UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

Facultad de Ciencias

Teoremas Límite

T E S I S

QUE PARA OBTENER EL TÍTULO DE :

A C T U A R I O

P R E S E N T A :

URSULA JACOBO ARTEAGA

D I R E C T O R A D E T E S I S :

DRA. GUADALUPE CARRASCO LICEA

FACULTAD DE CIENCIAS
UNAM

ENERO 2007

FACULTAD DE CIENCIAS

División de Estudios Profesionales

VNIVERSIDAD NACIONAL AVTÓNOMA DE MEXICO

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales
Facultad de Ciencias
P r e s e n t e .

Por este medio hacemos de su conocimiento que hemos revisado el trabajo escrito titulado:

"Teoremas Límite"

realizado por **Jacobo Arteaga Ursula** con número de cuenta **09638142-9,** quien opta por titularse en la opción de **Tesis** de la licenciatura en **Actuaría**. Dicho trabajo cuenta con nuestro voto aprobatorio.

Tutor(a) Propietario	Dra.	Guadalupe Carrasco Licea
Propietario	Dr.	Luis Antonio Rincón Solís
Propietario	M. en C.	Hugo Villaseñor Hernández
Suplente	M. en C.	Julio Cesar Cedillo Sánchez
Suplente	Act.	Jaime Vázquez Alamilla

A t e n t a m e n t e
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Ciudad Universitaria, D.F., a 7 de diciembre del 2006.
EL COORDINADOR DEL COMITÉ DE TITULACIÓN
DE LA LICENCIATURA EN ACTUARÍA

ACT. ROBERTO CÁNOVAS THERIOT

Señor sinodal: antes de firmar este documento, solicite al estudiante que le muestre la versión digital de su trabajo y verifique que la misma incluya todas las observaciones y correcciones que usted hizo sobre el mismo.

Quisiera agradecer antes que a nadie a mi familia que me ha dado a escoger, me ha apoyado en todas mis decisiones y me ha alentado a seguir adelante.

A Pita fuente de fortaleza, coraje y determinación, que además me dio la oportunidad de ver la probabilidad desde otro ángulo.

A todos mis compañeros que me hicieron valorar a Ciencias más que a cualquier otra escuela, no sólo por su excelente nivel educativo, si no también por siempre estar en pie de lucha.

Para Andrea, motor de mi vida.

El deber de un hombre está allí donde es más útil.

José Martí

CONTENIDO

PRESENTACIÓN

La idea de este trabajo comenzó al dar ayudantías de Probabilidad 1 y 2, en el Departamento de Matemáticas de la Facultad de Ciencias, de la UNAM. Estos cursos son obligatorios para los estudiantes de la carrera de Actuaría y optativos para estudiantes de las carreras de Física y Matemáticas. El primer curso es llevado en el tercer semestre y el segundo en el cuarto.

En el curso de Probabilidad 2, el programa incluye el tema de Convergencia, de sucesiones de variables aleatorias y teoremas límite; pero en muchas ocasiones no alcanza el semestre, para profundizar en ello. También, desafortunadamente, en gran parte de la bibliografía oficial del curso, introducen conceptos de teoría de la medida, con los que no cuentan los estudiantes de cuarto semestre de la carrera de actuaría, quienes apenas cursan Cálculo Diferencial e Integral 4. Esta situación hace complicado el estudio del tema mencionado.

En este trabajo se desarrolla una posible vía para presentar la convergencia de variables aleatorias y los principales teoremas acerca del límite de dichas sucesiones, usando resultados conocidos, y tomando en cuenta los conocimientos que tienen los estudiantes a los que está dirigido el curso de Probabilidad 2 en la Facultad de Ciencias de la UNAM.

Primero veremos algunos tipos de convergencia de cualquier sucesión de variables aleatorias, entre ellos convergencia en probabilidad; convergencia casi segura (o con probabilidad 1); convergencia en distribución y de la convergencia en L^P.

Teniendo estas bases hablaremos de la convergencia de un tipo de sucesión de variable aleatoria, (de la serie de variables aleatorias) donde se encuentran los teoremas límite muy "famosos" en la estadística *La ley de los grandes números* (*débil* y *fuerte*).

Se dice que una sucesión de variables aleatorias definidas sobre un espacio de probabilidad común obedece la *ley de los grandes números* cuando la media de una muestra $\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$ tiende a la media de las esperanzas de las variables aleatorias de la sucesión$\left(\frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right]\right)$, cuando el número n de variables aumenta.

La ley de los grandes números tiene sus inicios con Jakob Bernoulli. En el experimento que consiste repetir una prueba con la misma probabilidad de éxito un número grande de veces (experimento que más tarde recibiera el nombre de Bernoulli o Binomial), era consciente de que las frecuencias observadas se acercaban a un cálculo previo de su probabilidad al aumentar el número de repeticiones del experimento. Consciente de que en las situaciones de la vida real, la certeza absoluta (probabilidad 1) es imposible de alcanzar, Bernoulli introdujo la idea de la 'certeza moral': para que un resultado fuese moralmente cierto, debía tener una probabilidad no menor que 0.999, mientras que un resultado con probabilidad no mayor que 0.001 se consideraría 'moralmente imposible'. Fue para determinar la certeza moral de un suceso para lo que Bernoulli formuló su teorema, la Ley de los Grandes Números.

Para ilustrar este concepto, Bernoulli propuso el siguiente ejemplo: una urna con 30.000 bolas blancas y 20.000 negras, aunque el observador no lo sabe, pues lo que quiere es determinar la proporción entre bolas blancas y negras, sacando una de cada vez, anotando el resultado ("éxito" si es blanca y "fracaso" si es negra) y reintroduciéndola en la urna. Sea N el número de observaciones, X el número de éxitos y $p = \frac{r}{r+s}$ la probabilidad de éxito en cada prueba, siendo r el número de bolas blancas y s el de bolas negras. El teorema de Bernoulli afirma, en terminología moderna, que dada cualquier pequeña fracción ε (que Bernoulli siempre escribía en la forma $1/(r+s)$) y dado cualquier número entero positivo grande c, se puede hallar un número $N = N(c)$ tal que la probabilidad de que X/N difiera de p no más de ε es mayor que c veces la probabilidad de que X/N difiera de p más de ε. Con símbolos:

$P\left[\left|\frac{X}{N} - p\right| \leq \varepsilon\right] > c \cdot P\left[\left|\frac{X}{N} - p\right| > \varepsilon\right]$. O, en notación moderna:

Para todo $\varepsilon > 0$ y para toda $c \in Z^+$ existe N tal que $P\left[\left|\frac{X}{N} - p\right| > \varepsilon\right] < \frac{1}{c+1}$.

