finita en una vecindad del O.*

<u>*Teorema IO.*</u> *Sea* $M(S)$ *una función generatriz y sea* F *la correspondiente función de distribución. Entonces* $M(S)<\infty$ *en una vecindad del O si y sólo si las siguientes condiciones se satisfacen.*

I) F *tiene momentos* α_K *de todos los órdenes.*

II) Existe un número positivo γ *tal que* $|\alpha_K| \le \gamma^K K!$ *para todo* $K \geq 1$.

<u>*Demostración.*</u> $\Rightarrow]$ *Supongamos que* $M(S)<\infty$ *en un intervalo* $[-S_0, S_0]$, $S_0>0$, *o sea* $M(S)$ *es finita en una vecindad del cero, entonces por el Teorema 9 de este capítulo, tenemos que* $|\alpha_K|<\infty$ *para todo K=I,2,3,... y además* $M(S)=\sum_{K=0}^{\infty}\frac{\alpha_K}{K!}S^K$ *para* $S\in[-S_0,S_0]$ *, por lo tanto:*

en primer lugar $|\alpha_K|<\infty$ *para todo K=I,2,3,... es decir* F *tiene momentos* α_K *de todos los ordenes, ahora bien como* $M(S)<\infty$ *para todo* $S\in[-S_0,S_0]$ *, y además como* $M(S)=\sum_{K=0}^{\infty}\frac{\alpha_K}{K!}S^K$ *entonces vemos que* $\frac{\alpha_K}{K!}S^K \to 0$ *cuando* $K\to\infty$ *para cada* $S\in[-S_0,S_0]$

Entonces la sucesión $\frac{\alpha_K}{K!}S^K$ *está acotada, es decir existe* $N(S)>0$ *tal que:* $\left|\frac{\alpha_K}{K!}S^K\right|<N(S)$ $\forall$ K *y entonces*

$$\left|\frac{\alpha_K}{K!}S^K\right| = \left|\frac{\alpha_K}{K!}\right||S|^K \le \left|\frac{\alpha_K}{K!}\right||S_0|^K < N(S_0) \qquad \forall \quad S\in[-S_0,S_0]$$

Sea $N=N(S_0)$, *como siempre podemos tomar a* $N(S_0)>1$ *entonces* $N\le N^K$ *para todo K=I,2,3,...*

Por tanto $|\alpha_K| < \frac{N\,K!}{|S_0|^K} \le \left(\frac{N}{|S_0|}\right)^K K!$.

y claramente $\frac{N}{|S_0|}>0$

$\Leftarrow$ *Ahora bien en esta parte de la demostración tenemos que si X es una variable aleatoria arbitraria y si además suponemos que:*

I) Todos los momentos de X existen, es decir

$$E\left[|X|^n\right] = \int_{-\infty}^{\infty} |x|^n dF(x) < \infty$$

II) Sea $\alpha_k = \int_{-\infty}^{\infty} x^k dF(x)$ *y supongamos:* $|\alpha_k| \leq k!\,\gamma^k \quad \forall\, k$

donde γ *es una constante estrictamente positiva.*

Entonces queremos demostrar que la función generatriz de X existe en una vecindad del 0.

Sea $\phi(s) = E\left(e^{sx}\right)$ *en donde* $E\left(e^{sx}\right) = \int_{-\infty}^{\infty} e^{sx} dF(x)$

Dividamos la integral en 3 partes:

$$E(e^{sx}) = \int_{-\infty}^{-1} e^{sx} dF(x) + \int_{-1}^{1} e^{sx} dF(x) + \int_{1}^{\infty} e^{sx} dF(x)$$

La segunda integral es finita, solo hay que demostrar que las otras dos lo son también:

Sea $A > 1$

En el intervalo $[1, A]$ *la función* e^{sx} *es límite uniforme de* $f_n(x)$ *donde* $f_n(x) = \sum_{k=0}^{n} \frac{s^k x^k}{k!}$

Por lo tanto

$$\int_{1}^{A} e^{sx} dF(x) = \lim \int_{1}^{A} \sum_{k=0}^{n} \frac{s^k x^k}{k!}\, dF(x)$$

$$= \sum_{k=0}^{\infty} \int_{1}^{A} \frac{s^k x^k}{k!}\, dF(x)$$

También

$$\int_{-A}^{-1} e^{sx} dF(x) = \sum_{K=0}^{\infty} \int_{-A}^{-1} \frac{s^K x^K}{K!} dF(x)$$

$$\int_{-A}^{A} e^{sx} dF(x) = \int_{-1}^{1} e^{sx} dF(x) + \sum_{K=0}^{\infty} \left[\int_{-A}^{-1} \frac{s^K x^K}{K!} dF(x) + \int_{1}^{A} \frac{s^K x^K}{K!} dF(x) \right]$$

Si K es par:

$$\int_{1}^{A} \frac{s^K x^K}{K!} dF(x) + \int_{-A}^{-1} \frac{s^K x^K}{K!} dF(x)$$

$$\leq \int_{-\infty}^{\infty} \frac{s^K x^K}{K!} dF(x) = \frac{s^K \alpha_K}{K!} \leq s^K \gamma^K$$

si K es impar:

$$\left| \int_{-A}^{-1} \frac{s^K x^K}{K!} dF(x) + \int_{1}^{A} \frac{s^K x^K}{K!} dF(x) \right|$$

$$\leq \int_{-A}^{-1} \frac{s^K |x|^K}{K!} dF(x) + \int_{1}^{A} \frac{s^K x^K}{K!} dF(x)$$

$$\leq \int_{-A}^{-1} \frac{s^K |x|^{K+1}}{K!} dF(x) + \int_{1}^{A} \frac{s^K x^{K+1}}{K!} dF(x)$$

$$= \int_{-A}^{-1} \frac{s^K x^{K+1}}{K!} dF(x) + \int_{1}^{A} \frac{s^K x^{K+1}}{K!} dF(x)$$

$$\leq \int_{-\infty}^{\infty} \frac{s^K x^{K+1}}{K!} dF(x) = \frac{s^K}{K!} \alpha_{K+1} \leq \frac{s^K (K+1)! \, \gamma^{K+1}}{K!}$$

$$= (K+1) s^K \gamma^{K+1}$$

Por lo tanto $$\left| \int_{-A}^{A} e^{sx} dF(x) \right| \leq \int_{-1}^{1} e^{sx} dF(x) + \sum_{k=0}^{\infty} s^k \gamma^k + \sum_{k=1}^{\infty} (k+1) s^k \gamma^{k+1}$$

$\sum_{k=0}^{\infty} s^k \gamma^k$ *es finita para* $|s| < 1/\gamma$

si $|s| < 1/\gamma$ *la serie es diferenciable término a término:*

$$\therefore \sum_{k=0}^{\infty} k s^{k-1} \gamma^k < \infty \qquad \forall \quad |s| < 1/\gamma$$

es decir:

$$\sum_{k=0}^{\infty} (k+1) s^k \gamma^{k+1} < \infty$$

concluimos que:

Para cada $|s| < 1/\gamma$ *existe* $\Theta(s)$ *tal que*

$$\int_{-A}^{A} e^{sx} dF(x) < \Theta(s) < \infty \qquad \forall \ A$$

$$\therefore \int_{-\infty}^{\infty} e^{sx} dF(x) \leq \Theta(s) < \infty \qquad \forall \quad |s| < 1/\gamma .$$

<u>*Teorema II.*</u> *Si la Función Generatriz de Momentos* $M(S)$ *existe en una vecindad del 0, entonces ésta determina a la distribución.*

<u>*Demostración.*</u> *Supongamos que* $F(x)$ *tiene una Función Generatriz de Momentos* $M(S)$ *para* $S \in [-S_0, S_0]$, $S_0 > 0$; *ahora bien ya se ha demostrado en el Teorema 9, que:*

$$M(S) = \sum_{K=0}^{\infty} \frac{S^K}{K!} E(X^K) = \sum_{K=0}^{\infty} \frac{S^K}{K!} \int_{-\infty}^{\infty} x^K dF(x) \quad , \quad |S| \leq S_0 .$$

entonces vemos que las condiciones del Teorema 8 , del apéndice se satisfacen , y así de esta manera $F(x)$ *está determinada por sus momentos, que a su vez éstos están determinados por* $M(S)$ *por medio de* $M^{(K)}(0) = E(X^K) = \int_{-\infty}^{\infty} x^K dF(x)$ *, y de esta manera* $F(x)$ *está determinada por* $M(S)$.

CAPITULO III

CONVERGENCIA DE FUNCIONES DE DISTRIBUCION

En este capítulo veremos algunos resultados que nos servirán para estudiar otras propiedades de la Función Generatriz.

Definición I. *Una función* F *definida sobre la recta real es una función de distribución si:*

I) F *es no decreciente*

II) F *es continua por la derecha.*

III) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0 \qquad F(\infty) = \lim_{x \to \infty} F(x) < \infty$

Diremos que F *es una distribución de probabilidad si además* $F(\infty)=1$

A toda función de distribución F *podemos asociarle una medida acotada* μ *sobre* $\mathbb{R}$ *mediante la siguiente relación:*

$$\mu\big((a,b]\big) = F(b) - F(a)$$

Inversamente, toda medida acotada μ *sobre* $\mathbb{R}$ *define una función de distribución cuya medida asociada es* μ *, la cual está dada por la siguiente relación:*

$$F(x) = \mu\big((-\infty, x]\big)$$

Claramente esta correspondencia es uno a uno.

Sea F *una función de distribución,* μ *su medida asociada y* g *una función medible definida sobre* $\mathbb{R}$ *. Las siguientes notaciones serán utilizadas en adelante:*

$$E(g) = \int_{-\infty}^{\infty} g(x)\, dF(x) = \int_{-\infty}^{\infty} g(x)\, \mu(dx)$$

También, si A *es un conjunto Lebesgue medible, escribiremos:*

$$F\{A\} = \int_A dF(x) = \int_A \mu(dx) = \mu(A) .$$

Definición 2. Sea I un intervalo en $\mathbb{R}$, F una función de distribución y μ su medida asociada.

Diremos que I es un intervalo de continuidad de la distribución F si sus puntos extremos tienen medida μ nula.

La definición anterior es aplicable también cuando I es un intervalo no acotado, en cuyo caso $+\infty$ y $-\infty$ serán considerados como puntos de continuidad de F.

Definición 3. Sea $\{F_n\}$ una sucesión de funciones de distribución. Diremos que F_n converge débilmente a la distribución F si $F_n\{I\} \longrightarrow F\{I\}$ para todo I intervalo de continuidad de F abierto y acotado. En este caso escribiremos:

$$F_n \xrightarrow{d} F \qquad \text{ó} \qquad \lim_{d} F_n = F$$

Diremos que la convergencia es propia o completa si además

$$F_n(\infty) \longrightarrow F(\infty) \quad \text{, es decir si } F_n\{\mathbb{R}\} \longrightarrow F\{\mathbb{R}\} .$$

Obsérvese en particular que si $\{F_n\}$ es una sucesión de distribuciones de probabilidad y $F_n \xrightarrow{d} F$ entonces la convergencia es propia si y solo si F es una distribución de probabilidad.

Si F es una función de distribución, por ser una función no decreciente su conjunto de puntos de discontinuidad es a lo más numerable, de aquí se sigue que el conjunto de puntos de continuidad de F forma un conjunto denso en $\mathbb{R}$ y como F es continua por la derecha, F queda determinada entonces por sus valores en sus puntos de continuidad. Usaremos estas propiedades de una función de distribución en lo que sigue.

Demostraremos primero la unicidad del límite.

Proposición I. Sea $\{F_n\}$ una sucesión de funciones de distribución. Si $\{F_n\}$ converge débilmente entonces el límite es único.

Demostración. Supongamos
$$F_n \xrightarrow{d} F$$
$$F_n \xrightarrow{d} G$$

Sea entonces b un punto de continuidad de F y de G y sea $\{a_n\}$ una sucesión de puntos de continuidad de F y de G tal que $a_n \to -\infty$, $a_n \leq b$ $\forall n$, entonces

$$F(b) - F(a_n) = F\{(a_n, b)\} = G\{(a_n, b)\} = G(b) - G(a_n) \quad \forall n$$

De manera que tomando límites cuando $n \to \infty$ obtenemos:

$$F(b) = G(b)$$

Y como el conjunto de puntos de continuidad de F y de G forma un conjunto denso, el resultado se sigue de la continuidad por la derecha de F y G .

Lema I. $\{F_n\}$ es una sucesión de funciones de distribución que converge débilmente a la función de distribución F . Si μ_n es la medida asociada a F_n , se tiene entonces:

$$\mu_n(\{x\}) \to 0 \quad \forall x \text{ punto de continuidad de } F.$$

Demostración. Sea x punto de continuidad de F , sabemos que dado $\varepsilon > 0$ existe una vecindad $V(x)$ de x tal que:

$$|F(b) - F(a)| < \varepsilon \qquad \forall\, a, b \in V(x)$$

Ahora bien, el conjunto de discontinuidad de F es a lo más numerable, podemos tomar $a, b \in V(x)$ de manera que sean puntos de continuidad de F y tales que $a < x < b$.

Entonces:

$$F_n\{(a,b)\} \to F\{(a,b)\} = F\{(a,b]\} = F(b) - F(a)$$

Así que, existe N tal que:

$$F_n\{(a,b)\} < F(b) - F(a) + \varepsilon \qquad \forall\ n \geq N$$

Pero entonces:

$$\mu_n(\{x\}) \leq F_n\{(a,b)\} < F(b) - F(a) + \varepsilon < 2\varepsilon \qquad \forall\ n \geq N$$

$$\therefore\ \mu_n(\{x\}) \longrightarrow 0$$

Corolario I. Sea $\{F_n\}$ una sucesión de funciones de distribución, entonces $F_n \xrightarrow{d} F$ si y solo si $F_n\{I\} \longrightarrow F\{I\}$ para todo I intervalo de continuidad de F acotado.

Proposición 8. Sea $\{F_n\}$ una sucesión de funciones de distribución, entonces $F_n \xrightarrow{d} F$ propiamente si y solo si $F_n\{I\} \longrightarrow F\{I\}$ para todo I intervalo de continuidad de F acotado o no acotado.

Demostración.

$\Rightarrow]$ Supongamos $F_n \xrightarrow{d} F$ propiamente y sea x un punto de continuidad de F.

Vamos a demostrar que:

$$F_n\{(-\infty,x)\} \longrightarrow F\{(-\infty,x)\}$$

$$F_n\{(x,\infty)\} \longrightarrow F\{(x,\infty)\}$$

Con lo cual quedará probado el resultado pues como $F_n\{\{x\}\} \longrightarrow 0$ $\forall\ x$ punto de continuidad de F, las mismas relaciones se tendrán para intervalos infinitos cerrados en x.

Sea $M = F\{\mathbb{R}\}$ $\qquad M_n = F_n\{\mathbb{R}\}$

Dado $\varepsilon > 0$ tomemos $b > |x|$ tal que $-b$ y b sean puntos de continuidad de F tales que

$$F\{(-b,b)\} > M - \frac{\varepsilon}{2}$$

Como $F_n\{(-b,b)\} \longrightarrow F\{(-b,b)\}$ existen entonces N_1 tales que: $F_n\{(-b,b)\} > M - \frac{\varepsilon}{2} \quad \forall \ n \geq N_1$

como $M_n \rightarrow M$ existe N_2 tal que:

$$M > M_n - \frac{\varepsilon}{2} \qquad \forall \ n \geq N_2.$$

Entonces, si $N = \text{máx}\{N_1, N_2\}$, obtenemos:

$$F_n\{(-b,b)\} > M_n - \varepsilon \qquad \forall \ n \geq N$$

$$\therefore \quad F\{(-\infty,-b)\} < \varepsilon \qquad F_n\{(-\infty,-b)\} < \varepsilon \quad \forall \ n \geq N$$

$$F\{(b,\infty)\} < \varepsilon \qquad F_n\{(b,\infty)\} < \varepsilon \quad \forall \ n \geq N$$

Además:

$$F_n\{(-\infty,x)\} = F_n\{(-\infty,-b)\} + F_n\{(-b,x)\}$$

$$F\{(-\infty,x)\} = F\{(-\infty,-b)\} + F\{(-b,x)\}$$

$$F_n\{(x,\infty)\} = F_n\{(x,b)\} + F_n\{(b,\infty)\}$$

$$F\{(x,\infty)\} = F\{(x,b)\} + F\{(b,\infty)\}$$

Tomando ahora N suficientemente grande de manera que también se tenga:

$$\left| F_n\{(-b,x)\} - F\{(-b,x)\} \right| < \varepsilon \qquad \forall \quad n \geq N$$

$$\left| F_n\{(x,b)\} - F\{(x,b)\} \right| < \varepsilon \qquad \forall \quad n \geq N$$

Obtenemos:

$$\left| F_n\{(-\infty,x)\} - F\{(-\infty,x)\} \right| \leq \left| F_n\{(-\infty,-b)\} \right| + \left| F\{(-\infty,-b)\} \right|$$
$$+ \left| F_n\{(-b,x)\} - F\{(-b,x)\} \right|$$
$$< 3\,\varepsilon \qquad \forall \quad n \geq N$$

$$\left| F_n\{(x,\infty)\} - F\{(x,\infty)\} \right| \leq \left| F_n\{(b,\infty)\} \right| + \left| F\{(b,\infty)\} \right|$$
$$+ \left| F_n\{(x,b)\} - F\{(x,b)\} \right|$$
$$< 3\,\varepsilon \qquad \forall \quad n \geq N.$$

De lo cual se concluye el resultado buscado

$\Leftarrow]$ *Si* $F_n\{I\} \longrightarrow F\{I\}$ *para todo intervalo de continuidad de* F *acotado o no acotado se tiene directamente de la definición que* $F_n \xrightarrow{d} F$ *y como en particular* $F_n\{\mathbb{R}\} \longrightarrow F\{\mathbb{R}\}$ *la convergencia es propia.*

Como veremos más adelante, el siguiente concepto está relacionado con la convergencia propia.

Definición 4. Sea $\{F_n\}$ una sucesión de funciones de distribución y sea $M_n = F_n\{\mathbb{R}\}$. Diremos que la sucesión $\{F_n\}$ es uniforme si se satisfacen las siguientes dos condiciones:

I) $F_n(X) \longrightarrow 0$ uniformemente en n cuando $X \longrightarrow -\infty$

II) $M_n - F_n(X) \longrightarrow 0$ uniformemente en n cuando $X \longrightarrow \infty$

La siguiente proposición muestra que el concepto de sucesión uniforme de funciones de distribución corresponde a lo que se conoce como sucesión "tight".

Proposición 5. Sea $\{F_n\}$ una sucesión de funciones de distribución y sea $M_n = F_n\{\mathbb{R}\}$. Las siguientes condiciones son equivalentes:

I) $\{F_n\}$ es uniforme.

II) Para cada $\varepsilon > 0$ existe un intervalo acotado $(a, b]$ tal que:

$$F_n\{(a,b]\} = F_n(b) - F_n(a) > M_n - \varepsilon \quad \forall\, n$$

Demostración.

I) $\Rightarrow$ II). Supongamos que $\{F_n\}$ es uniforme, entonces dado $\varepsilon > 0$ existen $a < 0$ y $b > 0$ tales que:

$$F_n(X) < \frac{\varepsilon}{2} \quad \forall \; x \leq a \quad \forall\, n$$

$$M_n - F_n(X) < \frac{\varepsilon}{2} \quad \forall \; x \geq b \quad \forall\, n$$

Por lo tanto:

$$F_n\{(a,b]\} = F_n(b) - F_n(a) > M_n - \varepsilon \quad \forall\, n$$

II) $\Rightarrow$ I). Si (II) se satisface, dado $\varepsilon > 0$ tomemos un intervalo $(a,b]$ tal que $F_n\{(a,b]\} > M_n - \varepsilon \quad \forall\, n$

Entonces

$$F_n(x) < \varepsilon \qquad \forall \quad x \leq a \qquad \forall \ n.$$

$$F_n(x) > M_n - \varepsilon \qquad \forall \ x \geq b \quad \forall \ n.$$

Así que $F_n(x) \longrightarrow 0$ *uniformemente en* n *cuando* $x \to -\infty$ *y* $M_n - F_n(x) \to 0$ *uniformemente en* n *cuando* $x \to \infty$ *; es decir* $\{F_n\}$ *es uniforme.*

<u>Proposición 4.</u> Sean F_n $(n=1,2,\ldots)$ *y* F *funciones de distribución y supongamos* $F_n \xrightarrow{d} F$ *, entonces las siguientes condiciones son equivalentes:*

I) F_n *converge a* F *propiamente.*

IIO La sucesión $\{F_n\}$ *es uniforme.*

<u>Demostración,</u>

$\mathrm{I}) \Rightarrow \mathrm{II})$ *Supongamos que* F_n *converge a* F *propiamente, entonces dado* $\varepsilon > 0$ *tomemos un intervalo de continuidad de* F *de la forma* $(a, b]$ *y tal que:*

$$F\{(a,b]\} = F(b) - F(a) > F\{\mathbb{R}\} - \frac{\varepsilon}{3}$$

Como $F_n \xrightarrow{d} F$ *propiamente, podemos tomar ahora* N *tal que*

$$\left| F_n\{(a,b]\} - F\{(a,b]\} \right| < \frac{\varepsilon}{3} \qquad \forall \ n \geq N$$

$$\left| F_n\{\mathbb{R}\} - F\{\mathbb{R}\} \right| < \frac{\varepsilon}{3} \qquad \forall \quad n \geq N$$

Por lo tanto:

$$F_n\{\mathbb{R}\} - F_n\{(a,b]\} \leq |F_n(\mathbb{R}) - F(\mathbb{R})| + |F(\mathbb{R}) - F\{(a,b]\}| + |F\{(a,b]\} - F_n\{(a,b]\}| < \varepsilon \quad \forall n \geq N$$

Es decir:

$$F_n\{(a,b]\} > F_n\{\mathbb{R}\} - \varepsilon \qquad \forall \; n \geq N$$

Ahora para $J = 1, 2, \ldots, N-1$ *tomemos intervalos* $(a_J, b_J]$ *tales que:*

$$F_J\{(a_J, b_J]\} > F_J\{\mathbb{R}\} - \varepsilon$$

Entonces si $c = \text{mín}\{a, a_1, a_2, \ldots, a_{N-1}\}$ *y* $d = \text{máx}\{b, b_1, b_2, \ldots, b_{N-1}\}$ *se tiene:*

$$F_n\{(c,d]\} > F_n\{\mathbb{R}\} - \varepsilon \qquad \forall \; n$$

Así que la por la proposición 3, la sucesión $\{F_n\}$ *es uniforme.*

II) ⇒ I) *Supongamos ahora que la sucesión* $\{F_n\}$ *es uniforme. Por la proposición 3, para cada* $\varepsilon > 0$ *existe un intervalo* $(a, b]$ *tal que:*

$$F_n\{(a,b]\} = F_n(b) - F_n(a) > F_n\{\mathbb{R}\} - \frac{\varepsilon}{2} \qquad \forall \; n$$

Podemos tomar el intervalo $(a, b]$ *de manera que* a *y* b *sean puntos de continuidad de* F *y entonces como* $F_n \xrightarrow{d} F$ *se tiene:*

$$F_n\{(a,b]\} \longrightarrow F\{(a,b]\}$$

Así que existe N_1 *tal que:*

$$\left| F_n\{(a,b]\} - F\{(a,b]\} \right| < \frac{\varepsilon}{2} \qquad \forall \quad n \geq N_1$$

En particular:

$$F\{(a,b]\} > F_n\{(a,b]\} - \frac{\varepsilon}{2} \qquad \forall \quad n \geq N_1$$

Por lo tanto:

$$F\{\mathbb{R}\} \geq F\{(a,b]\} > F_n\{(a,b]\} - \frac{\varepsilon}{2} > F_n(\mathbb{R}\} - \varepsilon \qquad \forall \; n \geq N_1$$

Tomemos ahora un intervalo de continuidad de F *de la forma* $(c,d]$ *y tal que:*

$$F\{(c,d]\} = F(d) - F(c) > F\{\mathbb{R}\} - \frac{\varepsilon}{2}$$

Nuevamente, como $F_n \xrightarrow{d} F$, *existe* N_2 *tal que:*

$$\left| F_n\{(c,d]\} - F\{(c,d]\} \right| < \frac{\varepsilon}{2} \qquad \forall \quad n \geq N_2$$

En particular:

$$F\{(c,d]\} < F_n\{(c,d]\} + \frac{\varepsilon}{2} \qquad \forall \quad n \geq N_2$$

Y entonces

$$F\{\mathbb{R}\} < F\{(c,d]\} + \frac{\varepsilon}{2} < F_n\{(c,d]\} + \varepsilon \leq F_n\{\mathbb{R}\} + \varepsilon$$
$$\forall \; n \geq N_2$$

Así que si $N = \text{máx}\,\{N_1, N_2\}$ *tenemos:*

$$F\{\mathbb{R}\} > F_n\{\mathbb{R}\} - \varepsilon \qquad \forall \; n \geq N$$

$$F\{\mathbb{R}\} < F_n\{\mathbb{R}\} + \varepsilon \qquad \forall \; n \geq N$$

Es decir: $\left| F_n\{\mathbb{R}\} - F\{\mathbb{R}\} \right| < \varepsilon \qquad \forall \; n \geq N$

Así que $F_n\{\mathbb{R}\} \longrightarrow F\{\mathbb{R}\}$, *es decir la convergencia de* F_n a F *es propia.*

En ocasiones se habla de convergencia de funciones de distribución en el siguiente sentido:

$F_n \longrightarrow F$ si $F_n(x) \longrightarrow F(x)$ $\forall x$ punto de continuidad de F.

En general se usa esta definición cuando cada F_n como F son funciones de distribución de variables aleatorias, en cuyo caso se dice que la correspondiente sucesión de variables aleatorias converge en distribución.

En este caso esta definición de convergencia y la que dimos anteriormente coinciden como veremos más adelante. Sin embargo conviene precisar la relación que existe entre ellas en el caso general.

Lema 2. Sean F_n $(n=1,2,\dots)$ funciones de distribución y F una función no decreciente y acotada y μ_n la medida asociada a F_n.

Si $F_n(x) \longrightarrow F(x)$ $\forall x$ punto de continuidad de F entonces $\mu_n(\{x\}) \longrightarrow 0$ $\forall x$ punto de continuidad de F.

Demostración. Sea x un punto de continuidad de F.

Dado $\varepsilon > 0$ sea $V(x)$ una vecindad de x tal que:

$$|F(x) - F(y)| < \varepsilon \qquad \forall y \in V(x)$$

Tomemos $a, b \in V(x)$ de manera que sean puntos de continuidad de F y tales que $a < x < b$, entonces:

$$F_n\{(a,b]\} = F_n(b) - F_n(a) \longrightarrow F(b) - F(a)$$

Así que existe N tal que:

$$F_n\{(a,b]\} < F(b) - F(a) + \varepsilon \qquad \forall n \geq N$$

Pero entonces:

$$\mu_n(\{x\}) \leq F_n\{(a,b]\} < F(b) - F(a) + \varepsilon < 2\varepsilon \qquad \forall n \geq N$$

$$\therefore \quad \mu_n(\{x\}) \longrightarrow 0$$

<u>Proposición 5.</u> Sea $\{F_n\}$ una sucesión de funciones de distribución y supongamos $F_n(x) \longrightarrow F(x) \ \forall x$ punto de continuidad de la función de distribución F, entonces $F_n \xrightarrow{d} F$.

<u>Demostración.</u> Sea μ_n la medida generada por F_n y sea (a, b) un intervalo de continuidad de F, tenemos entonces:

$$F_n\{(a,b)\} = F_n\{(a,b]\} - \mu_n(\{b\})$$

$$= F_n(b) - F_n(a) - \mu_n(\{b\})$$

Pero $F_n(b) - F_n(a) \longrightarrow F(b) - F(a)$ por hipótesis

y $\mu_n(\{b\}) \longrightarrow 0$ por el lema anterior.

$$\therefore \quad F_n\{(a,b)\} \longrightarrow F(b) - F(a) = F\{(a, b]\}$$

Finalmente, $F\{(a,b]\} = F\{(a,b)\}$ ya que b es punto de continuidad de F.

El inverso de la proposición anterior no es válido, como se muestra con el ejemplo siguiente:

<u>Ejemplo.</u>

Definamos:

$$F_{2n}(x) = \begin{cases} 0 & \text{si } x < -2n \\ 1 & \text{si } x \geq -2n \end{cases}$$

$$F_{2n-1}(x) = \begin{cases} 0 & \text{si } x < 2n-1 \\ 1 & \text{si } x \geq 2n-1 \end{cases}$$

para $x \in \mathbb{R}$ y $n = 1, 2, 3, \ldots$

F_{a_n} y $F_{a_{n-1}}$ representan medidas de probabilidad concentradas en $-a_n$ y a_{n-1} respectivamente.

Entonces, cualquiera que sea el intervalo (a,b) $F_n(a,b) \to 0$.

Es decir, $F_n \xrightarrow{d} F$ donde F es la función identicamente cero.

Por otra parte:

$$F_{a_n}(x) \to 1 \quad \forall\, x \in \mathbb{R}$$

$$F_{a_{n-1}}(x) \to 0 \quad \forall\, x \in \mathbb{R}$$

$\therefore$ $F_n(x)$ no converge para ningún valor de $x \in \mathbb{R}$.

Los siguientes resultados demuestran en particular que cuando la convergencia es propia ambos tipos de convergencia coinciden.

<u>Proposición 6.</u> Sea $\{F_n\}$ una sucesión de funciones de distribución. Las siguientes condiciones son equivalentes:

I) $F_n \xrightarrow{d} F$ y $F_n(x) \to 0$ uniformemente en n cuando $x \to -\infty$

II) $F_n(x) \to F(x)$ $\forall\, x$ punto de continuidad de F.

<u>Demostración.</u> $(I) \Rightarrow (II)$. Supongamos (I) y sea x_0 punto de continuidad de F.

Dado $\varepsilon > 0$ tomemos a punto de continuidad de F tal que $a < x_0$ y tal que:

$$F_n(x) < \frac{\varepsilon}{3} \quad \forall\, x \leq a \quad \forall\, n$$

$$F(x) < \frac{\varepsilon}{3} \quad \forall\, x \leq a$$

Como $F_n \xrightarrow{d} F$ existe N tal que:

$$\left| F_n\{(a, x_0]\} - F\{(a, x_0]\} \right| < \frac{\varepsilon}{3} \quad \forall\, n \geq N$$

Entonces:

$$\left| F_n(x_0) - F(x_0) \right| = \left| F_n(a) + F_n\{(a,x_0]\} - F(a) - F\{(a,x_0]\} \right|$$

$$\leq F_n(a) + F(a) + \left| F_n\{(a,x_0]\} - F\{(a,x_0]\} \right|$$

$$< \varepsilon \qquad \forall \quad n \geq N$$

Así que $F_n(x_0) \longrightarrow F(x_0)$

II) $\Rightarrow$ *(I)*

Ya probamos en la proposición 5 que la condición (II) *implica*

$F_n \xrightarrow{d} F$

Para probar que $F_n(x) \rightarrow 0$ *uniformemente en* n *cuando* $x \rightarrow -\infty$ *, dada* $\varepsilon > 0$ *tomemos* a *punto de continuidad de* F *tal que:*

$$F(a) < \frac{\varepsilon}{2}$$

Usando (II) *existe entonces* N *tal que:*

$$\left| F_n(a) - F(a) \right| < \frac{\varepsilon}{2} \qquad \forall \quad n \geq N$$

$$\therefore \quad F_n(a) < F(a) + \frac{\varepsilon}{2} < \varepsilon \qquad \forall \quad n \geq N$$

Para $J = 1, 2, \ldots, N-1$ *tomemos ahora* a_J *tal que:*

$$F_n(a_J) < \varepsilon$$

Entonces si $b = \min\{a, a_1, a_2, \ldots, a_{N-1}\}$ *tenemos:*

$$F_n(b) < \varepsilon \qquad \forall \; n$$

$$\therefore \quad F_n(x) < \varepsilon \qquad \forall \; x \leq b \qquad \forall \; n$$

Así que $F_n(x) \longrightarrow 0$ *uniformemente en* n *cuando* $x \rightarrow -\infty$.

Entonces concluimos que en general la convergencia débil es menos fuerte que la convergencia en el otro sentido; pero también tenemos el siguiente corolario:

<u>*Corolario8.*</u> *Sean* F_n $(n=1,2,\ldots)$ *y* F *funciones de distribución y supongamos que* $F_n\{\mathbb{R}\} \to F\{\mathbb{R}\}$ *entonces las siguientes dos condiciones son equivalentes:*

a) $F_n \xrightarrow{d} F$

b) $F_n(x) \to F(x)$ $\forall x$ *punto de continuidad de* F.

<u>*Demostración.*</u> *Supongamos (a),* $F_n \xrightarrow{d} F$ *propiamente, así que por la proposición 4, la sucesión* $\{F_n\}$ *es uniforme y entonces se obtiene (b) por la proposición 6.*

Finalmente (b) $\Rightarrow$ *(a) está dada por la proposición 5.*

En seguida caracterizaremos la convergencia débil de una manera que nos será util más adelante.

Teorema I. *Sean* $F_n\ (n=1,2,\dots)$ *y* F *funciones de distribución entonces las siguientes dos condiciones son equivalentes:*

a) $F_n \xrightarrow{d} F$ *propiamente.*

b) $\int_{\mathbb{R}} g(x)\, dF_n(x) \longrightarrow \int_{\mathbb{R}} g(x)\, dF(x)$ $\forall$ *función continua y acotada* g.

Demostración.

Supongamos (a) y sea g *una función continua y acotada definida sobre* $\mathbb{R}$.