En su ejemplo, para $c = 1.000$, Bernoulli obtuvo como resultado que eran necesarias 25.550 observaciones. La intuición le decía que no hacían falta tantas y, por ello, lo intentó con otros valores de c. Desilusionado por sentir que había fracasado en su intento de cuantificar la certeza moral, Bernoulli no incluyó en su libro las aplicaciones prometidas. El que sí lo hizo fue su sobrino Niklaus (1687–

1759), que aplicó el resultado de su tío a registros de 14.000 nacimientos y llegó a la inesperada conclusión de que la frecuencia de nacimientos de niños es mayor que la de niñas, en la proporción de 18:17. Este resultado fue confirmado años después por Laplace.

Siméon Denis Poisson (1800) reintroduce el concepto de Ley de los Grandes Números y generaliza el teorema de Bernoulli. Poisson consideró una sucesión de n pruebas independientes, en cada una de las cuales el evento A puede ocurrir con probabilidad p_k, $k = 1, 2, ..., n$. Si X_n es el número de ocurrencias de A en las n observaciones, entonces para todo $\varepsilon > 0$ $P\left[\left|\frac{X_n}{n} - \frac{p_1+p_2+p_3+...+p_n}{n}\right| < \varepsilon\right] \underset{n\to\infty}{\longrightarrow} 0$. La forma generalizada de este resultado se debe a Andrei Anreyevich Markov (1880) quien plantea una condición para que se de el resultado generalizado. Más adelante Aleksandr Yakovlevich Kihintchine (1928) mostró que esa condición no era necesaria, pidiendo únicamente que la varianza común de la sucesión de variables aleatorias fuera finita.

En 1909 Émile Borel demostró que el experimento de Bernoulli llevaba a una afirmación más fuerte que la ley de los grandes números para $p = 0.5$. En 1917 Francesco Cantelli (1875-1966) extendió este hecho para p arbitrario: $P\left[\lim \frac{X_n}{n} = p\right] = 1$, afirmación que se conoce como la Ley Fuerte de los Grandes Números. Una generalización de este resultado fue dada por Andrei Kolmogorov en 1930.

La ley de los grandes números planteó el problema de estimar la suma de una sucesión de variables binomiales. Bernoulli había demostrado que la probabilidad de que el número de éxitos de un evento dado de probabilidad p en n pruebas estuviera entre A y B era $\sum_{A\leq k\leq B}\binom{n}{k}p^k(1-p)^{n-k}$. Al hacer crecer n lo más difícil era calcular $\binom{n}{k}=\frac{n!}{k!(n-k)!}$, ya que el numerador crece más rápido. Bernoulli recurrió a estimaciones poco precisas, De Moivre recurrió a una expresión asintótica y demostró que $m!\approx Be^{-m}m^{m+\frac{1}{2}}$, con B constante. Para determinar el valor de esta constante, construyó la siguiente expansión: $B=1-\frac{1}{12}+\frac{1}{360}-\frac{1}{1260}+\frac{1}{1680}-....,$ pero no quedó satisfecho así que pidió ayuda a su amigo James Stirling, quien demostró que $B=\sqrt{2\pi}$. Con este dato, De Moivre calculó una tabla para la función $m!$ con m entre 10 y 900, y enunció un resultado que, en notación moderna, dice

$$P\left[X=\frac{n}{2}+t\right]\approx P\left[X=\frac{n}{2}\right]e^{-\left(2t^2/n\right)}=\frac{2}{\sqrt{2\pi n}}e^{-\left(2t^2/n\right)}.$$

Laplace, en 1809, formuló un teorema de límites para la suma de variables aleatorias distribuidas uniformemente en un intervalo $(-h,h)$ y, considerando un número creciente de variables aleatorias discretas dio una aproximación de una distribución continua a partir de una discreta.

Chebyshev en 1887 demostró un teorema de acotamiento a la suma de variables aleatorias con esperanza 0. Para lograrlo Chebyshev ideó el Método de Momentos, de gran importancia en la estadística. Demostración que fue corregida por su discípulo Markov que en 1898 dió la siguiente versión al Teorema de Chebyshev. Sea $\{X_n\}$ una sucesión de variables aleatorias y $S_n=X_1+X_2+...+X_n$ y $\varphi_n(x)$ la distribución de X_n. Si para toda k se cumple $\int_{-\infty}^{\infty}x^k\varphi'(x)\,dx\rightarrow\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x^ke^{-x^2/2}dx$, entonces se cumple que $\forall a, b \; P\left[a < \frac{S_n}{Var(S_n)} < b\right] \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx$.

Y en 1900 A. Liapunov demostró que si las variables aleatorias tienen momento de orden 3 finito más una condición, llamada "Condición de Liapunov", entonces las funciones de distribución de S_n convergen a la distribución normal.

En 1922 Lindeberg fué mas allá que Liapunov, al pedir sólamente la existencia del momento de orden 2 más la condición: si

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{s_n^2} E\left[(X_k - \mu_k)^2 I_{[|X_k - \mu_k| > \varepsilon S_n]}\right] = 0,$$

con $s_n^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + ... + \sigma_n^2$, μ_k y σ_k^2 la media y la variaza respectivamente, de X_k, conocida como la condición de Lindeberg, la distribución de la suma de las variables aleatorias, centradas en sus esperanzas y normalizadas por la raíz cuadrada de la suma de sus varianzas converge a una distribución normal estándar.

En 1942 William Feller demostró que la condición suficiente de Lindeberg es también necesaria.

Así que por último desarrollaremos *El Teorema Central del Límite* desarrollando primero el resultado de De Moivre-Laplace y, tomando en cuenta que la condición de Lindeberg también se cumple para variables aleatorias idénticamente distribuidas, haremos la generalización a este tipo de sucesión de variables por medio de la función característica.

Llegamos hasta este resultado, no sólo porque para llegar a la generalización de Lindeberg necesitamos conocimientos de teoría de la medida, si no por su importancia tanto en probabilidad como estadística. Ya que es común suponer sucesiones de variables aleatorias independientes con la misma distribución, supuesto que se puede hacer y es de gran utilidad en la estadística. De hecho es esta independencia la que separa la Teoría de la Probabilidad de la Teoría de la Medida.

CAPÍTULO 1

CONVERGENCIA DE VARIABLES ALEATORIAS

1.1 Introducción

Consideremos un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ formado por un espacio muestral Ω, una $\sigma-$álgebra $\mathcal{A}$ de subconjuntos de Ω y una medida de probabilidad P.