Sea $M = F\{\mathbb{R}\}$ $M_n = F_n\{\mathbb{R}\}$ *y* K *una cota de superior* $|g|$

Dado $\varepsilon>0$ *tomemos* $a>0$ *tal que:*

$$F\{[-a,a]\} > M-\varepsilon$$

Entonces $F\{(-\infty,-a)\} + F\{(a,\infty)\} < \varepsilon$

Pero como $F_n \xrightarrow{d} F$ *propiamente:*

$$F_n\{(-\infty,-a)\} + F_n\{(a,\infty)\} \longrightarrow F\{(-\infty,-a)\} + F\{(a,\infty)\}$$

Por lo tanto, existe N_1 *tal que:*

$$F_n\{(-\infty,-a)\} + F_n\{(a,\infty)\} < \varepsilon \quad \forall\, n \geq N_1$$

De aquí se siguen las siguientes desigualdades:

$$\left| \int_{[-a,a]^c} g(x)\, dF_n(x) \right| \leq K\left(F_n\{(-\infty,-a)\} + F_n\{(a,\infty)\}\right) < K\varepsilon \quad \forall\, n \geq N_1$$

$$\left| \int_{[-a,a]^c} g(x)\, dF(x) \right| \leq K\left(F\{(-\infty,-a)\} + F\{(a,\infty)\}\right) < K\varepsilon ,$$

Ahora bien:

g *es una función continua en* $[-a,a]$ *, por lo tanto es uniformemente continua: consideremos entonces una partición de* $[-a,a]$ *en intervalos ajenos de continuidad* $I_1, I_2, \dots, I_r$ *en cada uno de los cuales la oscilación de* g *sea menor que* ε.

Sea entonces $x_J \in I_J$ *y definamos la función* $T: [-a,a] \longrightarrow \mathbb{R}$ *de la siguiente manera:* $T(x) = g(x_J)$ si $x \in I_J$

T *es entonces una función escalonada con la propiedad:*

$$|T(x) - g(x)| < \varepsilon \qquad \forall\ x \in [-a,a]$$

Por lo tanto, tenemos:

$$\left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} T(x)\, dF_n(x) \right| \le \int_{[-a,a]} |g(x) - T(x)|\, dF_n(x) < \varepsilon F_n\{[-a,a]\} \le \varepsilon M_n$$

$$\left| \int_{[-a,a]} g(x)\, dF(x) - \int_{[-a,a]} T(x)\, dF(x) \right| \le \int_{[-a,a]} |g(x) - T(x)|\, dF(x) < \varepsilon F\{[-a,a]\} \le \varepsilon M$$

Además:

$$\int_{[-a,a]} T(x)\, dF_n(x) = \sum_{J=1}^{r} x_J F_n\{I_J\} \longrightarrow \sum_{J=1}^{r} x_J F\{I_J\} = \int_{[-a,a]} T(x)\, dF(x)$$

$\therefore$ *existe* N_2 *tal que:*

$$\left| \int_{[-a,a]} T(x)\, dF_n(x) - \int_{[-a,a]} T(x)\, dF(x) \right| < \varepsilon \qquad \forall\ n \ge N_2$$

Combinando las desigualdades anteriores, obtenemos entonces:

$$\left| \int_{\mathbb{R}} g(x)\, dF_n(x) - \int_{\mathbb{R}} g(x)\, dF(x) \right| \leq \left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} g(x)\, dF(x) \right|$$

$$+ \left| \int_{[-a,a]^c} g(x)\, dF_n(x) \right| + \left| \int_{[-a,a]^c} g(x)\, dF(x) \right| < \left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} g(x)\, dF(x) \right| + 2K\varepsilon$$

$$\forall \quad n \geq N_1$$

Y también:

$$\left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} g(x)\, dF(x) \right| \leq \left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} T(x)\, dF_n(x) \right|$$

$$+ \left| \int_{[-a,a]} T(x)\, dF_n(x) - \int_{[-a,a]} T(x)\, dF(x) \right|$$

$$+ \left| \int_{[-a,a]} T(x)\, dF(x) - \int_{[-a,a]} g(x)\, dF(x) \right|$$

$$< \varepsilon M_n + \varepsilon M + \varepsilon \quad \forall \ n \geq N_2$$

Entonces tomando $N = \max\{N_1, N_2\}$ obtenemos

$$\left| \int_{\mathbb{R}} g(x)\, dF_n(x) - \int_{\mathbb{R}} g(x)\, dF(x) \right| < \varepsilon \left(M_n + M + 1 + 2K \right) \quad \forall\, n \geq N$$

Finalmente como la sucesión (M_n) es convergente, está acotada digamos por la constante C_0 .

Obtenemos entonces finalmente:

$$\left| \int_{\mathbb{R}} g(x)\, dF_n(x) - \int_{\mathbb{R}} g(x)\, dF(x) \right| < \varepsilon\, C \qquad \forall \ n \geq N$$

donde $C = M + C_0 + 2K + 1 = \text{Constante} < \infty$

Y como lo anterior se hizo para toda función continua g y todo $\varepsilon > 0$, se sigue (b).

Para demostrar la implicación $(b) \Rightarrow (a)$ demostremos primero el siguiente lema.

Lema 3. Usando las mismas notaciones y definiciones dadas en el Teorema, supongamos que (b) se cumple para toda función continua g con la propiedad:

$$\lim_{x \to \infty} g(x) = 0 \qquad \lim_{x \to -\infty} g(x) = 0$$

entonces $F_n \xrightarrow{d} F$

Demostración. Sea I un intervalo de continuidad de F, abierto y de longitud L. Dado $\varepsilon > 0$ sea I_δ un intervalo centrado en I y de longitud $L+\delta$ en donde $\delta > 0$ se escoge de tal manera que se tenga:

$$F\{I_\delta\} < F\{I\} + \varepsilon$$

Esto es posible ya que los extremos de I son puntos de continuidad de F.

Ahora tomemos una función continua $\mu(x)$ que dentro de I tome el valor constante 1, que se anule fuera de I_δ y tal que $0 \le \mu(x) \le 1$ en todas partes. Ahora bien observamos lo siguiente:

$$E_n(\mu) = \int_{-\infty}^{\infty} \mu(x)\, dF_n(x) = \int_{I_\delta} \mu(x)\, dF_n(x) + \int_{I_\delta^c} \mu(x)\, dF_n(x)$$

$$= \int_{I_\delta} \mu(x)\, dF_n(x)$$

$$= \int_{I} \mu(x)\, dF_n(x) + \int_{I_\delta - I} \mu(x)\, dF_n(x) = F_n\{I\} + \int_{I_\delta - I} \mu(x)\, dF_n(x)$$

por tanto $E_n(\mu) \geq F_n\{I\}$.

De igual forma observamos que:

$$E(\mu) = \int_{-\infty}^{\infty} \mu(x)\, dF(x) = \int_{I_\delta} \mu(x)\, dF(x) + \int_{I_\delta^c} \mu(x)\, dF(x)$$

$$= \int_{I_\delta} \mu(x)\, dF(x) \leq F\{I_\delta\} .$$

$$\therefore\ E(\mu) \leq F\{I_\delta\} < F\{I\} + \varepsilon$$

ahora para n *lo suficientemente grande, se tiene que:*

$E_n(\mu) < E(\mu) + \varepsilon$ *, por lo tanto tenemos:*

$$(a)\quad F_n\{I\} \leq E_n(\mu) < E(\mu) + \varepsilon \leq F\{I\} + 2\varepsilon$$

Ahora tomemos un intervalo I_δ *centrado en* I *de longitud* $L-\delta$ *donde se escoge a* $\delta > 0$ *de tal forma que:*

$$F\{I_\delta\} > F\{I\} - \varepsilon .$$

ahora sea $\mu(x)$ *una función continua que dentro de* I_δ *tome el valor constante* 1 *, y que se anule fuera de* I *y tal que* $0 \leq \mu(x) \leq 1$ *para todos los valores posibles.*

A continuación observemos las desigualdades:

$$E_n(\mu) = \int_{-\infty}^{\infty} \mu(x)\, dF_n(x) = \int_{I} \mu(x)\, dF_n(x) + \int_{I^c} \mu(x)\, dF_n(x)$$

$$= \int_{I} \mu(x)\, dF_n(x) \leq F_n\{I\} .$$

es decir $E_n(\mu) \leq F_n\{I\}$

Igualmente vemos que:

$$E(\mu) = \int_{-\infty}^{\infty} \mu(x)\, dF(x) = \int_{I} \mu(x)\, dF(x) + \int_{I^c} \mu(x)\, dF(x)$$

$$= \int_{I} \mu(x)\, dF(x)$$

$$= \int_{I_\delta} \mu(x)\, dF(x) + \int_{I-I_\delta} \mu(x)\, dF(x)$$

$$= F\{I_\delta\} + \int_{I-I_\delta} \mu(x)\, dF(x)$$

Por tanto $E(\mu) \geq F\{I_\delta\}$.

Ahora para n *suficientemente grande, se tiene que:*

$E(\mu) - \varepsilon < E_n(\mu)$ *por lo que se tiene:*

$$\text{(b)}\ F_n\{I\} \geq E_n(\mu) > E(\mu) - \varepsilon \geq F\{I_\delta\} - \varepsilon > F\{I\} - 2\varepsilon$$

Combinando (a) *y* (b) *obtenemos* $F_n \xrightarrow{d} F$, *ya que* I *fue cualquier intervalo abierto de continuidad de* F.

Supongamos ahora que (b) del teorema es válido, entonces como toda función continua g *con la propiedad del lema es acotada, se cumple la condición del lema y por lo tanto* $F_n \xrightarrow{d} F$.
Finalmente para demostrar que la convergencia es propia, apliquemos (b) a la función identicamente 1 *, con lo cual obtenemos* $F_n\{\mathbb{R}\} \longrightarrow F\{\mathbb{R}\}$.
En adelante diremos que una función continua $g: \mathbb{R} \to \mathbb{R}$ *se anula en el infinito si* $\lim_{x\to\infty} g(x) = 0$ *y* $\lim_{x\to-\infty} g(x) = 0$
Observemos ahora que el inverso del lema no es válido como se muestra con el siguiente ejemplo.

Ejemplo.

Definamos:

$$F_{2n}(x) = \begin{cases} 0 & \text{si } x < -2n \\ n & \text{si } x \geq -2n \end{cases}$$

$$F_{2n-1}(x) = \begin{cases} 0 & \text{si } x < 2n-1 \\ 2n-1 & \text{si } x \geq 2n-1 \end{cases}$$

Entonces, cualquiera que sea el intervalo (a,b) *, se tiene:*

$$F_n\{(a,b)\} \longrightarrow 0$$

Es decir $F_n \xrightarrow{d} F$ *donde* F *es la función de distribución identicamente cero.*
Por otra parte, definamos la siguiente función $g: \mathbb{R} \to \mathbb{R}$

$$g(x) = \begin{cases} 1 & \text{si } -1 \leq x \leq 1 \\ \left|\frac{1}{x}\right| & \text{en otro caso.} \end{cases}$$

g es continua y se anula en el infinito, sin embargo:

$$\int_{\mathbb{R}} g(x)\, dF_{2n}(x) = \frac{1}{2} \quad \forall\, n$$

$$\int_{\mathbb{R}} g(x)\, dF_{2n-1}(x) = 1 \quad . \quad \forall\, n$$

$$\therefore \int_{\mathbb{R}} g(x)\, dF_n(x) \quad \text{no converge.}$$

Se requiere entonces una condición más para asegurar que la convergencia débil implica la validez de la propiedad (b) del teorema para todas las funciones continuas g que se anulen en el infinito. El siguiente teorema establece la equivalencia bajo una condición adicional.

<u>Teorema 2.</u> Sean F_n $(n=1,2,\ldots)$ y F funciones de distribución y sea $M_n = F_n\{\mathbb{R}\}$. Supongamos que la sucesión (M_n) está acotada, entonces las siguientes condiciones son equivalentes:

a) $F_n \xrightarrow{d} F$

b) $\int_{\mathbb{R}} g(x)\, dF_n(x) \longrightarrow \int_{\mathbb{R}} g(x)\, dF(x)$ $\forall$ función continua g que se anule en el infinito.

<u>Demostración.</u>

a) $\Rightarrow$ b) Sea $M = F\{\mathbb{R}\}$, K una cota superior de (M_n) mayor que M y sea g una función continua que se anule en el infinito.

Dado $\varepsilon > 0$ tomamos $a > 0$ tal que:

$$|g(x)| < \varepsilon \quad \text{si} \quad x \in [-a, a]^c$$

$$F\{[-a, a]\} > M - \varepsilon$$

Se tiene entonces:

$$\left| \int_{[-a,a]^c} g(x)\, dF_n(x) \right| \leq \varepsilon F_n\left\{[-a,a]^c\right\} \leq \varepsilon F_n\left\{\mathbb{R}\right\} \leq \varepsilon K$$

$$\left| \int_{[-a,a]^c} g(x)\, dF(x) \right| \leq \varepsilon F\left\{[-a,a]^c\right\} \leq \varepsilon F\left\{\mathbb{R}\right\} \leq \varepsilon K$$

Procedamos entonces exactamente como en la primera parte de la demostración del Teorema I para obtener N *tal que:*

$$\left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} g(x)\, dF(x) \right| < \varepsilon M_n + \varepsilon M + \varepsilon \qquad \forall\, n \geq N$$

Entonces:

$$\left| \int_{\mathbb{R}} g(x)\, dF_n(x) - \int_{\mathbb{R}} g(x)\, dF(x) \right| \leq \left| \int_{[-a,a]} g(x)\, dF_n(x) - \int_{[-a,a]} g(x)\, dF(x) \right|$$

$$+ \left| \int_{[-a,a]^c} g(x)\, dF_n(x) \right| + \left| \int_{[-a,a]^c} g(x)\, dF(x) \right|$$

$$< \varepsilon \left(M_n + M + 1 + 2K \right)$$

$$\leq \varepsilon \left(M + 1 + 3K \right) \qquad \forall\, n \geq N$$

Y como lo anterior se hizo para toda función continua g *y todo* $\varepsilon>0$ *, se concluye la validez de (b).*

Finalmente, la implicación $b) \Rightarrow a)$ *se sigue inmediatamente del lema3.*

Como caso particular sean X_n $(n=1,2,\dots)$ *y* X *variables aleatorias con funciones de distribución* F_n $(n=1,2,\dots)$ *y* F *respectivamente, sean además* $M_n=F_n(\mathbb{R})$ *,* $M=F(\mathbb{R})$ *; entonces en este caso* (M_n) *es una sucesión constante e igual a* M *, así que podemos aplicar el corolario 2 de la proposición 6, el teorema I y el teorema 2 para obtener el siguiente corolario.*

__Corolario 3.__ *Sean* X_n $(n=1,2,\dots)$ *y* X *variables aleatorias con funciones de distribución* F_n $(n=1,2,\dots)$ *y* F *respectivamente, entonces las siguientes condiciones son equivalentes:*

a) $F_n \xrightarrow{d} F$

b) $F_n(x) \longrightarrow F(x)$ $\forall\, x$ *punto de continuidad de* F .

c) $\int_{\mathbb{R}} g(x)\, dF_n(x) \longrightarrow \int_{\mathbb{R}} g(x)\, dF(x)$ $\forall$ *función continua y acotada* g .

d) $\int_{\mathbb{R}} g(x)\, dF_n(x) \longrightarrow \int_{\mathbb{R}} g(x)\, dF(x)$ $\forall$ *función continua* g *que se anule en el infinito.*

Los siguientes resultados son claves para el estudio de la convergencia de funciones de distribución pues dan un criterio muy útil para probar la convergencia débil de una sucesión de funciones de distribución.

Teorema 3. (Teorema de Helley)

Para toda sucesión $\{F_n\}$ *de funciones de distribución tales que la sucesión* $(F_n(\mathbb{R}))$ *está acotada, existe una subsucesión* $\{F_{n_K}\}$ *y una función de distribución* F *tal que* $F_{n_K} \xrightarrow{d} F$.

Demostración. Sean $r_1, r_2, r_3, \ldots$ los racionales en $\mathbb{R}$

Para cada J , la sucesión $(F_n(r_J))$ está acotada, por lo tanto aplicando el método diagonal (teorema 6 del apéndice) existe una sucesión creciente (n_K) de números naturales tal que

$\lim_K F_{n_K}(r_J)$ existe $\forall J$.

Sea $G_1(r_J) = \lim_K F_{n_K}(r_J)$

Evidentemente G_1 es una función monótona no decreciente definida sobre $\mathbb{Q}$.

Definamos ahora $G_2 : \mathbb{R} \longrightarrow [0,1]$ de la siguiente manera:

$$G_2(x) = \inf \{ G_1(r_J) : x < r_J \}$$

Sigue siendo evidente que G_2 es una función monótona no decreciente.

Demostremos que G_2 es continua por la derecha:

Por definición, dado $x \in \mathbb{R}$ y $\varepsilon > 0$ existe $r \in \mathbb{Q}$ tal que $x < r$ y tal que: $G_1(r) < G_2(x) + \varepsilon$

Entonces si $x \leq y < r$ obtenemos:

$$G_2(y) \leq G_1(r) < G_2(x) + \varepsilon$$

Es decir:

$$G_2(y) - G_2(x) < \varepsilon$$

Lo cual prueba que G_2 es continua por la derecha en el punto x.

Sea ahora x un punto de continuidad de G_2

Dado $\varepsilon>0$ tomemos $y<x$ tal que

$$G_2(x)-\varepsilon < G_2(y)$$

Y tomemos $r, s \in \mathbb{Q}$ de modo que $y<r<x<s$ y tal que:

$$G_1(s) < G_2(x)+\varepsilon$$

Obtenemos entonces:

$$G_2(x)-\varepsilon < G_2(y) \leq G_1(r) \leq G_1(s) < G_2(x)+\varepsilon$$

Por otra parte, sabemos que:

$$G_1(r) = \lim_k F_{n_k}(r)$$

$$G_1(s) = \lim_k F_{n_k}(s)$$

Por lo tanto, existe N tal que:

$$F_{n_k}(r) > G_2(x)-\varepsilon \qquad \forall\ k \geq N$$

$$F_{n_k}(s) < G_2(x)+\varepsilon \qquad \forall\ k \geq N$$

Pero:

$$F_{n_k}(r) \leq F_{n_k}(x) \leq F_{n_k}(s) \qquad \forall\ k$$

Por lo tanto

$$G_2(x)-\varepsilon < F_{n_k}(x) < G_2(x)+\varepsilon \qquad \forall\ k \geq N$$

Y entonces:

$$\lim_k F_{n_k}(x) = G_2(x)$$

Definamos ahora:

$$F(x) = G_2(x) - \lim_{y \to -\infty} G_2(y)$$

F es entonces una función de distribución.

Por el lema2. $F_{n_k}\{\{x\}\} \longrightarrow 0$ $\forall x$ punto de continuidad de F

Y entonces si I es un intervalo acotado de continuidad de F con extremos a y b se tiene:

$$F_{n_k}\{I\} \longrightarrow F(b) - F(a) = F\{I\}$$

Es decir $F_{n_k} \xrightarrow{d} F$

Finalizamos este capítulo con un criterio para la convergencia débil de una sucesión de funciones de distribución de probabilidad.

<u>Teorema 4.</u> Sea $\{F_n\}$ una sucesión de funciones de distribución de probabilidad y supongamos que toda subsucesión de $\{F_n\}$ débilmente convergente converge a la misma distribución F, entonces $F_n \xrightarrow{d} F$.

<u>Demostración.</u>

Supongamos que $\{F_n\}$ no converge a F, usando entonces el teorema 2 sabemos que debe existir una función g que se anule en el infinito y tal que:

$$\int_{\mathbb{R}} g(x)\, dF_n(x) \quad \text{no converge a} \quad \int_{\mathbb{R}} g(x)\, dF(x)$$

Esto implica la existencia de $\varepsilon > 0$ y una subsucesión $\{F_{n_k}\}$ de $\{F_n\}$ tal que:

$$\left| \int_{\mathbb{R}} g(x)\, dF_{n_k}(x) - \int_{\mathbb{R}} g(x)\, dF(x) \right| \geq \varepsilon \qquad \forall \quad K$$

Pero entonces $\{F_{n_k}\}$ no puede tener ninguna subsucesión débilmente convergente pues por hipótesis cualquiera de ellas debe converger a F.

Esto contradice el teorema de Helley y por lo tanto nuestra hipótesis es falsa, lo cual prueba el teorema.

CAPITULO IV

LA TRANSFORMADA DE LAPLACE

Anteriormente hemos definido la Función generatriz de momentos de una variable aleatoria X con función de distribución F_X como:

$$M(s) = E\left(e^{sX}\right) = \int_{-\infty}^{\infty} e^{sx} d F_X(x)$$

para toda s en donde la integral anterior sea finita.

Cuando F_X está concentrada en $[0, \infty)$ se ve inmediatamente que $M(s)$ es finita para toda $s \leq 0$.

En general la función generatriz de una variable aleatoria es útil cuando se encuentra definida en una vecindad del cero. Sin embargo, veremos en este capítulo que cuando la función de distribución de la variable aleatoria está concentrada en $[0, \infty)$ los valores de la función generatriz en $(-\infty, 0]$ son igualmente útiles, aún cuando la función generatriz puede no estar definida en una vecindad del 0.

Para esto, damos la siguiente definición:

Definición I. Sea F una función de distribución concentrada en $[0, \infty)$. La transformada de Laplace de F es la función $\varphi : \mathbb{R}^+ \longrightarrow \mathbb{R}$ definida por:

$$\varphi(s) = \int_{0}^{\infty} e^{-sx} dF(x)$$

Los siguientes resultados servirán para demostrar más adelante que la transformada de Laplace determina a la distribución.

Definición 2. Una sucesión de números reales $\{C_k\}$ es llamada completamente monótona si: $(-1)^r \Delta^r C_k \geq 0$ para todo $r=0,1,2,\ldots$ donde ya se ha explicado el significado del factor diferencias para una sucesión $\{C_k\}$ (ver la definición 2 del capítulo II).

Teorema I. Una sucesión $\{C_k\}$ define los momentos de una función de distribución concentrada en $[0,1]$ si, y solo si la sucesión es completamente monótona.
En este caso la función de distribución F queda determinada por sus momentos.

Demostración. $\Rightarrow]$ Tomemos una distribución F concentrada en $[0,1]$, el k-ésimo momento de F está dado por:

$$C_k = E(X^k) = \int_0^1 X^k \, dF(X) \quad \text{, ahora bien}$$

$$\begin{aligned} (-1)^0 \Delta^0 C_k &= C_k \\ &= E(X^k) \\ &= E\left(X^k (1-X)^0\right) \end{aligned}$$

observemos también que:

$$\begin{aligned} (-1)\Delta C_k &= (-1)\left(C_{k+1} - C_k\right) = C_k - C_{k+1} \\ &= E\left(X^k - X^{k+1}\right) = E\left(X^k(1-X)\right) \end{aligned}$$

Igualmente observamos que:

$$\begin{aligned} (-1)^2 \Delta^2 C_k &= \Delta^2 C_k = C_{k+2} - 2C_{k+1} + C_k \\ &= E\left(X^{k+2} - 2X^{k+1} + X^k\right) = E\left(X^k(1-X)^2\right) \end{aligned}$$

Así este proceso inductivo nos lleva a:

$$\text{(A)} \quad (-1)^r \Delta^r C_k = E\left[X^k (1-X)^r\right]$$

y claramente

$$E\left[X^k(1-X)^r\right] = \int_0^1 X^k(1-X)^r dF(X) \geq 0$$

puesto que $F(X)$ *está concentrada en* $[0,1]$.

Por lo tanto la sucesión $\{C_K\}$ *es una sucesión completamente monótona.*

$\Leftarrow$] *Antes de probar el inverso.*

Primero veamos lo siguiente:

Tomemos $F(X)$ *una distribución concentrada en* $[0,1]$.

Sea $\{C_K\}$ *su correspondiente sucesión de momentos y sea* $\mu(X)$ *una función continua arbitraria.*

Definamos, como antes:

$$E(\mu) = \int_{-\infty}^{\infty} \mu(X)\, dF(X)$$

Recordemos que si el polinomio de Bernstein de grado n *correspondiente a* $\mu(X)$ *(definición I del Cap. II) está dado por:*

$$B_{n,\mu}(\theta) = \sum_{J=0}^{n} \mu(Jh) \binom{n}{J} \theta^J (1-\theta)^{n-J} \qquad \text{donde se toma } h = 1/n$$

Ahora integrando $B_{n,\mu}$ *con respecto a* F, *obtenemos lo siguiente:*

$$E(B_{n,\mu}) = \int_0^1 \sum_{J=0}^{n} \mu(Jh)\binom{n}{J} \theta^J(1-\theta)^{n-J}\, dF(\theta)$$

$$= \sum_{J=0}^{n} \mu(Jh)\binom{n}{J} \int_0^1 \theta^J(1-\theta)^{n-J}\, dF(\theta)$$

y aplicando la igualdad (A), obtenemos:

$$E(B_{n,\mu}) = \sum_{J=0}^{n} \mu(Jh)\binom{n}{J}(-1)^{n-J}\Delta^{n-J}C_J$$

$$(B) \qquad = \sum_{J=0}^{n} \mu(Jh)\, P_J^{(n)}$$

donde por abreviar, hemos hecho:

$$(C) \quad \rho_J^{(n)} = \binom{n}{J}(-1)^{n-J}\Delta^{n-J} C_J \qquad J = 0, 1, 2, \ldots, n$$

Como la sucesión $\{C_k\}$ *es completamente monótona, se tiene*

$$\rho_J^n \geq 0 \quad \forall \; n, J$$

Sea entonces μ_n *una medida sobre* $\mathbb{R}$ *concentrada en el conjunto* $\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\right\}$ *y tal que:*

$$\mu_n\left(\left\{\frac{J}{n}\right\}\right) = \rho_J^{(n)}$$

Sea F_n *la función de distribución asociada a* μ_n *y como antes denotemos por* E_n *a las esperanzas con respecto a* F_n

La relación (B) *puede entonces escribirse de la siguiente manera:* $E(B_{n,\mu}) = E_n(\mu)$

Pero $B_{n,\mu} \to \mu$ *uniformemente, cuando* $n \to \infty$, *por lo tanto:* $\lim\limits_{n\to\infty} E(B_{n,\mu}) = E(\mu)$.

Y entonces también:

$$\lim_{n\to\infty} E_n(\mu) = E(\mu)$$

Y como μ *fué una función continua arbitraria, concluimos entonces (usando el teoremaI del capítulo III):*

$$F_n \xrightarrow{d} F \qquad \textit{propiamente.}$$

Lo anterior nos da un método para probar lo que deseamos pues si ahora $\{C_k\}$ *es una sucesión completamente monótona arbitraria y deseamos probar que* $\{C_k\}$ *define los momentos de una función de distribución* F, *entonces podemos encontrar* F *definiendo* $\rho_J^{(n)}$ *como en (C) y luego las correspondientes* F_n *y entonces* F *deberá ser el límite débil de la sucesión* (F_n). *Procederemos de esta manera a continuación.*

Sea $\{C_K\}$ *una sucesión completamente monótona arbitraria.*

Definamos entonces:

$$P_J^{(n)} = \binom{n}{J}(-1)^{n-J}\Delta^{n-J}C_J \qquad J = 0, 1, 2, \ldots, n$$
$$n = 1, 2, \ldots$$

Como $\{C_K\}$ *es completamente monótona, se tiene* $P_J^{(n)} \geq 0 \quad \forall J, n$

Por otra parte, sea μ *una función arbitraria y definamos:*

$$a_i = \mu\left(\frac{i}{n}\right) \qquad i = 0, 1, 2, \ldots$$

Por la proposición 3 del capítulo II, tenemos:

$$\sum_{J=0}^{n} \mu\left(\frac{J}{n}\right)\binom{n}{J}(-1)^{n-J}\Delta^{n-J}C_J = \sum_{r=0}^{n} C_r \binom{n}{r}\Delta^r a_0$$

Es decir:

$$\sum_{J=0}^{n} \mu\left(\frac{J}{n}\right) P_J^{(n)} = \sum_{r=0}^{n} C_r \binom{n}{r}\Delta^r a_0$$

Tomemos ahora μ *identicamente* 1 *, entonces:*

$$\Delta^0 a_0 = a_0 = 1 \qquad \Delta^r a_0 = 0 \qquad r = 1, 2, \ldots$$

Por lo tanto, obtenemos:

$$\sum_{J=0}^{n} P_J^{(n)} = C_0 \qquad \forall n$$

Definamos entonces las medidas μ_n *concentradas en el conjunto* $\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\right\}$ *como antes y denotemos por* F_n *a su correspondiente función de distribución.*

Se tiene $F_n(\mathbb{R}) = C_0 \quad \forall n$ *así que por el teorema de Helley existe una subsucesión* (F_{n_K}) *de* (F_n) *que converge débilmente.*

Sea F *el límite débil de esta subsucesión, es decir:*

$$F_{n_K} \xrightarrow{d} F$$

Evidentemente, F *está concentrada en* $[0,1]$ *y la convergencia es propia*

Ahora bien, por el teorema I del capítulo III, si μ *es cualquier función continua definida en el intervalo* $[0,1]$ *, tenemos:*

$$\lim_{K\to\infty} \int_{-1}^{1} \mu(x)\, dF_{n_K}(x) = \int_{-1}^{1} \mu(x)\, dF(x)$$

En particular, tomando $\mu(x) = x^K$ *, obtenemos que el k-ésimo momento de* F_{n_K} *converge al k-ésimo momento de* F *.*
Para probar entonces que C_K *es el k-ésimo momento de* F *bastará con demostrar entonces que el k-ésimo momento de* F_n *converge a* C_K *, lo cual haremos a continuación.*

Sea μ *un polinomio definido en todo* $\mathbb{R}$ *, tomemos* $h = 1/n$ *y consideremos las razones de diferencias de* μ *dadas por la definición 3 del capítulo III.*

$$V_1(x) = \underset{h}{\Delta} \mu(x) = \frac{\mu(x+h) - \mu(x)}{h}$$

$$V_2(x) = \underset{h}{\Delta}^2 \mu(x) = \underset{h}{\Delta} V_1(x) = \frac{V_1(x+h) - V_1(x)}{h}$$

$$\vdots$$

$$V_J(x) = \underset{h}{\Delta}^J \mu(x) = \underset{h}{\Delta} V_{J-1}(x) = \frac{V_{J-1}(x+h) - V_{J-1}(x)}{h}$$

Por el teorema del valor medio, podemos escribir:

$$V_J^{(m)}(x) = \frac{V_{J-1}^{(m)}(x+h) - V_{J-1}^{(m)}(x)}{h}$$

$$= V_{J-1}^{(m+1)}(x + \theta h) \qquad \text{con} \quad 0 < \theta < 1$$

Fijemos entonces $r \in \mathbb{N}$ *y tomemos* $x = 0$ *, entonces:*

$$V_r(0) = \frac{V_{r-1}(h) - V_{r-1}(0)}{h} = V'_{r-1}(\theta_1 h)$$

$$= V_{r-2}^{(2)}(\theta_1 h + \theta_2 h) = \cdots = V_1^{(r-1)}(\theta_1 h + \theta_2 h + \cdots + \theta_{r-1} h)$$

$$= \mu^{(r)}(\theta_1 h + \theta_2 h + \cdots + \theta_r h)$$

donde $0 < \theta_i < 1 \qquad \forall \ i = 1, 2, \ldots, r$

Sea $\theta = \theta_1 + \theta_2 + \cdots + \theta_r$

Entonces $0 < \theta < r$ *y además:*

$$\underset{h}{\Delta}^r \mu(0) = V_r(0) = \mu^{(r)}(\theta h)$$

$$\therefore \lim_{n\to\infty} \underset{h}{\Delta}^{r} \mu(0) = \mu^{(r)}(0)$$

Definamos ahora:

$$a_i = \mu(ih)$$

Ya hemos visto que se tiene:

$$\underset{h}{\Delta}^{n}\mu(0) = \frac{\Delta^{n}(a_0)}{h^{n}}$$

Aplicando la fórmula de reciprocidad obtenida en la proposición3 del capítulo II a las sucesiones $\{a_i\}$ y $\{C_k\}$ *obtenemos entonces:*

$$\sum_{r=0}^{m} C_r \binom{m}{r} h^{r} \underset{h}{\Delta}^{r}\mu(0) = \sum_{J=0}^{m} \mu(Jh) \binom{m}{J} (-1)^{m-J} \Delta^{m-J} C_J \qquad \forall\, m \in \mathbb{Z}.$$

Por lo tanto, podemos escribir:

$$E_n(\mu) = \sum_{J=0}^{n} \mu(Jh)\, \rho_J^{(n)} = \sum_{J=0}^{n} \mu(Jh) \binom{n}{J} (-1)^{n-J} \Delta^{n-J} C_J$$

$$= \sum_{r=0}^{n} C_r \binom{n}{r} h^{r} \underset{h}{\Delta}^{r}\mu(0)$$

Tomemos en particular $\mu(x) = x^{k}$

Entonces, como se tiene:

$$\underset{h}{\Delta}^{r}\mu(0) = \mu^{(r)}(\theta h) \qquad \text{con} \quad 0<\theta<r \qquad \forall\, r \in \mathbb{N}$$

$$\underset{h}{\Delta}^{r}\mu(0) = 0 \qquad \forall\, r > k$$

podemos escribir (Tomando $n > k$)

$$E_n(\mu) = \sum_{r=0}^{k} C_r \binom{n}{r} h^{r} \underset{h}{\Delta}^{r}\mu(0)$$

Tomando ahora límites cuando $n \to \infty$, *obtenemos:*

$$\lim_{n\to\infty} E_n(\mu) = \sum_{r=0}^{K} C_r \lim_{n\to\infty}\left[\binom{n}{r} h^r\right] \lim_{n\to\infty} \Delta_h^r \mu(0)$$

Pero:

$$\binom{n}{r} h^r = \frac{1}{r!} \; \frac{n(n-1)\cdots(n-r+1)}{n^r}$$

$$= \frac{1}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{r-1}{n}\right)$$

$$\therefore \lim_{n\to\infty}\binom{n}{r} h^r = \frac{1}{r!}$$

Por otra parte:

$$\lim_{n\to\infty} \Delta_h^r \mu(0) = \mu^{(r)}(0) = \begin{cases} 0 & \text{si } r \neq K \\ r! & \text{si } r = K \end{cases}$$

Por lo tanto, concluimos:

$$\lim_{n\to\infty} E_n(\mu) = C_K$$

Para concluir la demostración del teorema, sea F *una función de distribución concentrada en* $[0,1]$.
Demostraremos que F *queda determinada por sus momentos.*

Sea $\{C_k\}$ *los momentos de* F *y definamos como antes para cada n una medida* μ_n *concentrada en el conjunto* $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\}$ *y tal que:*

$$\mu_n\left(\left\{\frac{J}{n}\right\}\right) = P_J^n = \binom{n}{J}(-1)^{n-J}\Delta^{n-J}C_J \qquad J = 0, 1, 2, \ldots, n$$

Si F_n *es la función de distribución asociada a* μ_n *, ya hemos visto que* $F_n \xrightarrow{d} F$ *propiamente.*
Por lo tanto, usando la proposición C del capítulo III, resulta que $F_n(x) \longrightarrow F(x) \quad \forall x$ *punto de continuidad de* F.
Pero:

$$F_n(x) = \sum_{\frac{J}{n} \leq x} \binom{n}{J}(-1)^{n-J}\Delta^{n-J}C_J$$

De manera que los momentos de F *determinan sus valores en sus puntos de continuidad, los cuales determinan a su vez todos los valores de* F.