Informalmente, una variable aleatoria es una función del espacio muestral a los reales que nos permite analizar un tipo particular de eventos de un experimento aleatorio. Por ejemplo, al lanzar dos dados regulares, es posible calcular la probabilidad de que la suma sea 5, de que la diferencia absoluta sea 3, de que el mayor de los números obtenidos sea 6, y de otros muchos eventos. Pero si queremos analizar sólo los eventos relacionados con la suma de los números obtenidos, definimos una variable aleatoria $X : \Omega \to R$, dada por $X(a,b) = a + b$ donde $(a,b) \in \{(x,y) \,|\, x, y \in \{1,2,3,4,5,6\}\}$. Definir la función anterior nos permite pasar del cálculo de la probabilidad de un evento particular al estudio de una ley de probabilidades, es decir, al análisis del comportamiento global de las distintas probabilidades que corresponden a los valores que puede tomar la suma.

La medida de probabilidad P tiene como dominio una $\sigma-$álgebra de subconjuntos de Ω, de manera que el dominio de esta función particular no es un subconjunto de los números reales. Construir una variable aleatoria nos permite llevar los elementos del espacio muestral al conjunto de los reales logrando así que la ley de probabilidades que se construye posteriormente, sea una función real de valores reales, en la cual es posible aplicar los conceptos del cálculo diferencial e integral. Para que esto ocurra, la variable aleatoria debe mantenter la estructura de eventos en Ω, es decir, debe mantener la estructura de la $\sigma-$álgebra construida en el espacio muestral. Por ello, una variable aleatoria no es otra cosa que una función medible.

Definición 1.1.1 *Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Una variable aleatoria (v.a.) es una función real, $X : \Omega \to \mathbb{R}$, tal que el conjunto*

$$X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B, \ con\ B = (-\infty, x]\} \in \mathcal{A}$$

para todo $B \in \mathfrak{B}(\mathbb{R})$ (para todo $x \in \mathbb{R}$).

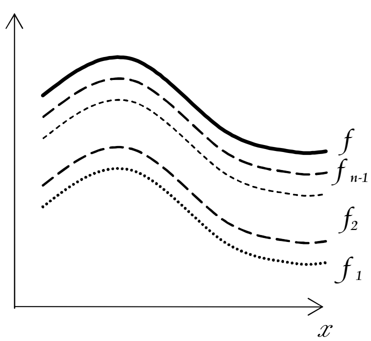
Entonces, dado que una sucesión de variables aleatorias $\{X_n\}$ no es otra cosa que una sucesión de funciones medibles, podría pensarse que el análisis de la convergencia de una sucesión de variables aleatorias se reduce al estudio de la convergencia determinista de sucesiones de funciones. Esta conjetura es falsa pues el azar no cabe dentro del concepto determinista de convergencia. Para mostrar lo anterior, empecemos por recordar las dos formas conocidas de convergencia determinista de sucesiones de funciones:

1. Convergencia puntual

Sea $\{f_n\}$ una sucesión de funciones definidas en un subconjunto A de $\mathbb{R}$ y sea f otra función definida en el mismo dominio. Se dice que f_n converge de manera puntual a f si

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \forall\, x \in A.$$

En tal caso, escribimos $\lim_{n\to\infty} f_n = f$.



En este tipo de convergencia la idea es que la sucesión de imágenes $\{f_1(x), f_2(x), f_3(x), \dots\}$ sea convergente a $f(x)$ para toda x del dominio.

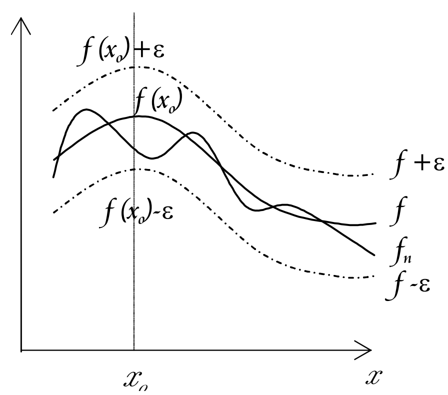
2. Convergencia uniforme

Sea $\{f_n\}$ una sucesión de funciones definidas en un subconjunto A de $\mathbb{R}$ y sea f otra función definida en el mismo dominio. Se dice que f_n converge uniformemente a f si para cada $\varepsilon > 0$ existe un número $N_\varepsilon \in \mathbb{N}$ tal que, para todo $x \in A$,

$$|f_n(x) - f(x)| < \varepsilon \qquad \forall\, n > N_\varepsilon.$$

En tal caso, escribimos $f_n \xrightarrow{u} f$.

La esencia de la convergencia uniforme de una sucesión de funciones consiste en que para cualquier $\varepsilon > 0$ se puede elegir un número N_ε dependiendo sólo de ε dado, y no depende de la elección del punto x en A, tal que para $n \geq N_\varepsilon$ la desigualdad $|f_n(x) - f(x)| < \varepsilon$ se cumplirá en todos los puntos sobre el conjunto A, es decir "las gráficas" de las funciones f_n estarán situadas en la "ε-franja" que rodea a la gráfica de la función f.

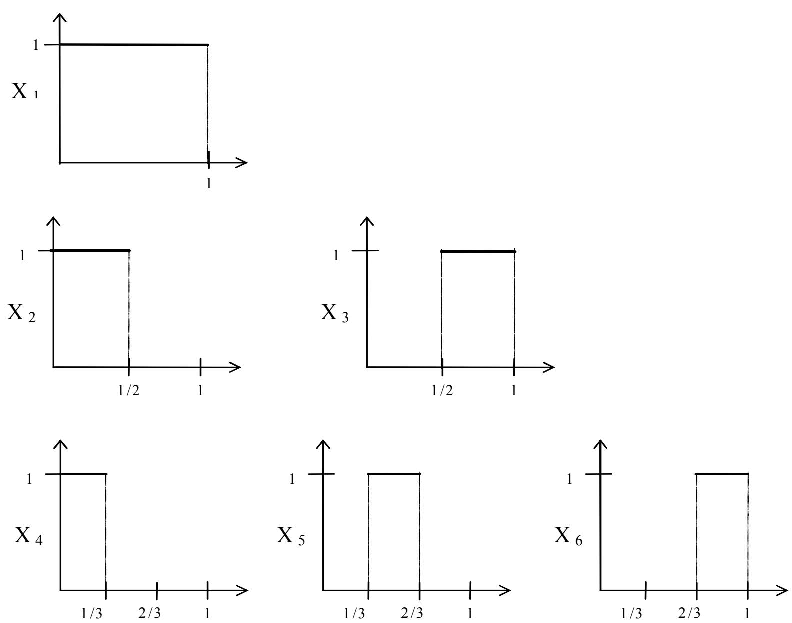


La convergencia uniforme es más fuerte que la convergencia puntual porque en esta última la rapidez de convergencia de cada sucesión de imágenes puede ser muy diferente, mientras que en la convergencia uniforme es necesario que a partir de algún elemento de la sucesión todas las imágenes esten dentro de una vecindad dada de $f(x)$. Es fácil observar que si una sucesión de funciones converge uniformemente a una función f, entonces converge puntualmente a f, pero el recíproco no se cumple.