Teorema 2. Si $F(x)$ y $H(y)$ son distribuciones concentradas en $[0,\infty)$ y además $\varphi_1(s)=\int_0^\infty e^{-sx}dF(x)$ y $\varphi_2(s)=\int_0^\infty e^{-sy}dH(y)$ son tales que $\varphi_1(s)=\varphi_2(s)$ para todo s, entonces $F=H$.

Demostración. Hagamos $y=e^{-x}$, entonces observamos que conforme x toma valores desde 0 a ∞, entonces la variable y toma valores de 1 a 0.

Ahora bien definimos a la distribución $G(y)$ de tal forma que este concentrada en $[0,1]$ y $G(y)=1-F(x)$ para todo punto de continuidad.

Entonces vemos que:

$$\varphi(s)=\int_0^\infty e^{-sx}dF(x)=\int_0^1 y^s\, dG(y)$$

Ahora bien observamos que la siguiente expresión $\varphi(s)=\int_0^1 y^s dG(y)$ calculada en los valores $s=0,1,2,3,\ldots$ nos da los momentos de la distribución $G(y)$, es decir $\varphi(k)=\alpha_k$, y por lo tanto la sucesión de momentos $\alpha_0,\alpha_1,\alpha_2,\ldots$ es completamente monótona, entonces en virtud del Teorema 1 anteriormente demostrado, G está determinada de forma única por sus momentos, y por lo tanto el conocimiento de $\varphi(0),\varphi(1),\ldots$ determina a $G(y)$ y por tanto a $F(x)$ también.

A continuación veremos otra demostración de este Teorema, donde se tiene una fórmula explicita de inversión que permite calcular $F(x)$ cuando se conoce su Transformada $\varphi(s)$.

Antes veamos los siguientes 2 resultados:

Lema I. Para $\lambda > 0$ definimos $G_\lambda : \mathbb{R}^+ \to \mathbb{R}$ de la siguiente manera:

$$G_\lambda(\tau) = \sum_{k=0}^{[\lambda\tau]} \frac{e^{-\lambda}\lambda^k}{k!}$$

donde $[\lambda\tau]$ denota la parte entera de $\lambda\tau$.

entonces:
$$\lim_{\lambda\to\infty} G_\lambda(\tau) = \begin{cases} 1 & \text{si } \tau > 1 \\ 0 & \text{si } \tau < 1 \end{cases}$$

Demostración.

Sea Y una variable aleatoria tipo Poisson de parámetro λ

Entonces: $G_\lambda(\tau) = P[Y \leq \lambda\tau]$

Si $\tau > 1$, sea $\tau_0 = \tau - 1$, entonces $\tau_0 > 0$ y usando la desigualdad de Chebichev, tenemos:

$$P[Y > \lambda\tau] = P[Y > \lambda\tau_0 + \lambda] = P[Y - \lambda > \lambda\tau_0]$$
$$= P[Y - E(Y) > \lambda\tau_0]$$
$$\leq P[|Y - E(Y)| > \lambda\tau_0] \leq \frac{Var(Y)}{\lambda^2\tau_0^2} = \frac{1}{\lambda\tau_0^2}$$

$$\therefore \lim_{\lambda\to\infty} P[Y > \lambda\tau] \leq \lim_{\lambda\to\infty} \frac{1}{\lambda\tau_0^2} = 0$$

Así que
$$\lim_{\lambda\to\infty} G_\lambda(\tau) = \lim_{\lambda\to\infty} P[Y \leq \lambda\tau]$$
$$= 1 - \lim_{\lambda\to\infty} P[Y > \lambda\tau] = \underline{1}$$

Ahora, si $\tau < 1$, sea $\tau_0 = 1 - \tau$, entonces $\tau_0 > 0$ y además:

$$P[Y \leq \lambda\tau] = P[Y \leq \lambda - \lambda\tau_0] = P[Y-\lambda \leq -\lambda\tau_0]$$

Pero, usando nuevamente la desigualdad de Chebichev, para λ *suficientemente grande de manera que* $\lambda\tau_0 > 1$ *, tenemos:*

$$\begin{aligned} P[Y-\lambda \leq -\lambda\tau_0] &\leq P[Y-\lambda < -\lambda\tau_0 + 1] \\ &\leq P[|Y-E(Y)| > \lambda\tau_0 - 1] \\ &\leq \frac{Var(Y)}{(\lambda\tau_0-1)^2} = \frac{\lambda}{(\lambda\tau_0-1)^2} \end{aligned}$$

$$\therefore \lim_{\lambda\to\infty} P[Y-\lambda \leq -\lambda\tau_0] \leq \lim_{\lambda\to\infty} \frac{\lambda}{(\lambda\tau_0-1)^2} = 0$$

Es decir,

$$\lim_{\lambda\to\infty} G_\lambda(\tau) = 0 \quad .$$

Lema 2. Sea F una función de distribución concentrada en $[0,\infty)$ y definamos:

$$\varphi(s) = \int_0^\infty e^{-sy}\, dF(y) \qquad \text{para } s>0$$

entonces la función φ es infinitamente diferenciable en $(0,\infty)$ y además:

$$\varphi^{(K)}(s) = (-1)^K \int_0^\infty y^K e^{-ys}\, dF(y) \qquad \forall\ s>0 .$$

Demostración.

Primero observamos que:

$$\lim_{y\to\infty} \frac{y^K}{e^{sy}} = 0 \qquad \forall\ s>0$$

Así que $\forall\ s>0$ la función $y \longrightarrow y^K e^{-sy}$ definida en $[0,\infty)$ es acotada y por lo tanto integrable con respecto a F.

Pero como F está concentrada en $[0,\infty)$, la función $y \longrightarrow y^K e^{-sy}$ definida en todo $\mathbb{R}$ también es integrable.

Sea ahora $I=(a,b)$ donde $b>a>0$ y definamos $f: \mathbb{R}\times I \to \mathbb{R}$ de la siguiente manera $f(y,\tau) = e^{-\tau y}$

Tomando como espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}), dF)$ tenemos:

a) La función $y \longrightarrow f(y,\tau)$ es medible e integrable $\forall\ \tau\in I$

b) La función $\tau \longrightarrow f(y,\tau)$ es derivable en I $\forall\ y\in\mathbb{R}$

c) $\dfrac{\partial}{\partial\tau} f(y,\tau) = -y\, e^{-\tau y}$ $\forall y\in\mathbb{R}$, así que:

$$\left| \frac{\partial}{\partial\tau} f(y,\tau) \right| = |y|\, e^{-\tau y} \le y e^{-ay} \qquad \forall\ \tau\in I,\ y\geq 0$$

Entonces tomando $A=(-\infty,0)$, la función f satisface las condiciones del Teorema 5 del apéndice y entonces aplicando éste obtenemos:

$$\varphi'(s) = \int_0^\infty -y e^{-sy} dF(y) \qquad \forall \ s \in I$$

Y aplicando el mismo argumento sucesivamente, obtenemos:

$$\varphi^{(k)}(s) = (-1)^k \int_0^\infty y^k e^{-sy} dF(y) \qquad \forall \ s \in I \ ; \quad k = 1, 2, \ldots$$

Finalmente, como el intervalo $I = (a, b)$ con $b > a > 0$ fué arbitrario, obtenemos:

$$\varphi^k(s) = (-1)^k \int_0^\infty y^k e^{-sy} dF(y) \qquad \forall \ s > 0, \quad k = 1, 2, \ldots$$

Teorema 3.
Sea $F(x)$ *la función de distribución de una variable aleatoria* X *concentrada en* $[0,\infty)$ *, entonces la Transformada de Laplace*

$$\varphi(s) = \int_0^\infty e^{-sx}\,dF(x) \qquad s \geq 0$$

determina a la Función de Distribución.

Demostración. *En el Lema anterior, hemos visto que para* s *positiva*

$$\varphi^{(k)}(s) = (-1)^k \int_0^\infty y^k e^{-sy}\,dF(y)$$

por lo tanto para x *y* s *positivos*

$$\sum_{k=0}^{[sx]} \frac{(-1)^k}{k!} s^k \varphi^{(k)}(s) = \sum_{k=0}^{[sx]} \frac{(-1)^k}{k!} s^k (-1)^k \int_0^\infty y^k e^{-sy}\,dF(y)$$

$$= \sum_{k=0}^{[sx]} \frac{s^k}{k!} \int_0^\infty y^k e^{-sy}\,dF(y)$$

$$= \int_0^\infty \sum_{k=0}^{[sx]} \frac{s^k y^k}{k!} e^{-sy}\,dF(y)$$

$$= \int_0^\infty \sum_{k=0}^{[sx]} \frac{(sy)^k}{k!} e^{-sy}\,dF(y) = \int_0^\infty \sum_{k=0}^{[sx]} \frac{e^{-sy}(sy)^k}{k!}\,dF(y)$$

$$= \int_0^\infty G_{sy}\left(\frac{x}{y}\right) dF(y)$$

ahora fijemos $x>0$ y observamos lo siguiente:

Si $0 \leq y < x$, entonces $G_{sy}\left(\frac{x}{y}\right) \longrightarrow 1$ cuando $s \to \infty$ y si $y > x$ el límite es 0, por el Lema 1 anterior.

Entonces si $y \neq x$

$$G_{sy}\left(\frac{x}{y}\right) \longrightarrow I_{[0,x]}(y) \quad \text{cuando} \quad s \longrightarrow \infty .$$

Entonces por el Teorema de la Convergencia Dominada (ver apéndice) si $P[X=x]=0$ tenemos:

$$(*) \lim_{s\to\infty} \sum_{k=0}^{[sx]} \frac{(-1)^k}{k!} s^k \varphi^{(k)}(s) = \int_0^\infty I_{[0,x]}(y)\, dF(y) = P[X \leq x] = F(x)$$

De esta forma $\varphi(s)$ determina el valor de F en x si $x>0$ y $P[X=x]=0$, pero $\{x>0 : P[X=x]=0\}$ es un conjunto denso en $[0,\infty)$ y como $F(x)$ es continua por la derecha, entonces está determinada por $\varphi(s)$ por medio de la ecuación $(*)$.

CAPITULO V

Teoremas de Continuidad.

En este capítulo veremos tres teoremas de Continuidad, uno para los Momentos, uno para la Función Generatriz y otro para la Transformada de Laplace.

El teorema de continuidad para las Transformadas de Laplace establece una correspondencia entre las Funciones De Distribución concentradas en $[0, \infty)$ y las Transformadas de Laplace, que no sólo es una correspondencia uno a uno sino también una correspondencia bicontinua.

Los Teoremas de Continuidad pueden ser usados para el estudio de Distribuciones Límite.

En el capítulo II ya hemos visto un Teorema de esta naturaleza (Teorema 3) para el caso discreto.

Definición I. Sean X_n $(n=1,2,\ldots)$ y X variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Sea F_n la función de distribución de X_n para cada n y sea F la función de distribución de X. Diremos que la sucesión (X_n) converge en distribución a X si $F_n \xrightarrow{d} F$. En este caso escribiremos $X_n \xrightarrow{D} X$.

El corolario al teorema 2 del capítulo III nos da inmediatamente el siguiente resultado:

Teorema I. Sean X_n $(n=1,2,\ldots)$ y X variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Las siguientes condiciones son equivalentes:

a) $X_n \xrightarrow{D} X$

b) $E\left[g(X_n)\right] \longrightarrow E\left[g(X)\right]$ función continua y acotada g.

c) $E\left[g(X_n)\right] \longrightarrow E\left[g(X)\right]$ función continua g que se anule en el infinito.

Proposición I. Sean X_n $(n=1,2,\ldots)$ y X variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, sea $h: \mathbb{R} \to \mathbb{R}$ una función continua y supongamos $X_n \xrightarrow{D} X$. Entonces $h(X_n) \xrightarrow{D} h(X)$

Demostración. Sea $g: \mathbb{R} \to \mathbb{R}$ una función continua y acotada entonces $g \circ h: \mathbb{R} \to \mathbb{R}$ es también una función continua y acotada, por lo tanto, aplicando el teorema I del capítulo III obtenemos:

$$E\left[(g\circ h)(X_n)\right] \longrightarrow E\left[(g\circ h)(X)\right]$$

Es decir:

$$E\left[g(h(X_n))\right] \longrightarrow E\left[g(h(X))\right]$$

Y como g *fué una función continua y acotada arbitraria, aplicando el mismo teorema I del capítulo III obtenemos el resultado.*

Teorema 2. (Skorohod). *Sean* $F_n \ (n=1,2,\dots)$ *y* F *distribuciones de probabilidad y supongamos* $F_n \xrightarrow{d} F$ *. Entonces existe un espacio de probabilidad* $(\Omega, \mathcal{F}, P)$ *y variables aleatorias* $Y_n \ (n=1,2,\dots)$, Y *sobre ese espacio tales que* Y_n *tiene a* F_n *como función de distribución para cada* n , Y *tiene a* F *como función de distribución y* $Y_n(\omega) \longrightarrow Y(\omega) \ \forall \omega \in \Omega$.

Demostración. *Tomamos* $\Omega = (0,1)$, $\mathcal{F}$ *los conjuntos Lebesgue-medibles en* $(0,1)$ *y* P *la medida de Lebesgue sobre* $(0,1)$ *.* *Definamos entonces* $Z_n : \Omega \rightarrow \mathbb{R}$, $Y : \Omega \rightarrow \mathbb{R}$ *de la siguiente manera:*

$$Y(\omega) = \inf \{ x \in \mathbb{R} : \omega \leq F(x) \}$$

$$Z_n(\omega) = \inf \{ x \in \mathbb{R} : \omega \leq F_n(x) \}$$

tenemos entonces:

(I) $\{ \omega \in (0,1) : Y(\omega) \leq x \} = \{ \omega \in (0,1) : \omega \leq F(x) \} = (0, F(x)] \cap (0,1)$

En efecto:

Si $Y(\omega) = x$ *entonces existe una sucesión* (x_n) *tal que* $x_n \downarrow x$ *y satisfaciendo* $\omega \leq F(x_n)$ *, así que siendo* F *continua por la derecha se tiene* $\omega \leq F(x)$ *.*

Si $Y(\omega) < x$ *existe* $x' < x$ *tal que* $\omega \leq F(x')$ *y como* F *es no decreciente se tiene también* $\omega \leq F(x)$ *.*

Por otra parte:

Si $\omega \leq F(x)$ *por definición se tiene* $Y(\omega) \leq x$ *.*

Así que entonces, si F_Y es la función de distribución de Y se tiene:

$$F_Y(x) = P[w \in (0,1): Y(w) \leq x] = P[(0, F(x)] \cap (0,1)] = F(x)$$

De manera análoga, se tiene:

$$\text{(II)} \quad \{w \in (0,1): Z_n(w) \leq x\} = \{w \in (0,1): w \leq F_n(x)\} = (0, F(x)] \cap (0,1)$$

De esta manera si F_{Z_n} es la función de distribución de Z_n obtenemos: $F_{Z_n}(x) = F_n(x)$

Sea ahora $w \in (0,1)$ punto de continuidad de Y.