Analicemos una sucesión de variables aleatorias

Ejemplo 1.1.2 *Considere el espacio* $((0,1], \mathfrak{B}(0,1), P)$, *donde* P *es la medida* *de Lebesgue sobre* Ω. *Sean*

$$
\begin{aligned}
X_1 &= I_{(0,1]}, \\
X_2 &= I_{(0,1/2]}, \quad X_3 = I_{(1/2,1]}, \\
X_4 &= I_{(0,1/3]}, \quad X_5 = I_{(1/3,2/3]}, \quad X_6 = I_{(2/3,1]}, \\
X_7 &= I_{(0,1/4]}, \quad X_8 = I_{(1/4,2/4]}, \quad X_9 = I_{(2/4,3/4]}, \quad X_{10} = I_{(3/4,1]}, \\
& \ldots\ldots\ldots\ldots \\
X_n &= I_{\left(\frac{j_n - 1}{k_n}, \frac{j_n}{k_n}\right]}, \qquad \text{donde } k_n \text{ se incrementa con } n,
\end{aligned}
$$



Se tiene que para cada $\omega \in (0,1]$, *éste cae en algún subintervalo de longitud* $1/k_n$, *y para cada* k_n, *sucede que* $X_n(\omega)$ *va brincando entre* 0 *y* 1 *al ir tomando* n *cada vez más grande. Por tanto,* $\{X_n\}$ *no converge determinísticamente como sucesión de funciones.*

Hay diversas formas de convergencia de sucesiones de variables aleatorias. En las siguientes secciones analizaremos las formas de convergencia que se presentan con mayor frecuencia.

Los libros que se utilizaron para recopilar este tema son [1, K.L. Chung], [2, Dudley], [5, García, Miguel], [8, Heinz Bauer], [11, Rincón, Luis]

1.2 Convergencia en probabilidad

Analicemos, nuevamente, el ejemplo anterior (ejemplo 1.1.2).

Ejemplo 1.2.1 *Se observa que las variables aleatorias X_n se van acercando a $X = 0$. Así, si $\varepsilon < 1$, $P(|X_n > \varepsilon|) = 1/k_n$ y cuando $n \to \infty$, se tiene que $k_n \to \infty$, de donde concluimos que*

$$\lim_{n\to\infty} P[|X_n| > \varepsilon] = \lim_{n\to\infty} \frac{1}{k_n} = 0.$$

Definición 1.2.2 *Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $\{X_n\}$ una sucesión de variables aleatorias y X otra variable aleatoria ambas definidas en Ω. Se dice que X_n converge en probabilidad a X, si y sólo si para cada $\varepsilon > 0$ se cumple que*

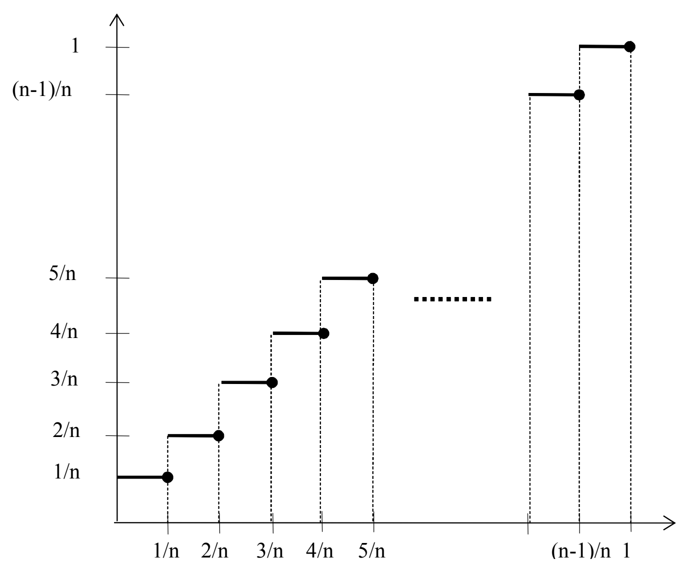
$$\lim_{n\to\infty} P[|X_n - X| > \varepsilon] = 0,$$

lo que denotaremos por $X_n \xrightarrow{P} X$

Ejemplo 1.2.3 *Se elige un punto al azar en el intervalo $(0, 1]$. Sea X la variable aleatoria que indica el punto elegido. Tomemos la sucesión de variables aleatorias*

$\{X_n\}$ *definidas por*

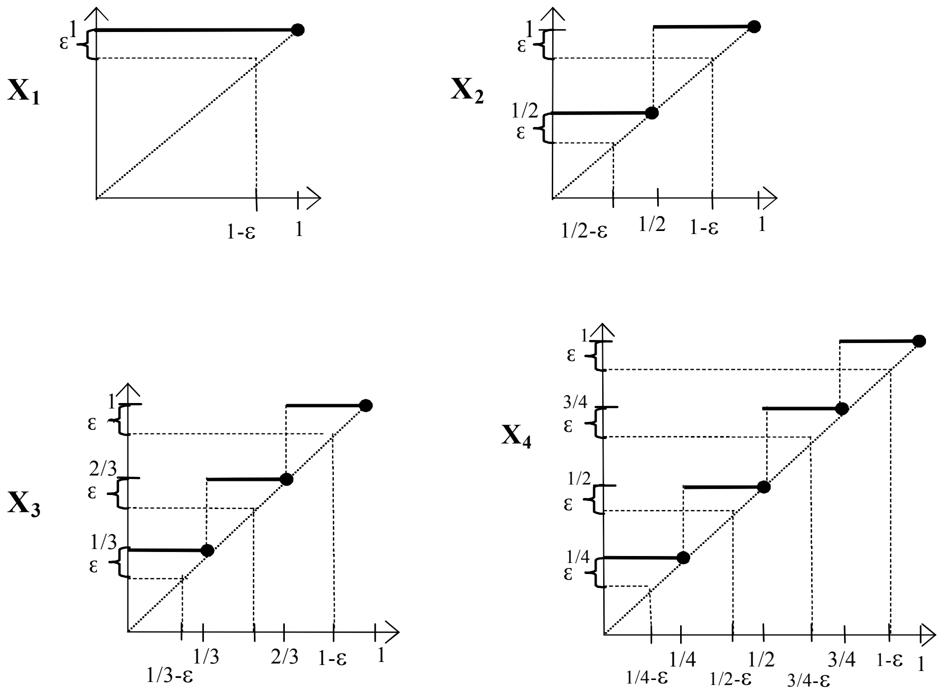
$$X_n = \sum_{k=1}^{n} \frac{k}{n} I_{\left(\frac{k-1}{n}, \frac{k}{n}\right]}.$$