Dado $\varepsilon > 0$ tomemos $x \in \mathbb{R}$ punto de continuidad de F tal que

$Y(w) - \varepsilon < x < Y(w)$ se tiene entonces: $F(x) < w$

y como $F_n(x) \longrightarrow F(x)$ por hipótesis entonces, existe N tal que $F_n(x) < w \quad \forall \; n \geq N$

Por lo tanto, usando (II):

$$x \leq Z_n(w) \quad \forall \; n \geq N$$

y entonces:

$$\liminf_n Z_n(w) \geq x > Y(w) - \varepsilon$$

y como ε fué arbitrario:

$$\liminf_n Z_n(w) \geq Y(w)$$

Ahora:

Dado $\varepsilon > 0$ tomemos $w' > w$ tal que $Y(w') < Y(w) + \varepsilon$

Tomemos entonces $y \in \mathbb{R}$ punto de continuidad de F tal que:

$$Y(w') < y < Y(w) + \varepsilon$$

Como $Y(w') < y$ se tiene por la relación (I): $w' < F(y)$

Y entonces $\omega < F(y)$

Como $F_n(y) \longrightarrow F(y)$ *existe* N *tal que*

$$\omega < F_n(y) \qquad \forall\, n \geq N$$

Y entonces por (II):

$$Z_n(\omega) \leq y \qquad \forall\, n \geq N$$

$$\therefore\; Z_n(\omega) < Y(\omega) + \varepsilon \qquad \forall\, n \geq N$$

Por lo tanto: $\lim\sup_n Z_n(\omega) \leq Y(\omega) + \varepsilon$

Y como ε *fué arbitrario*

$$\lim\sup_n Z_n(\omega) \leq Y(\omega)$$

concluimos entonces que:

$Z_n(\omega) \longrightarrow Y(\omega)$ $\forall\, \omega \in (0,1)$ *punto de continuidad de* Y

Como Y *es una función monótona no decreciente su conjunto de discontinuidad es a lo más numerable.*

Definamos entonces:

$$Y(\omega) = \begin{cases} Z(\omega) & \text{si } Z \text{ es continua en } \omega \\ 0 & \text{en otro caso} \end{cases}$$

$$Y_n(\omega) = \begin{cases} Z_n(\omega) & \text{si } Z_n \text{ es continua en } \omega \\ 0 & \text{en otro caso} \end{cases}$$

Entonces:

$Y = Z$ P- casi seguramente.

$Y_n = Z_n$ P- casi seguramente.

$$\therefore\; F_Y = F_Z = F$$

$$F_{Y_n} = F_{Z_n} = F_n$$

Y evidentemente $Y_n(\omega) \longrightarrow Y(\omega) \quad \forall\, \omega \in \Omega.$

Definición 2. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y γ una familia de variables aleatorias reales definidas en ese espacio. Diremos que la familia γ es uniformemente integrable si se satisface la siguiente relación:

$$\lim_{\alpha\to\infty} \sup_{X\in\gamma} \int_{[|X|\geq\alpha]} |X|\, dP = 0$$

En el apéndice se dan condiciones para que una familia de variables aleatorias sea uniformemente integrable.

Teorema 3. Sean X_n $(n=1,2,\ldots)$ y X variables aleatorias definidas en un espacio de probabilidad $(\Omega_1, \mathcal{F}_1, P_1)$. Supongamos que la familia $\{X_n\}$ es uniformemente integrable y que $X_n \xrightarrow{D} X$, entonces X es integrable y $E(X_n)\to E(X)$

Demostración. Sea F_n la función de distribución de X_n para cada n y sea F la función de distribución de X.

Sean Y_n $(n=1,2,\ldots)$ y Y las variables dadas por el Teorema de Skorohod, definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y tales que Y_n tiene a F_n como función de distribución para cada n, Y tiene a F como función de distribución y $Y_n(\omega)\to Y(\omega)$ $\forall\, \omega\in\Omega$.

si $\alpha\in\mathbb{R}$

$$\int_{[|X_n|\geq\alpha]} |X_n|\, dP_1 = \int_{\{x\in\mathbb{R}:|x|\geq\alpha\}} |x|\, dF_n(x) = \int_{\{|Y_n|\geq\alpha\}} |Y_n|\, dP$$

Entonces, la familia $\{X_n\}$ es uniformemente integrable si y solo si lo es la familia $\{Y_n\}$.

Además:

$$E\left[|X|\right] = \int_{\mathbb{R}} |x|\, dF(x) = E\left[|Y|\right]$$

Por lo tanto, X *es integrable si y solo si lo es* Y.

Finalmente:

$$E(X_n) = \int_{\mathbb{R}} x \, dF_n(x) = E(Y_n)$$

$$E(X) = \int_{\mathbb{R}} x \, dF(x) = E(Y)$$

$$\therefore E(X_n) \to E(X) \quad \text{si y solo si} \quad E(Y_n) \to E(Y)$$

El resultado se sigue entonces inmediatamente aplicando el teorema 3 del apéndice a la sucesión $\{Y_n\}$ *y a* Y.

Ya con estos elementos podemos pasar a demostrar los teoremas de continuidad. Estos nos dan condiciones para que una sucesión de variables aleatorias $\{X_n\}$ *converja en distribución a una variable aleatoria* X.

Teorema 4. Sean X_n $(n=1,2,\ldots)$ y X variables aleatorias definidas sobre un espacio de probabilidad $(\Omega_1, \mathcal{F}_1, P_1)$. Supongamos que cada X_n y X tienen momentos de todos los ordenes, que la distribución de X está determinada por sus momentos y que $E[X_n^r] \longrightarrow E[X^r]$ $\forall\, r \in \mathbb{N}$.

Entonces $X_n \xrightarrow{D} X$.

Demostración. Sea F_n la función de distribución de X_n para cada n y F la función de distribución de X.

Sea (F_{n_k}) una subsucesión débilmente convergente de (F_n) y sea G la función de distribución límite de esa subsucesión.

Como la sucesión $(E[X_n^2])$ converge, entonces está acotada por una constante K, así que usando la desigualdad de Chebyshev, podemos escribir:

$$P\left(|X_n| \geq x\right) \leq \frac{1}{x^2} E\left[X_n^2\right] \leq \frac{K}{x^2} \qquad \forall\; x \in \mathbb{R}$$

Es decir:

$$P\left(-x < X_n < x\right) \geq 1 - \frac{K}{x^2} \qquad \forall\; x \in \mathbb{R}$$

Por lo tanto, obtenemos:

$$F_n(x) - F_n(-x) \geq 1 - \frac{K}{x^2} \qquad \forall\; x \in \mathbb{R}$$

Dado $\varepsilon > 0$ podemos tomar entonces $x_0 > 0$ tal que:

$$F_n(x_0) - F_n(-x_0) > 1 - \varepsilon$$

De lo cual concluimos que la sucesión $\{F_n\}$ es uniforme.

Por lo tanto, aplicando la proposición 4 del capítulo III, la convergencia de F_{n_k} a G es propia, es decir G es una distribución de probabilidad.

El teorema de Skorohod nos asegura ahora la existencia de un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y de variables aleatorias Y_{n_k} $(k=1,2,\ldots)$ y Y definidas sobre ese espacio tales que:

Y_{n_k} tiene como distribución a F_{n_k}

Y tiene como distribución a G

$$Y_{n_k}(w) \to Y(w) \quad \forall\, w \in \Omega .$$

Ahora bien:

$$E\left[Y_{n_k}^r\right] = \int_{\mathbb{R}} x^r dF_{n_k}(x) = E\left[X_{n_k}^r\right] \quad \forall\; k, r \in \mathbb{N}$$

Entonces:

$$E\left[Y_{n_k}^r\right] \to E\left[X^r\right] \quad \forall\; r \in \mathbb{N}$$

En particular, la sucesión $\left(E\left[Y_{n_k}^r\right]\right)$ está acotada $\forall r \in \mathbb{N}$.

Sea ahora $r \in \mathbb{N}$ fijo y tomemos s un natural por tal que $s > r$. Entonces si $\alpha > 1$ tenemos:

$$E\left[Y_{n_k}^s\right] \geq \int_{[|Y_{n_k}|>\alpha]} Y_{n_k}^s \, dP = \int_{[|Y_{n_k}|>\alpha]} \left|Y_{n_k}^{s-r}\right| \left|Y_{n_k}^r\right| dP$$

$$\geq \alpha^{s-r} \int_{[|Y_{n_k}|>\alpha]} \left|Y_{n_k}^r\right| dP \geq \alpha^{s-r} \int_{[|Y_{n_k}^r|>\alpha]} \left|Y_{n_k}^r\right| dP$$

Por lo tanto:

$$\int_{[|Y^r_{n_k}|>\alpha]} |Y^r_{n_k}|\, dP \leq \frac{1}{\alpha^{s-r}} E\left[Y^s_{n_k}\right]$$

Si C *es una cota de la sucesión* $(E[Y^s_{n_k}])$ *obtenemos entonces:*

$$\int_{[|Y^r_{n_k}|>\alpha]} |Y^r_{n_k}|\, dP \leq \frac{C}{\alpha^{s-r}} \quad \forall\ k$$

$$\therefore \lim_{\alpha\to\infty} \left(\sup_k \int_{[|Y^r_{n_k}|>\alpha]} |Y^r_{n_k}|\, dP \right) = 0$$

Es decir, la familia $\{Y^r_{n_k}\}$ *es uniformemente integrable.*
Además, por la proposición I $Y^r_{n_k} \xrightarrow{D} Y^r$.
Entonces, aplicando el teorema 3 obtenemos que Y^r *es integrable y* $E[Y^r_{n_k}] \longrightarrow E[Y^r]$
Pero como $E[Y^r_{n_k}] \to E[X^r]$, *obtenemos:* $E[Y^r]=E[X^r]$
$\forall\ r\in\mathbb{N}$.

Y como la distribución de X *está determinada por sus momentos, concluimos que* X *tiene como función de distribución a* G .
Se concluye entonces el resultado aplicando el teorema 4 del capítulo III.

Teorema 5. Sean X_n $(n=1,2,\dots)$ y X variables aleatorias definidas en un espacio de probabilidad $(\Omega_1, \mathcal{F}_1, P_1)$.
Sea M_n la función generatriz de X_n para cada n y sea M la función generatriz de X . Supongamos que todas las M_n y M existen en una vecindad común $V(0)$ de cero y que M_n converge puntualmente a M en $V(0)$. Entonces $X_n \xrightarrow{D} X$

Demostración. Sea F_n la función de distribución de X_n para cada n y F la función de distribución de X .
Sea (F_{n_k}) una subsucesión débilmente convergente de (F_n) y sea G la función de distribución límite de esa subsucesión.
Demostremos primero que G es una distribución de probabilidad mostrando que la familia $\{F_n\}$ es uniforme.
Sea $[-s_0, s_0]$ con $s_0>0$ un intervalo contenido en $V(0)$ y sea $a>0$, tenemos entonces:

$$M_n(s_0) + M_n(-s_0) = \int_{-\infty}^{\infty} e^{s_0 x}\, dF_n(x) + \int_{-\infty}^{\infty} e^{-s_0 x}\, dF_n(x)$$

$$\geq \int_{(-a,a)^c} \left(e^{s_0 x} + e^{-s_0 x}\right) dF_n(x)$$

Pero si $x \in (-a,a)^c$:

$$e^{s_0 x} + e^{-s_0 x} \geq e^{s_0 a}$$

$$\therefore \quad M_n(s_0) + M_n(-s_0) \geq e^{s_0 a} \int_{(-a,a)^c} dF_n(x) = e^{s_0 a} F_n\{(-a,a)^c\}$$

Es decir:

$$F_n\{(-a,a)^c\} \leq e^{-s_0 a}\left[M_n(s_0) + M_n(-s_0)\right]$$

Pero por hipótesis, las sucesiones $\left(M_n(s_0)\right)$ *y* $\left(M_n(-s_0)\right)$ *son convergentes, así que en particular están acotadas.*
Sea K *una cota de la suma* $M_n(s_0) + M_n(-s_0)$

Entonces:

$$F_n\left\{(-a,a)^c\right\} \leq K e^{-s_0 a}$$

Ahora dado $\varepsilon>0$ *tomemos* a *tal que* $K e^{-s_0 a} < \varepsilon$

entonces $F_n\left\{(-a,a)\right\} > 1-\varepsilon$

Así que por la proposición 3 del capítulo III la familia $\{F_n\}$ *es uniforme.*

Entonces usando la proposición 4 del capítulo III obtenemos que:

$$F_{n_k} \xrightarrow{d} G$$

donde G *es una distribución de probabilidad.*

El teorema de Skorohod nos asegura la existencia de un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ *y de variables aleatorias* Y_{n_k} $(k=1,2,\ldots)$ *y* Y *definidas sobre ese espacio tales que:*

Y_{n_k} *tiene como función de distribución a* F_{n_k}

Y *tiene como función de distribución a* G

$$Y_{n_k}(w) \longrightarrow Y(w) \qquad \forall \quad w \in \Omega.$$

Ahora bien, tenemos:

$$E\left(e^{sX_{n_k}}\right) = \int_{-\infty}^{\infty} e^{sx}\, dF_{n_k}(x) = E\left(e^{sY_{n_k}}\right)$$

Es decir, X_{n_k} *y* Y_{n_k} *tienen la misma función generatriz.*

Por lo tanto:

$$E\left(e^{sY_{n_k}}\right) \longrightarrow M(s) \qquad \forall \ s \in V(0)$$

El resultado buscado quedará entonces demostrado probando que

$E(e^{sY_{n_k}}) \longrightarrow E(e^{sY}) \quad \forall s \in V(0)$ porque entonces tendríamos

$$M(s) = E(e^{sY}) \qquad \forall \quad s \in V(0)$$

Y como la función generatriz de una variable aleatoria, cuando existe en una vecindad del 0, determina de manera única la distribución de la variable aleatoria, tendríamos $F = G$ y por lo tanto $F_n \xrightarrow{d} F$ de acuerdo al teorema 4 del capítulo III.

En base al teorema 3 del apéndice, para probar que

$E(e^{sY_{n_k}}) \longrightarrow E(e^{sY})$ basta con demostrar que la familia

$\{e^{sY_{n_k}} \quad k=1,2,\ldots\}$ es uniformemente integrable.

Para probar este último resultado usaremos el teorema 4 del apéndice. De hecho se puede probar que la familia

$\{e^{sY_n} \quad n=1,2,\ldots\}$ es uniformemente integrable $\forall s \in V(0)$ y esto es lo que demostramos a continuación:

Es claro que podemos suponer que $V(0)$ es un intervalo abierto.

Sea entonces $s \in V(0)$

Si $s=0$ es trivial ver que $\{e^{sY_n} \quad n=1,2,\ldots\}$ es uniformemente integrable.

Tomemos entonces $s \neq 0$

Si $s>0$ tomemos $s' \in V(0)$ tal que $s'>s$

Si $s<0$ tomemos $s' \in V(0)$ tal que $s'<s$

Definamos $\delta = \frac{s'}{s} - 1$

Sea entonces $H: \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ definida por:

$$H(\tau) = \tau^{1+\delta}$$

Como $\delta > 0$ *se tiene evidentemente:*

$$\lim_{\tau \to \infty} \frac{H(\tau)}{\tau} = +\infty$$

Además, si $z_n = e^{s y_n}$ *tenemos:*

$$H(z_n) = e^{s(1+\delta) y_n} = e^{s' y_n}$$

Pero:

$$E(e^{s' y_n}) = E(e^{s' x_n}) \longrightarrow M(s')$$

En particular, la sucesión $\{E(e^{s' y_n})\}$ *está acotada y entonces:*

$$\sup_n \{E(e^{s' y_n})\} < \infty$$

Así que por el teorema 4 del apéndice la familia $\{e^{s' y_n}\}$ *es uniformemente integrable.*

Teorema 6. Sea $\{F_n\}$ una sucesión de distribuciones de probabilidad concentradas en $[0, \infty)$.
Sea φ_n la transformada de Laplace de F_n para cada n, entonces las siguientes condiciones son equivalentes:

(I) $\varphi_n(s)$ converge $\forall\ s \geq 0$

(II $\{F_n\}$ converge débilmente

Además, si $F_n \xrightarrow{d} F$ y φ es la transformada de Laplace de F , entonces $\varphi_n(s) \longrightarrow \varphi(s) \quad \forall\ s \geq 0$

Demostración. (II) $\Longrightarrow$ (I)] Supongamos que $F_n \xrightarrow{d} F$

Evidentemente F debe estar concentrada en $[0, \infty)$.

Sea entonces φ la transformada de Laplace de F .

Para $s \geq 0$ definamos $g: \mathbb{R} \longrightarrow \mathbb{R}$ de la siguiente manera:

$$g(x) = e^{-s|x|}$$

Entonces:

$$\int_{\mathbb{R}} g(x)\, dF_n(x) = \varphi_n(s)$$

$$\int_{\mathbb{R}} g(x)\, dF(x) = \varphi(s)$$

Además g es una función continua que se anula en el infinito, así que por el teorema 2 del capítulo III, se tiene:

$$\varphi_n(s) \longrightarrow \varphi(s)$$

(I) $\Longrightarrow$ (II)]

Supongamos que $\varphi_n(s)$ converge $\forall\ s \geq 0$

Sea (F_{n_k}) una subsucesión débilmente convergente de (F_n) y sea G la distribución límite de esa subsucesión.

Por la primera parte de la demostración, la sucesión de transformadas de Laplace de F_{n_k} converge puntualmente a la transformada de Laplace de G. Entonces si φ es la transformada de Laplace de G, se tiene:

$$\varphi(s) = \lim_{k} \varphi_{n_k}(s) = \lim_{n} \varphi_n(s) \qquad \forall\ s \geq 0$$

Así que cualquiera que sea la subsucesión débilmente convergente la transformada de Laplace de la distribución límite está dada por $\lim_{n} \varphi_n(s)$.