X_n *es una función escalonada con saltos en* $\{1/n, 2/n, 3/n, ..., n/n\}$, *en donde la función es continua en un intervalo de tamaño* $\frac{1}{n}$, *longitud que disminuye a medida que* n *crece por lo que* X_n *se comporta cada vez más como* X. *Entonces tenemos que probar que para todo* $\epsilon > 0$

$$\lim_{n\to\infty} P\left[|X_n - X| > \varepsilon\right] = 0.$$

Para ayudarnos veamos la siguiente gráfica.



Es decir que

$$
\begin{aligned}
& \{\omega \in (0,1] \mid |X_n(\omega) - X(\omega)| > \varepsilon\} \\
= \;& \left\{ \left(0, \frac{1}{n} - \varepsilon\right), \left(\frac{1}{n}, \frac{2}{n} - \varepsilon\right), \left(\frac{2}{n}, \frac{3}{n} - \varepsilon\right), ..., \left(\frac{n-1}{n}, \frac{n}{n} - \varepsilon\right) \right\},
\end{aligned}
$$

entonces

$$P[|X_n - X| > \varepsilon] = P\left[\bigcup_{i=1}^{n} \left(\frac{i-1}{n}, \frac{i}{n} - \varepsilon\right)\right] \leq \sum_{i=1}^{n} \left(\frac{1}{n} - \varepsilon\right) = 1 - n\varepsilon.$$

Para n lo suficientemente grande sabemos que existe $N \in \mathbb{N}$ tal que para toda $n \geq N$, $\left(\frac{1}{n} \leq \frac{1}{N}\right)$ podemos elegir $\varepsilon \geq \frac{1}{N}$ tal que

$$P[|X_n - X| > \varepsilon] \leq 1 - n\varepsilon \leq 1 - n\frac{1}{N} \leq 1 - n\frac{1}{n} = 0,$$

concluyendo que

$$\lim_{n\to\infty} P\left[|X_n - X| > \varepsilon\right] = 0.$$

Ejemplo 1.2.4 *Sea* $\{Y_n\}$ *una sucesión de variables aleatorias independientes, con distribución uniforme en el intervalo* $[0, \alpha]$*, entonces*

$$X_n = \min\{Y_1, Y_2, ..., Y_n\},$$

converge en probabilidad a 0 *y*

$$Z_n = \max\{Y_1, Y_2, ...Y_n\},$$

converge en probabilidad a α*. Tomemos* ε *positivo que cumple* $\varepsilon < \alpha$ *ya que de no ser así la probabilidad sería cero*

$$\begin{aligned} P\left(|X_n| > \varepsilon\right) &= P\left[|\min\{Y_1, Y_2, ..., Y_n\}| > \varepsilon\right] \\ &= P\left[\bigcap_{i=1}^{n} (Y_i > \varepsilon)\right] = \left(1 - \frac{\varepsilon}{\alpha}\right)^n. \end{aligned}$$

Tomando el límite se obtiene el resultado. Ahora bien

$$P\left[|Z_n - \alpha| > \varepsilon\right] = P\left[Z_n < \alpha - \varepsilon\right] = P\left[\bigcap_{i=1}^{n} (Y_i < \alpha - \varepsilon)\right] = \left(1 - \frac{\varepsilon}{\alpha}\right)^n.$$

Que coincide con el resultado anterior así también al tomar el límite.

Proposición 1.2.5 *Sean* $\{X_n\}$ *y* $\{Y_n\}$ *dos sucesiones de variables aleatorias y sea* c *una constante:*

1. *Si* $X_n \xrightarrow{P} X$ *y* $X_n \xrightarrow{P} Y$*, entonces* $P(X = Y) = 1$.

2. *Si* $X_n \xrightarrow{P} X$ *, entonces* $cX_n \xrightarrow{P} cX$.

3. *Si* $X_n \xrightarrow{P} X$ *y* $Y_n \xrightarrow{P} Y$, *entonces* $X_n + Y_n \xrightarrow{P} X + Y$.

4. *Si* $X_n \xrightarrow{P} X$ *, entonces* $X_n^2 \xrightarrow{P} X^2$.

5. *Si* $X_n \xrightarrow{P} X$ *y* $Y_n \xrightarrow{P} Y$, *entonces* $X_nY_n \xrightarrow{P} XY$.

Demostración.

1.
$$|X - Y| \leq |X_n - X| + |X_n - Y|\,,$$

entonces

$$[|X - Y| > \varepsilon] \subset [(|X_n - X| + |Y_n - Y|) > \varepsilon],$$

y para cualquier $\varepsilon > 0$ *se cumple*

$$[(|X_n - X| + |X_n - Y|) > \varepsilon] \subset [|X_n - X| > \frac{\varepsilon}{2}] \cup [|X_n - Y| > \frac{\varepsilon}{2}],$$

entonces

$$P[|X - Y| > \varepsilon] \leq P[|X_n - X| > \frac{\varepsilon}{2}] + P[|X_n - Y| > \frac{\varepsilon}{2}].$$

Tomando el límite cuando $n \to \infty$ *se tiene que*

$$P[|X - Y| > \varepsilon] = 0 \qquad \textit{para cualquier } \varepsilon > 0.$$

2.
$$\lim_{n\to\infty} P[|cX_n - cX| > \varepsilon] = \lim_{n\to\infty} P\left[|X_n - X| > \frac{\varepsilon}{|c|}\right] = 0.$$

3.

$$|(X_n + Y_n) - (X + Y)| \leq |X_n - X| + |Y_n - Y|,$$

entonces

$$[|(X_n + Y_n) - (X + Y)| > \varepsilon] \subset [|X_n - X| > \frac{\varepsilon}{2}] \cup [|Y_n - Y| > \frac{\varepsilon}{2}],$$

implica

$$P[|(X_n + Y_n) - (X + Y)| > \varepsilon] \leq P[|X_n - X| > \frac{\varepsilon}{2}] + P[|Y_n - Y| > \frac{\varepsilon}{2}].$$

tomando el límite se tiene

$$\lim_{n\to\infty} P[|(X_n + Y_n) - (X + Y)| > \varepsilon] = 0.$$

4. *Sabemos que* $\sum_{k=0}^{\infty} P[k \leq |X| < k+1] = 1$, *entonces dada* $\delta > 0$ *existe* M *tal que*

$$P[|X| > M] \leq P[|X| \geq M] = \sum_{k=M}^{\infty} P[k \leq |X| < k+1] < \frac{\delta}{2},$$

también como $X_n \xrightarrow{P} X$, *entonces podemos afirmar que existe una* N *tal que para toda* $n \geq N$, *se cumple que*

$$P[|X_n - X| > M] < \frac{\delta}{2},$$

y como

$$|X_n + X| \leq |X_n - X| + 2|X|,$$

entonces

$$[|X_n + X| > 4M] \subseteq [|X_n - X| > 2M] \cup [|X| > M],$$

entonces para $n \geq N$

$$P\left[|X_n + X| > 4M\right] \leq P\left[|X_n - X| > 2M\right] + P\left[|X| > M\right] < \delta,$$

Por lo que para $\varepsilon > 0$

$$\begin{aligned}
P\left[\left|X_n^2 - X^2\right| > \varepsilon\right] &= P\left[\left|\left(X_n - X\right)\left(X_n + X\right)\right| > \varepsilon\right] \\
&= P\left[\left|\left(X_n - X\right)\left(X_n + X\right)\right| > \varepsilon, |X_n + X| \leq 4M\right] \\
&\quad + P\left[\left|\left(X_n - X\right)\left(X_n + X\right)\right| > \varepsilon, |X_n + X| > 4M\right] \\
&\leq P\left[\left|\left(X_n - X\right)\left(X_n + X\right)\right| > \varepsilon, 0 < |X_n + X| \leq 4M\right] \\
&\quad + P\left[\left|\left(X_n - X\right)\left(X_n + X\right)\right| > \varepsilon, |X_n + X| > 4M\right] \\
&= P\left[|X_n - X| > \frac{\varepsilon}{4M}, 0 < |X_n + X| \leq 4M\right] \\
&\quad + P\left[\left|\left(X_n - X\right)\left(X_n + X\right)\right| > \varepsilon, |X_n + X| > 4M\right] \\
&\leq P\left[|X_n - X| > \frac{\varepsilon}{4M}\right] + P\left[|X_n + X| > 4M\right] \\
&< P\left[|X_n - X| > \frac{\varepsilon}{4M}\right] + \delta.
\end{aligned}$$

Tomando límite se tiene que

$$\limsup_{n \to \infty} P\left[\left|X_n^2 - X^2\right| > \varepsilon\right] \leq \delta$$

y como esto se cumple para toda $\delta > 0$ *entonces*

$$\lim_{n \to \infty} P\left[\left|X_n^2 - X^2\right| > \varepsilon\right] = 0.$$

5. *Como* $X_n Y_n = \frac{1}{4}\left[\left(X_n + Y_n\right)^2 - \left(X_n - Y_n\right)^2\right]$, *el resultado se sigue de los resultados anteriores.*

■

1.3 Convergencia casi segura

Entonces ya vimos que si una sucesión de funciones $\{X_n\}$ no converge determinísticamente, puede ser que converja en probabilidad, es decir

$$\lim_{n\to\infty} P\left[|X_n - X| > \varepsilon\right] = 0.$$

Y de ello surge la idea de ver qué pasa con $P\left(\lim_{n\to\infty} X_n = X\right)$, nos preguntamos por la probabilidad de una convergencia puntual. Lo cual de inmediato nos lleva a pensar que existe un conjunto, en donde $\lim_{n\to\infty} X_n \neq X$, si este conjunto es de probabilidad cero, entonces $P\left(\lim_{n\to\infty} X_n = X\right) = 1$, que es la definición de convergencia casi segura, convergencia más fuerte que la definida anteriormente.

Definición 1.3.1 *Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $\{X_n\}$ una sucesión de variables aleatorias, se dice que X_n converge casi seguramente a X, si y sólo si,*

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1,$$

lo que denotaremos como $X_n \overset{cs}{\to} X$.

Ejemplo 1.3.2 *Se elige un punto al azar en el intervalo $(0,1]$. Sea X la variable aleatoria que indica el punto elegido. Tomemos la sucesión de variables aleatorias $\{X_n\}$ definidas por*

$$X_n = \sum_{k=1}^{n} \frac{k}{n} I_{\left(\frac{k-1}{n}, \frac{k}{n}\right]},$$

(que corresponde al primer ejemplo de la anterior sección). Definamos

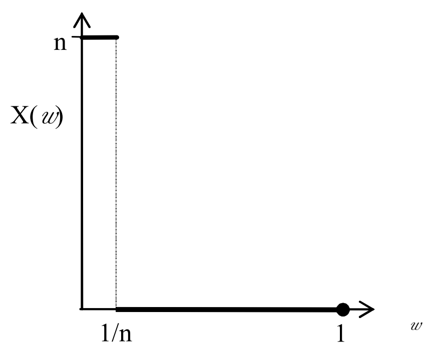
$$A = \left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} = (0,1] = \Omega.$$

Entonces el conjunto N donde X_n no converge a X es vacío

Ejemplo 1.3.3 *Considere el espacio de probabilidad* $([0,1], \mathfrak{B}[0,1], P)$, *con P la medida de probabilidad uniforme (es decir $P[(a,b)] = b-a$). Demostraremos que*

$$X_n = nI_{\left[0,\frac{1}{n}\right]} \xrightarrow{cs} 0$$

en otras palabras demostraremos que $P[\lim_{n\to\infty} X_n = 0] = 1$. Vemos que



el punto $\omega = 0$ es el único punto de Ω para el que $X_n(0) \overset{cs}{\nrightarrow} 0$, es decir que

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X\right\} = (0,1],$$

y como para $N = \{0\}$ $P(N) = 0$, entonces

$$P\left[\lim_{n\to\infty} X_n = 0\right] = 1.$$

Proposición 1.3.4 *Sean $\{X_n\}$ y $\{Y_n\}$ dos sucesiones de variables aleatorias y sea c una constante:*

1. *Si $X_n \xrightarrow{cs} X$ y $X_n \xrightarrow{cs} Y$, entonces $P(X = Y) = 1$.*

2. *Si $X_n \xrightarrow{cs} X$, entonces $cX_n \xrightarrow{cs} cX$.*

3. *Si* $X_n \overset{cs}{\to} X$ *y* $Y_n \overset{cs}{\to} Y$, *entonces* $X_n + Y_n \overset{cs}{\to} X + Y$.