Ahora, como la transformada de Laplace de una distribución determina a ésta (teorema 3 del capítulo IV) se sigue que todas las subsucesiones débilmente convergentes convergen a la misma distribución límite G. Por lo tanto, usando el teorema 4 del capítulo III, concluimos que $F_n \xrightarrow{d} G$.

<u>Corolario I.</u> Sean X_n $(n=1,2,\ldots)$ y X variables aleatorias Sean F_n la función de distribución de X_n para cada n y F la función de distribución de X.
Supongamos que todas las F_n y F están concentradas en $[0,\infty)$ y sea φ_n la transformada de Laplace de F_n para cada n y φ la transformada de Laplce de F.
Entonces las siguientes condiciones son equivalentes:

(I) $\varphi_n(s) \longrightarrow \varphi(s) \quad \forall\ s \geq 0$

(II) $X_n \xrightarrow{D} X$

APENDICE

A. Teorema Límite de Abel

Teorema I. "Teorema Límite de Abel"
Supongamos que la serie $f(x) = \sum_{n=0}^{\infty} a_n x^n$ converge para el intervalo $-1 < x < 1$.

Si la serie converge también para $x = 1$ entonces el límite $\lim_{x \to 1^-} f(x)$ existe y se tiene: $\lim_{x \to 1^-} f(x) = \sum_{n=0}^{\infty} a_n$

Demostración.
Definamos $f(1) = \sum_{n=0}^{\infty} a_n$

Para $0 < x < 1$ multipliquemos la serie $\sum_{n=0}^{\infty} a_n x^n$ por la serie geométrica $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. Obtenemos:

$$\frac{1}{1-x} f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad \text{donde} \qquad c_n = \sum_{k=0}^{n} a_k$$

Entonces tenemos:

$$\text{(A)} \quad f(x) - f(1) = (1-x) \sum_{n=0}^{\infty} [c_n - f(1)]\, x^n \qquad \forall\ 0 < x < 1$$

Pero $\lim_{n \to \infty} c_n = f(1)$, por lo tanto, dada $\varepsilon > 0$ existe N tal que $x \geqslant N$ implica $|c_n - f(1)| < \varepsilon/2$

Dividamos entonces la sumatoria de (A) de la siguiente manera:

$$\text{(B)} \quad f(x) - f(1) = (1-x) \sum_{n=0}^{N-1} [c_n - f(1)]\, x^n + (1-x) \sum_{n=N}^{\infty} [c_n - f(1)]\, x^n$$

Sea: $M = \text{máx} \{ |c_n - f(1)| : n = 0, 1, 2, \dots, N-1 \}$

Entonces, a partir de (B) obtenemos:

$$|f(x)-f(1)| \leq (1-x)\, M \sum_{n=0}^{N-1} x^n + (1-x)\, \frac{\varepsilon}{2} \sum_{n=N}^{\infty} x^n$$

$$\leq (1-x) M N + (1-x)\, \frac{\varepsilon}{2}\, \frac{x^N}{1-x} \leq (1-x) M N + \frac{\varepsilon}{2}$$

Sea entonces $0 < \delta < 1$ *tal que* $\delta < \dfrac{\varepsilon}{2MN}$

Obtenemos:

$$|f(x) - f(1)| < \varepsilon \quad \forall \quad 1 - \delta < x < 1$$

Es decir: $\lim\limits_{x \to 1^-} f(x) = f(1)$

lo cual demuestra el resultado.

(B) INTEGRABILIDAD UNIFORME Y TEOREMAS DE CONVERGENCIA

Teorema 2. "Teorema de la convergencia Dominada"

Sean X_n $(n=1,2,\ldots)$ y Y funciones medibles reales definidas en un espacio medible $(E, \mathcal{Q}, \mu)$.

Supongamos que Y es integrable, que $|X_n| \leq Y$ casi en todas partes y que $X_n \longrightarrow X$ casi en todas partes, entonces X y las X_n son integrables y se tiene que $\int X_n\, d\mu \longrightarrow \int X\, d\mu$

Demostración. Ya que $|X_n| \leq Y$ casi en todas partes, X y todas las X_n son integrables.

Sea $Z_n = |X_n - X|$ y $Z = 2Y$, entonces $Z_n \to 0$ y $0 \leq Z_n \leq Z$ casi en todas partes, además Z es integrable y como sabemos que:

$$\left| \int X_n\, d\mu - \int X\, d\mu \right| \leq \int |X_n - X|\, d\mu$$

es suficiente probar que

$$\int Z_n\, d\mu \longrightarrow 0$$

Por el Lema de Fatou (H.L. Royden, Real Analysis. Second Edition, Collier Macmillan International Editions 1968 p. 83).

tenemos:

$$\int Z\, d\mu = \int \text{Lím} \inf_n (Z - Z_n)\, d\mu \leq \text{Lím} \inf_n \int (Z - Z_n)\, d\mu =$$

$$= \int Z\, d\mu - \text{Lím} \sup_n \int Z_n\, d\mu$$

Por tanto

$$\int Z\, d\mu \leq \int Z\, d\mu - \text{Lím} \sup_n \int Z_n\, d\mu \leq \int Z\, d\mu$$

Entonces

$$\text{Lím} \sup_n \int Z_n \, d\mu = 0.$$

Así que

$$\int Z_n \, d\mu \longrightarrow 0.$$

Los siguientes dos teoremas admiten un inverso, sin embargo se enuncia y demuestra solo una parte de cada uno porque es la única que se utiliza en este trabajo.
Los resultados completos pueden consultarse en C. Dellacherie et P.A. Meyer, Probabilités et potentiel , Ch.I,p.33-39.

__*Teorema 3.*__ *Sean* $X_n \ (n=1,2,\dots)$ *y* X *variables aleatorias reales definidas en un espacio de probabilidad* $(\Omega, \mathcal{F}, P)$:
Supongamos que X_n *es integrable para cada n y que* $X_n \to X$ *casi en todas partes. Entonces si la familia* $\{X_n\}$ *es uniformemente integrable se obtiene que* X *es integrable y* $E[|X_n - X|] \to 0$.

__*Demostración.*__

Sea $\alpha > 0$ *, entonces:*

$$E[|X_n|] = \int_{\{|X_n| \geq \alpha\}} |X_n| \, dP + \int_{\{|X_n| < \alpha\}} |X_n| \, dP$$

$$\leq \alpha P[|X_n| < \alpha] + \int_{\{|X_n| \geq \alpha\}} |X_n| \, dP$$

$$\leq \alpha + \int_{\{|X_n| \geq \alpha\}} |X_n| \, dP$$

Tenemos α_0 *tal que:*

$$\int_{\{|X_n| \geq \alpha_0\}} |X_n| \, dP \leq 1 \qquad \forall n$$

lo cual es posible porque la familia $\{X_n\}$ *es uniformemente integrable.*

Entonces

$$E\left[|X_n|\right] \leq \alpha_0 + 1 \quad \forall n$$

Así que usando el lema de Fatou:

$$E\left[|X|\right] = E\left[\text{Lím } |X_n|\right] \leq \text{Lím inf } E\left[|X_n|\right] \leq \alpha_0 + 1 < \infty$$

Es decir, X *es integrable.*

Por otra parte, si $\alpha > 0$

Sea $A = \left[|X_n| \leq \alpha \,;\; |X| \leq \alpha\right]$

Entonces:

$$A^c = \left[|X_n| > \alpha \,;\, |X| \leq \alpha\right] \cup \left[|X| > \alpha\right]$$

$$= \left[|X| > \alpha \,;\; |X_n| \leq \alpha\right] \cup \left[|X_n| > \alpha\right]$$

Así que: $E\left[|X_n - X|\right] = \int_A |X_n - X|\, dP + \int_{A^c} |X_n - X|\, dP$

$$\leq \int_A |X_n - X|\, dP + \int_{A^c} |X_n|\, dP + \int_{A^c} |X|\, dP$$

$$= \int_A |X_n - X|\, dP + \int_{[|X| > \alpha \,;\, |X_n| \leq \alpha]} |X_n|\, dP + \int_{[|X_n| > \alpha]} |X_n|\, dP$$

$$+ \int_{[|X_n| > \alpha \,;\, |X| \leq \alpha]} |X|\, dP + \int_{[|X| > \alpha]} |X|\, dP$$

Definamos:

$$X_n^{\alpha} = \begin{cases} X_n & \text{si } |X_n| \leq \alpha \\ 0 & \text{en otro caso} \end{cases}$$

$$X^{\alpha} = \begin{cases} X & \text{si } |X| \leq \alpha \\ 0 & \text{en otro caso} \end{cases}$$

Entonces:

$$\int_A |X_n - X|\, dP + \int_{[|X|>\alpha;\, |X_n|\leq\alpha]} |X_n|\, dP + \int_{[|X_n|>\alpha;\, |X|\leq\alpha]} |X|\, dP$$

$$= E\left[\left| X_n^{\alpha} - X^{\alpha} \right|\right]$$

Por lo tanto:

$$(A)\; E\left[|X_n - X|\right] \leq E\left[|X_n^{\alpha} - X^{\alpha}|\right] + \int_{[|X_n|>\alpha]} |X_n|\, dP + \int_{[|X|>\alpha]} |X|\, dP$$

Ahora bien, si α *es tal que* $P\left[|X| = \alpha\right] = 0$, *tenemos:*

$$X_n^{\alpha} \longrightarrow X^{\alpha}$$

casi seguramente

Además: $\left| X_n^{\alpha} - X^{\alpha} \right| \leq \alpha$

Entonces por el teorema de la convergencia dominada obtenemos:

$$E\left[\left| X_n^{\alpha} - X^{\alpha} \right|\right] \longrightarrow 0$$

Dada $\varepsilon>0$ *, tomemos entonces* $\alpha_0>0$ *tal que:*

$$\int_{[|X_n|>\alpha]} |X_n|\,dP < \varepsilon \qquad \forall \quad \alpha \geq \alpha_0 \qquad \forall n$$

$$\int_{[|X|>\alpha]} |X|\,dP < \varepsilon \qquad \forall \quad \alpha \geq \alpha_0$$

Ahora, el conjunto $A=\left\{\alpha\in \mathbb{R} : P\left[|X|=\alpha\right] > \frac{1}{m}\right\}$ *es finito* $\forall\, m\in \mathbb{N}$ *pues de otra manera se tendría* $P(\Omega)=P[X\in \mathbb{R}]=\infty$

Entonces el conjunto $\left\{\alpha\in\mathbb{R} : P\left[|X|=\alpha\right]>0\right\}$ *es a lo más numerable.*

Existe entonces $\alpha\geq\alpha_0$ *tal que* $\left\{\alpha\in\mathbb{R} : P\left[|X|=\alpha\right]>0\right\}$.

Para esta α *sea* N *tal que:*

$$E\left[\left|X_n^{\alpha}-X^{\alpha}\right|\right] < \varepsilon \qquad \forall \quad n\geq N$$

Entonces de (A) obtenemos

$$E\left[|X_n-X|\right] < 3\varepsilon \qquad \forall \quad n\geq N$$

$$\therefore \; E\left[|X_n-X|\right] \longrightarrow 0$$

Teorema 4. " *Teorema de la Vallée-Poussin*"

Sea $(\Omega, \mathcal{F}, P)$ *un espacio de probabilidad y* $\mathcal{H}$ *una familia de variables aleatorias reales definidas en ese espacio e integrables. Supongamos que existe una función* $H: \mathbb{R}^+ \to \mathbb{R}^+$ *con las siguientes dos propiedades:*

a) $\lim\limits_{t \to \infty} \dfrac{H(t)}{t} = +\infty$

b) $\sup\limits_{X \in \mathcal{H}} \left\{ E\left[H(|X|)\right] \right\} < \infty$

entonces la familia $\mathcal{H}$ *es uniformemente integrable.*

Demostración.

Sea $M = \sup\limits_{X \in \mathcal{H}} \left\{ E\left[H(|X|)\right] \right\}$

Dada $a > 0$ *sea* $\alpha_0 > 0$ *tal que:*

$$\frac{H(t)}{t} > a \qquad \forall \quad \alpha \geq \alpha_0$$

Entonces:

$$\frac{H(|X|)}{|X|} > a \qquad \text{sobre } [|X| \geq \alpha] \quad \forall \ X \in \mathcal{H}, \ \forall \alpha \geq \alpha_0$$

Así que, si $X \in \mathcal{H}$ *:*

$$\int_{[|X| \geq \alpha]} |X| \, dP \leq \frac{1}{a} \int_{[|X| \geq \alpha]} H(|X|) \, dP \leq \frac{1}{a} E\left[H(|X|)\right] \leq \frac{1}{a} M \quad \forall \ \alpha \geq \alpha_0$$

Por lo tanto:

$$\sup_{X \in \mathcal{H}} \int_{[|X| \geq \alpha]} |X| \, dP \leq \frac{M}{a} \qquad \forall \ \alpha \geq \alpha_0$$

Dada $\varepsilon > 0$*, tomemos entonces* $a = \dfrac{M}{\varepsilon}$ *, de manera que considerando la correspondiente* α_0 *, obtenemos:*

$$\sup_{X \in \mathcal{X}} \int_{[|X| \geq \alpha]} |X| \, dP \leq \varepsilon \qquad \forall \quad \alpha \geq \alpha_0 \tag{A}$$

Es decir, dada $\varepsilon > 0$ *existe* α_0 *satisfaciendo (A); por lo tanto:*

$$\lim_{\alpha \to \infty} \sup_{X \in \mathcal{X}} \int_{[|X| \geq \alpha]} |X| \, dP = 0$$

Así que la familia $\mathcal{X}$ *es uniformemente integrable.*

C. INTERCAMBIO DE INTEGRAL Y DERIVADA

Teorema 5.
Sea $(E, \mathcal{Q}, \mu)$ *un espacio medible. Sea* I *un intervalo abierto en* $\mathbb{R}$ *y* $f: E \times I \longrightarrow \mathbb{R}$ *una función con las siguientes propiedades:*

a) Para cada $\tau \in I$ *la función* $x \longrightarrow f(x, \tau)$ *es medible e integrable.*

b) Existe $A \subset E$ *con* $\mu(A) = 0$ *tal que la función* $\tau \longrightarrow f(x, \tau)$ *es derivable en* I $\forall x \notin A$.

c) Existe una función $g: E \longrightarrow \mathbb{R}^+$ *integrable tal que* $|f'(x, \tau)| \leq g(x) \quad \forall \ \tau \in I, \ x \notin A.$

Entonces, la función $h: \tau \longrightarrow \int_E f(x, \tau) \, \mu(dx)$ *es derivable en* I *y además:*

$$h'(\tau) = \int_E f'(x, \tau) \, \mu(dx) \qquad \forall \ \tau \in I$$

donde se ha definido $f'(x, \tau)$ *fuera de* A *de manera arbitraria.*

__*Demostración.*__ *Sea* $\tau \in I$ *y* (h_n) *una sucesión de números reales estrictamente positivos tal que* $h_n \longrightarrow 0$.

Sea $H_n : E \to \mathbb{R}$ *definida por:*

$$H_n(x) = \frac{f(x, \tau + h_n) - f(x, \tau)}{h_n}$$

Por el teorema del valor medio para $x \notin A$ *, existe* $s_n \in (\tau, \tau + h_n)$ *tal que:*

$$H_n(x) = f'(x, s_n)$$

Por lo tanto:

$|H_n| \leq g$ *casi en todas partes*

Además $H_n(x) \to f'(x, \tau) \quad \forall \; x \notin A$

Entonces, por el teorema de la convergencia dominada, obtenemos que la función $x \to f'(x, \tau)$ *así como todas las* H_n *son integrables y además:*

$$\int_E H_n(x)\, \mu(dx) \longrightarrow \int_E f'(x, \tau)\, \mu(dx)$$

Es decir:

$$\lim_{n \to \infty} \frac{1}{h_n}\left[\int_E f(x, \tau + h_n)\, \mu(dx) - \int_E f(x, \tau)\, \mu(dx)\right] = \int_E f'(x, \tau)\, \mu(dx)$$

Y como esto vale para toda sucesión (h_n) *de números reales estrictamente positivos tal que* $h_n \to 0$ *, obtenemos:*

$$\lim_{h \to 0} \frac{1}{h}\left[\int_E f(x, \tau + h)\, \mu(dx) - \int_E f(x, \tau)\, \mu(dx)\right] = \int_E f'(x, \tau)\, \mu(dx)$$

lo cual demuestra el resultado.

D. EL METODO DIAGONAL

Recordemos el siguiente resultado (Teorema de Bolzano- Weierstrass).

Si $X_1, X_n, \ldots$ *es una sucesión acotada de números reales entonces existe una sucesión creciente* $n_1, n_2, \ldots$ *de números naturales para los cuales el límite* $\lim_{K} X_{n_K}$ *existe.*

Es decir, es posible seleccionar una subsucesión convergente $X_{n_1}, X_{n_2}, \ldots$

Esta es una de las propiedades fundamentales de los números reales,

Veamos a continuación el siguiente Teorema.

__Teorema 6.__ *Supongamos que cada renglón del siguiente arreglo;*

$$\begin{array}{l} X_{11},\ X_{12},\ X_{13}, \ldots \\ X_{21},\ X_{22},\ X_{23}, \ldots \qquad\qquad (I) \\ \ldots\ldots\ldots\ldots \end{array}$$

es una sucesión acotada de números reales.

Entonces existe una sucesión creciente $n_1, n_2, n_3, \ldots$ *de números naturales tales que el límite* $\lim_{K} X_{r,n_K}$ *existe para r=I, 2, ...*

__Demostración.__Del primer renglón de (I), seleccionamos una subsucesión convergente:

(2) $X_{1,n_{11}}, X_{1,n_{12}}, X_{1,n_{13}}, \ldots$ *donde* $\{n_{1,K}\}$ *es una sucesión creciente de números naturales y* $\lim_{K} X_{1,n_{1K}}$ *existe, ahora bien a partir del segundo renglón de (I) tomemos la subsucesión:*

(3) $X_{2,n_{11}}, X_{2,n_{12}}, X_{2,n_{13}}, \ldots$

que claramente es acotada; seleccionamos de ésta una subsucesión convergente: $X_{2,n_{21}}, X_{2,n_{22}}, X_{2,n_{23}}$ donde $\{n_{2k}\}$ es una sucesión creciente de números naturales que a su vez es una subsucesión de $\{n_{1,k}\}$ y $\lim_K X_{2,n_{2K}}$ existe.
Continuando inductivamente de esta forma obtenemos el siguiente arreglo:

$$(4) \qquad \begin{array}{l} n_{11}, \; n_{12}, \; n_{13}, \dots \\ n_{21}, \; n_{22}, \; n_{23}, \dots \\ \dots\dots\dots\dots\dots \end{array}$$

Con tre propiedades:
A) Cada renglón de (4) es una sucesión creciente de números naturales.
B) El r-ésimo renglón es una subsucesión del (r-I)-ésimo.
C) Para cada r, $\lim_K X_{r,n_{rK}}$ existe.
de esta manera:

(5) $X_{r,n_{r1}}, X_{r,n_{r2}}, X_{r,n_{r3}}, \dots$

es una subsucesión convergente del r-ésimo renglón de (I),
Ahora sea $n_K = n_{KK}$, ya que cada renglón de (4) constituye una sucesión creciente y está contenido en el anterior renglón, $n_1, n_2, n_3, \dots$ es una sucesión creciente de números naturales.

Por lo tanto n_r, n_{r+1}, n_{r+2} es una subsucesión de (5) y es por lo tanto convergente.
De esta forma $\lim_K n_{r,n_K}$ existe $\forall\, r = 1, 2, \dots$

<u>Nota</u> Ya que $\{n_K\}$ es la diagonal del arreglo (4), aplicaciones de este Teorema son llamadas aplicaciones del Método Diagonal.

E. FUNCIONES DE DISTRIBUCION DETERMINADAS POR SUS MOMENTOS.

En esta parte del apéndice definimos lo que es la función característica de una función de distribución y demostramos algunas de sus propiedades. Sin embargo, no es nuestro objetivo hacer un estudio de la función característica; por tal motivo solo expondremos aquí lo estrictamente necesario para demostrar el teorema 8 el cual utilizamos en el estudio de la función generatriz.
Por lo mismo el principal resultado que exponemos sobre la función característica no es demostrado aquí, una prueba de éste puede encontrarse en cualquier texto de probabilidad que incluya el estudio de la función característica, en particular puede consultarse: Billingsley; sección 26.
El teorema 8 es utilizado específicamente para demostrar que una función de distribución queda determinada por su función generatriz cuando esta existe en una vecindad del 0.

Como se verá más adelante, la demostración del teorema 7 utiliza el hecho de que una función de distribución queda determinada por su función característica.

Así, nuestro resultado para funciones generatrices es demostrado a partir del resultado análogo para funciones características.

Este es el único resultado expuesto en la tesis que requiere del uso de la función característica. Hubieramos querido no hacer uso de esta última pero no encontramos en la literatura, ni logramos producir, una demostración directa del resultado mencionado para funciones generatrices.

Definición I. Sea X una variable aleatoria con función de distribución F. La función característica de X es la función: $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ definida por:

$$\varphi(\tau) = E\left[e^{i\tau X}\right] = \int_{\mathbb{R}} e^{i\tau x} dF(x)$$

El siguiente resultado, cuya demostración puede consultarse en Billingsley sección 26, es la principal propiedad de las funciones características que utilizaremos en este trabajo.

Teorema 7. Sean X, Y variables aleatorias con funciones de distribución F y G respectivamente. Si X, Y tienen la misma función característica entonces $F = G$.

Los siguientes dos lemas se utilizarán más adelante en la demostración del resultado principal de esta parte del apéndice.

Lema I. Sea $x \in \mathbb{R}$, $n \in \mathbb{Z}^+$, entonces:

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \frac{|x|^{n+1}}{(n+1)!}$$

Demostración.

Integrando por partes se obtiene:

$$(A) \quad \int_0^x (x-s)^n e^{is} ds = \frac{x^{n+1}}{n+1} + \frac{i}{n+1} \int_0^x (x-s)^{n+1} e^{is} ds \quad \forall x \in \mathbb{R},\ n = 0, 1, 2, \ldots$$

De aquí se sigue:

$$(B) \quad e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is} ds \quad \forall x \in \mathbb{R},\ n = 0, 1, 2, \ldots$$

En efecto, la fórmula vale trivialmente para $n = 0$ y si vale para m entonces:

$$e^{ix} = \sum_{k=0}^{m} \frac{(ix)^k}{k!} + \frac{i^{m+1}}{m!} \int_0^x (x-s)^m e^{is} ds$$

$$= \sum_{k=0}^{m} \frac{(ix)^k}{k!} + \frac{i^{m+1}}{m!} \left[\frac{x^{m+1}}{m+1} + \frac{i}{m+1} \int_0^x (x-s)^{m+1} e^{is} ds \right]$$

$$= \sum_{k=0}^{m+1} \frac{(ix)^k}{k!} + \frac{i^{m+2}}{(m+1)!} \int_0^x (x-s)^{m+1} e^{is} ds$$

Entonces por inducción concluimos que la fórmula (B) es válida $\forall \; n=0,1,2,\ldots$

De (B) obtenemos ahora:

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| = \left| \frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is} ds \right|$$

$$= \frac{1}{n!} \left| \int_0^x (x-s)^n e^{is} ds \right|$$

$$\leq \frac{1}{n!} \left| \int_0^x |x-s|^n \, |e^{is}| ds \right|$$

$$= \frac{1}{n!} \left| \int_0^x |x-s|^n ds \right|$$

$$= \frac{1}{n!} \, \frac{|x|^{n+1}}{n+1} = \frac{|x|^{n+1}}{(n+1)!}.$$

lo cual se quería demostrar.

Lema 8. Sea X una variable aleatoria con función de distribución F. Sea $K \in \mathbb{N}$ y supongamos que $E\left[|X|^K\right] < \infty$, entonces la k-ésima derivada $\varphi^{(K)}$ de la función característica de X satisface la relación:

$$\varphi^{(K)}(\tau) = E\left[i^K X^K e^{i\tau X}\right] = \int_{\mathbb{R}} i^K x^K e^{i\tau x}\, dF(x)$$

Demostración.

Como $E\left[|X|^K\right] < \infty$, entonces $E\left[|X|^J\right] < \infty \quad \forall\, J = 1, 2, \ldots, K$

Demostraremos entonces la siguiente relación:

$$(A) \quad \varphi^{(J)}(\tau) = \int_{\mathbb{R}} i^J x^J e^{i\tau x}\, dF(x) \qquad J = 0, 1, 2, \ldots, K$$

Esta relación vale trivialmente para $J = 0$.

Supongamos entonces que vale para $m < K$, entonces:

$$\varphi^{(m+1)}(\tau) = \lim_{h \to 0} \frac{\varphi^{(m)}(\tau + h) - \varphi^{(m)}(\tau)}{h}$$

$$= \lim_{h \to 0} \int_{\mathbb{R}} \frac{i^m x^m e^{i(\tau+h)x} - i^m x^m e^{i\tau x}}{h}\, dF(x)$$

$$= \lim_{h \to 0} \int_{\mathbb{R}} i^m x^m e^{i\tau x}\, \frac{e^{ihx} - 1}{h}\, dF(x)$$

Pero:

$$\lim_{h\to 0} \frac{e^{ihx}-1}{h} = \lim_{h\to 0} ix\, e^{ihx} = ix$$

Así que:

$$\lim_{h\to 0} i^m x^m e^{i\tau x} \frac{e^{ihx}-1}{h} = i^{m+1} x^{m+1} e^{i\tau x}$$

Además, usando el Lema I anterior con $n=0$, *obtenemos :*

$$|e^{ihx}-1| \leq |hx|$$

Por lo tanto:

$$\left| i^m x^m e^{i\tau x} \frac{e^{ihx}-1}{h} \right| \leq |x|^{m+1}$$

Entonces, como la función $x \to |x|^{m+1}$ *es integrable con respecto a* F *, podemos aplicar el teorema de la convergencia dominada para obtener:*

$$\varphi^{(m+1)}(\tau) = \int_{\mathbb{R}} \lim_{h\to 0} i^m x^m e^{i\tau x} \frac{e^{ihx}-1}{h}$$

$$= \int_{\mathbb{R}} i^{m+1} x^{m+1} e^{i\tau x}\, dF(x)$$

Así que la relación (A) vale para $m+1$.
Entonces por inducción (A) es válida para $J=1,2,\ldots,K$ *de lo cual se concluye el Lema 2.*

Teorema 8. Sea F una distribución de probabilidad sobre $\mathbb{R}$ tal que:

$$\int_{\mathbb{R}} |x|^k \, dF(x) < \infty \qquad \forall \; k = 1, 2, \ldots$$

Sea $\alpha_k = \int_{\mathbb{R}} x^k \, dF(x)$ el momento de orden k de F y supongamos que la serie de potencias $\sum_{k=1}^{\infty} \frac{\alpha_k \mu^k}{k!}$ tiene un radio positivo de convergencia, entonces F es la única distribución de probabilidad con momentos $\alpha_0, \alpha_1, \ldots$

Demostración.
Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad sobre el cual podemos definir una variable aleatoria X con función de distribución F y sea φ la función característica de X. Demostraremos primero que φ tiene un desarrollo en serie de Taylor alrededor de cualquier punto $\tau \in \mathbb{R}$.

En efecto; usando los dos lemas anteriores podemos escribir:

$$\left| \varphi(\tau+h) - \sum_{k=0}^{n} \frac{\varphi^{(k)}(\tau)\, h^k}{k!} \right|$$

$$= \left| \int_{\mathbb{R}} e^{i(\tau+h)x} \, dF(x) - \sum_{k=0}^{n} \frac{h^k}{k!} \int_{\mathbb{R}} (ix)^k e^{i\tau x} \, dF(x) \right|$$

$$= \left| \int_{\mathbb{R}} \left[e^{i(\tau+h)x} - \sum_{k=0}^{n} \frac{h^k}{k!} (ix)^k e^{i\tau x} \right] dF(x) \right|$$

$$\leq \int_{\mathbb{R}} |e^{itx}| \left| e^{ihx} - \sum_{k=0}^{n} \frac{h^k}{k!} (ix)^k \right| dF(x)$$

$$\leq \int_{\mathbb{R}} \frac{|hx|^{n+1}}{(n+1)!} dF(x) = \frac{|h|^{n+1}}{(n+1)!} \int_{\mathbb{R}} |x|^{n+1} dF(x)$$

Vamos a probar que tomando h *suficientemente pequeño esta última expresión tiende a cero cuando* $n \to \infty$.

Sea $\beta_n = \int_{\mathbb{R}} |x|^n dF(x)$

Como la serie de potencias $\sum_{k=1}^{\infty} \frac{\alpha_k u^k}{k!}$ *tiene un radio positivo de convergencia, sabemos que existe* $0 < s < 1$ *tal que la serie* $\sum_{k=1}^{\infty} \frac{\alpha_k u^k}{k!}$.

converge $\forall\ |u| \leq s$

En particular:

$$\lim_{k \to \infty} \frac{\alpha_k s^k}{k!} = 0$$

Sea $0 < r < s$ *entonces:*

$$\lim_{k \to \infty} k \left(\frac{r}{s}\right)^k = 0$$

Por lo tanto existe $k_0 \in \mathbb{N}$ *tal que:*

$$k \left(\frac{r}{s}\right)^k < \frac{r}{2} \qquad \forall\ k \geq k_0$$

Es decir:

$$\gamma^{k-1} < \frac{5^k}{2k} \qquad \forall \quad k \geq k_0$$

Además:

$$|x|^{k-1} \leq 1 + |x|^k \qquad \forall \; k \in \mathbb{N} \;, \forall \; x \in \mathbb{R}$$

Por lo tanto:

$$B_{k-1} \leq 1 + B_k \qquad \forall \quad k \in \mathbb{N}$$

De manera que:

$$\frac{B_{2k-1}\,\gamma^{2k-1}}{(2k-1)!} \leq \frac{\gamma^{2k-1}}{(2k-1)!} + \frac{B_{2k}\,\gamma^{2k-1}}{(2k-1)!}$$

$$\leq \frac{\gamma^{2k-1}}{(2k-1)!} + \frac{B_{2k}\,5^{2k}}{(2k)!} = \frac{\gamma^{2k-1}}{(2k-1)!} + \frac{\alpha_{2k}\,5^{2k}}{(2k)!}$$

Y entonces:

$$\lim_{k\to\infty} \frac{B_{2k-1}\,\gamma^{2k-1}}{(2k-1)!} = \lim_{k\to\infty} \frac{\gamma^{2k-1}}{(2k-1)!} + \lim_{k\to\infty} \frac{\alpha_{2k}\,5^{2k}}{(2k)!}$$

$$= 0$$

Por otra parte:

$$\lim_{k\to\infty} \frac{B_{2k}\,\gamma^{2k}}{(2k)!} = \lim_{k\to\infty} \frac{\alpha_{2k}\,\gamma^{2k}}{(2k)!} < \lim_{k\to\infty} \frac{\alpha_{2k}\,5^{2k}}{(2k)!} = 0$$

$$\therefore \quad \lim_{k\to\infty} \frac{B_k\,\gamma^k}{k!} = 0$$

Y también, $\forall \ |h| \leq r$

$$\lim_{n\to\infty} \frac{|h|^{n+1}}{(n+1)!} \int |x|^{n+1} dF(x) = \lim_{n\to\infty} \frac{B_{n+1} |h|^{n+1}}{(n+1)!} = 0$$

Concluimos entonces:

$$\varphi(t+h) = \sum_{k=0}^{\infty} \frac{\varphi^k(\tau)\, h^k}{k!} \qquad \forall \ |h| \leq r \quad \forall \ \tau \in \mathbb{R}$$

Supongamos ahora que G es otra distribución de probabilidad sobre $\mathbb{R}$ con momentos $\alpha_1, \alpha_2, \ldots$

Por el mismo argumento de antes, si ψ es la función característica de una variable aleatoria Y con función de distribución G obtenemos:

$$\psi(\tau+h) = \sum_{k=0}^{\infty} \frac{\psi^k(\tau)\, h^k}{k!} \qquad \forall \ |h| \leq r \quad \forall \ \tau \in \mathbb{R}$$

Pero: $\varphi^k(0) = i^k \alpha_k = \psi^k(0) \quad \forall \ k$

$\therefore \ \varphi$ y ψ coinciden en el intervalo $[-r, r]$

Tomemos ahora $x \in (-2r, 2r)$, x es de la forma:

$$x = \tau_0 + h \quad \text{con} \quad |h| = r \qquad \tau_0 \in (-r, r)$$

Pero como φ y ψ coinciden en $[-r, r]$, tienen identicas derivadas de todos los ordenes en $(-r, r)$, en particular:

$$\varphi^k(\tau_0) = \psi(\tau_0) \quad \forall \ k$$

$$\therefore \ \varphi(x) = \psi(x)$$

Es decir, φ y ψ coinciden en el intervalo $(-2r, 2r)$

Ahora, si $x \in (-3r, 3r)$, x es de la forma:

$x = \tau_0 + h$ con $|h| = r$ $\tau_0 \in (-2r, 2r)$

Pero $\varphi^K(\tau_0) = \psi^K(\tau_0)$

$\therefore \varphi(x) = \psi(x)$

Es decir, φ y ψ coinciden en el intervalo $(-3r, 3r)$.

Siguiendo con este camino se demuestra que φ y ψ coinciden en todo $\mathbb{R}$.

Finalmente, aplicando el teorema 7 concluimos que $F = G$.

B I B L I O G R A F I A

I. William Feller: Introducción a la teoría de Probabilidades y sus Aplicaciones, Tomo I,II, Ed. Limusa.

2. R. G. Laha and V.K. Rohatgi: Probability Theory, Ed. John Wiley and Sons.

3. Patrick Billingsley: Probability and Measure, Ed. John Wiley and Sons.

4. Robert G. Bartle: The Elements of Real Analysis: Ed John Wiley and Sons.

5. P.A. Meyer: Probabilités et Potential, Ed. Hermann.

6. H.L. Royden, Real Analysis, Collier Macmillan International Editions.