4. *Si* $X_n \overset{cs}{\to} X$ *y* $Y_n \overset{cs}{\to} Y$, *entonces* $X_n Y_n \overset{cs}{\to} XY$.

5. *Si* $X_n \overset{cs}{\to} Y_j$ *y* $Y_j \overset{cs}{\to} X$, *entonces* $X_n \overset{cs}{\to} X$.

Demostración.

1. *Tenemos que se cumple que*

$$P\left[Y = \lim_{n\to\infty} X_n = X\right] = 1.$$

2.

$$\begin{aligned} P\left[\lim_{n\to\infty} X_n = X\right] &= P\left[c\lim_{n\to\infty} X_n = cX\right] \\ &= P\left[\lim_{n\to\infty} cX_n = cX\right] = 1. \end{aligned}$$

3. *Llamemos*

$$A = \left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}$$

y

$$B = \left\{\omega \in \Omega \mid \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\right\},$$

por hipótesis existen conjuntos N_1 *y* N_2 *tales que*

$$A \cup N_1 = \Omega \quad y \quad B \cup N_2 = \Omega,$$

donde A *y* N_1 *así como* B *y* N_2 *son mutuamente excluyentes y además*

$$P(N_1) = P(N_2) = 0,$$

entonces,

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} \left(X_n(\omega) + Y_n(\omega)\right) = X\left(\omega\right) + Y\left(\omega\right)\right\} = A \cap B,$$

así que

$$P\left[\lim_{n\to\infty} \left(X_n + Y_n\right) = X + Y\right] = 1 - P\left(N_1 \cup N_2\right) = 1.$$

4. *Usando los mismos conjuntos que en el caso anterior*

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} \left(X_n(\omega) Y_n(\omega)\right) = X\left(\omega\right) Y\left(\omega\right)\right\} = A \cap B,$$

ya que ambos convergen, por lo que

$$P\left[\lim_{n\to\infty} \left(X_n Y_n\right) = XY\right] = 1 - P\left(N_1 \cup N_2\right) = 1.$$

5. *Llamemos*

$$\Omega_0 = \left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = Y_j\left(\omega\right)\right\},$$

y

$$\Omega_1 = \left\{\omega \in \Omega \mid \lim_{j\to\infty} Y_j(\omega) = X\left(\omega\right)\right\},$$

por hipótesis conjuntos de probabilidad uno. Entonces existen N_1, $N_2 \in \mathbb{N}$ *tales que*

$$\left|X_n\left(\omega\right) - Y_j\left(\omega\right)\right| < \frac{\varepsilon}{2},\ \text{para } n \geq N_1 \text{ y } \omega \in \Omega_0,$$

$$\left|Y_j\left(\omega\right) - X\left(\omega\right)\right| < \frac{\varepsilon}{2},\ \text{para } j \geq N_2 \text{ y } \omega \in \Omega_1,$$

entonces,

$$\left|X_n\left(\omega\right) - X\left(\omega\right)\right| \leq \left|X_n\left(\omega\right) - Y_j\left(\omega\right)\right| + \left|Y_j\left(\omega\right) - X\left(\omega\right)\right| < \varepsilon,$$

para $n \geq \max\{N_1, N_2\}$ *y* $\omega \in \Omega_0 \cap \Omega_1 = \Omega_2$. *Siendo*

$$\Omega_2 = \left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\},$$

tenemos que

$$P(\Omega_2^c) = P(\Omega_0^c \cup \Omega_1^c) \leq P(\Omega_0^c) + P(\Omega_1^c) = 0,$$

por lo que $P(\Omega_2) = 1$, *es decir,* $X_n \overset{cs}{\to} X$.

■

1.3.1 Lema de Borel-Cantelli

Teorema 1.3.5 *Sea* $P : \mathcal{A} \to [0,1]$ *una medida de probabilidad. Entonces*

1. *Si* $E_1, E_2, E_{3,} \dots \in \mathcal{A}$, $E_1 \subset E_2 \subset E_3 \subset \dots$, *entonces*

$$P\left(\bigcup_{j=1}^{\infty} E_j\right) = \lim_{n\to\infty} P(E_n).$$

2. *Si* $E_1, E_2, E_{3,} \dots \in \mathcal{A}$, $E_1 \supset E_2 \supset E_3 \supset \dots$, *entonces*

$$P\left(\bigcap_{j=1}^{\infty} E_j\right) = \lim_{n\to\infty} P(E_n).$$

Demostración.

1. *Sean* $A_1 = E_1$ *y* $A_k = E_k - E_{k-1}$ *entonces los eventos* A_k *son ajenos entre sí y además* $\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} E_j$. *Así que por la* $\sigma-$*aditividad y tomando*

$E_0 = \emptyset$, se tiene que

$$\begin{aligned} P\left(\bigcup_{j=1}^{\infty} A_j\right) &= \sum_{j=1}^{\infty} P\left(A_j\right) \\ &= \sum_{j=1}^{\infty} P\left(E_j - E_{j-1}\right) \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} P\left(E_j - E_{j-1}\right) \\ &= \lim_{n\to\infty} P\left(E_n\right). \end{aligned}$$

2. *Si $E_1 \supset E_2 \supset E_3 \supset \ldots$, entonces $E_1^c \subset E_2^c \subset E_3^c \subset \ldots$ y por el inciso anterior*

$$\begin{aligned} P\left(\bigcap_{j=1}^{\infty} E_j\right) &= P\left(\bigcup_{j=1}^{\infty} E_j^c\right)^c = 1 - P\left(\bigcup_{j=1}^{\infty} E_j^c\right) \\ &= 1 - \lim_{n\to\infty} P\left(E_n^c\right) = \lim_{n\to\infty}\left[1 - P\left(E_n^c\right)\right] \\ &= \lim_{n\to\infty} P\left(E_n\right). \end{aligned}$$

■

Definición 1.3.6 *Sea E_n cualquier sucesión de eventos de Ω , definimos*

$$\lim_n \ \sup E_n = \bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} E_n \qquad\qquad \lim_n \ \inf E_n = \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty} E_n$$

Notemos que $\lim_n \inf E_n = \left(\lim_n \sup E_n^c\right)^c$

Usando el teorema anterior (1.3.5) y el hecho de que $C_k = \bigcap\limits_{n=k}^{\infty} B_n$ es una sucesión decreciente, entonces

$$P\left(\lim_n \ \sup E_n\right) = \lim_{m\to\infty} P\left(\bigcup_{n=m}^{\infty} E_n\right)$$

y para $D_k = \bigcup_{n=k}^{\infty} F_n$, una sucesión creciente, entonces

$$P\left(\lim_n \ \inf E_n\right) = \lim_{m\to\infty} P\left(\bigcap_{n=m}^{\infty} E_n\right)$$

Axioma 1.3.7 *Sea $\{E_n\}$ cualquier sucesión de eventos de Ω entonces*

$$\limsup_n E_n = \{\omega \in \Omega \mid \omega \in E_n \text{ para una infinidad de valores de } n\}$$

Demostración. *Sea $F_m = \bigcup_{n=m}^{\infty} E_n$, entonces*

$$\bigcap_{m=1}^{\infty} F_m = \limsup_n E_n.$$

Observemos que si $\omega \in \bigcap_{m=1}^{\infty} F_m$ entonces $\omega \in F_m$ para toda m. Donde $\omega \in \bigcup_{n=m}^{\infty} E_n$ si ω pertenece sólo a un subconjunto finito de $\{E_n\}$ entonces podemos encontrar una m^1 tal que, para $n \geq m^1$, $\omega \notin \bigcup_{n=m^1}^{\infty} E_n$, lo cual implicaría que $\omega \notin F_{m^1}$ lo que es una contradicción. Por eso es que ω pertenece a un conjunto infinito de $\{E_n\}$. ∎

Teorema 1.3.8 (Lema de Borel-Cantelli) *Sea $\{E_n\}$ una sucesión de eventos arbitrarios tales que*

$$\sum_{n=1}^{\infty} P(E_n) < \infty, \text{ entonces, } P(\lim{}_n \sup E_n) = 0.$$

Demostración. $P(\lim_n \sup E_n) = \lim_{m\to\infty} P\left(\bigcup_{n=m}^{\infty} E_n\right)$

$$\leq \lim_{m\to\infty} \sum_{n=m}^{\infty} P(E_n) = 0.$$ ∎

Proposición 1.3.9 *Sea $\{X_n\}$ una sucesión de variables aleatorias entonces*

$$X_n \xrightarrow{cs} X,$$

si y sólo si, para cualquier $\varepsilon > 0$

$$P\left[\{\omega \in \Omega \mid (|X_n(\omega) - X(\omega)| > \varepsilon) \text{ para una infinidad de valores de } n\}\right] = 0.$$

Demostración. *Primero supongamos que se cumple que $X_n \xrightarrow{cs} X$ es decir que*

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} = \Omega_0,$$

es un conjunto de probabilidad uno, y cumple que para todo $\omega \in \Omega_0$ y $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que

$$|X_n - X| < \varepsilon \text{ para } n \geq N.$$

Ahora bien si

$$\omega \in A(\varepsilon) = \{\omega \in \Omega \mid (|X_n(\omega) - X(\omega)| > \varepsilon) \text{ para una infinidad de valores}\},$$

entonces $\omega \in \Omega_0^c$, es decir que $A(\varepsilon) \subseteq \Omega_0^c$ por lo que

$$P[A(\varepsilon)] \leq P[\Omega_0^c] = 0,$$

Ahora supongamos que $P[A(\varepsilon)] = 0$, es decir, por (1.3.7) que

$$A(\varepsilon) = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{\omega \in \Omega \mid (|X_n(\omega) - X(\omega)| > \varepsilon)\},$$

es un conjunto de probabilidad cero para cualquier $\varepsilon > 0$, en particular para $\varepsilon = \frac{1}{n}$ para cada $n \in \mathbb{N}$, sea

$$B_m = A\left(\frac{1}{m}\right),$$

entonces $\{B_m^c\}$ *es una sucesión monótona decreciente por lo que*

$$P\left(\bigcap_{m=1}^{\infty} B_m^c\right) = \lim_{m\to\infty} P\left(\bigcap_{m=1}^{\infty} B_m^c\right) = 1,$$

como $B_m^c = \left\{\omega \in \Omega \mid \exists N \text{ tal que } \left(|X_n(\omega) - X(\omega)| \leq \frac{1}{m}\right) \ \forall n \geq N\right\}$, *lo que implica que para cada* $\omega \in \bigcap_{m=1}^{\infty} B_m^c$, *y para cualquier* $m \in \mathbb{N}$ *existe* N *tal que* $|X_n(\omega) - X(\omega)| \leq \frac{1}{m}$, *para cualquier* $n \geq N$, *es decir, que*

$$\lim_{n\to\infty} X_n(\omega) = X(\omega),$$

o bien, como

$$\bigcap_{m=1}^{\infty} B_m^c \subset \Omega_0,$$

entonces

$$1 = P\left[\bigcap_{m=1}^{\infty} B_m^c\right] \leq P[\Omega_0]. \quad \blacksquare$$

Corolario 1.3.10 *Sea* $\{X_n\}$ *una sucesión de variables aleatorias si*

$$\sum_{n=1}^{\infty} P(|X_n(\omega) - X(\omega)| > \varepsilon) < \infty,$$

entonces $X_n \xrightarrow{cs} X$.

Demostración. *Por el* (1.3.8) *si* $\sum_{n=1}^{\infty} P(|X_n(\omega) - X(\omega)| > \varepsilon) < \infty$, *entonces*

$$P[\lim\ _n \sup(|X_n(\omega) - X(\omega)| > \varepsilon)] = 0.$$

Así que por (1.3.9), $X_n \xrightarrow{cs} X$. $\blacksquare$

1.4 Convergencia en distribución

Un tipo de convergencia que parece ser muy natural en probabilidad es el de convergencia en distribución, ya que continuamente nos encontramos con problemas que tienen una solución, posiblemente sencilla, pero no así su cálculo que se complica a cada paso. Por ejemplo nos planteamos encontrar la probabilidad de obtener al menos 65 dobles seises en 100 lanzamientos de dos dados. El resultado numérico es, en este caso, el que se va haciendo complicado conforme el número de repeticiones del experimento. Lo que nos lleva a buscar una variable aleatoria con la que podamos aproximar este resultado.

Definición 1.4.1 *Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $\{X_n\}$ una sucesión de variables aleatorias, se dice que X_n converge en distribución a la variable aleatoria X, si y sólo si*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{para todo } x \in \mathbb{R} \text{ donde } F_X \text{ sea continua,}$$

lo que denotaremos como $X_n \xrightarrow{D} X$.

Ejemplo 1.4.2 *Sea $\{X_n\}$ una sucesión de variables donde X_n se distribuye como una normal de parámetros 0 y $\frac{\sigma^2}{n}$ y sea X la constante cero. Entonces*

$$P[X_n \le x] = P\left[\frac{X_n}{\sqrt{\sigma^2/n}} \le \frac{x}{\sqrt{\sigma^2/n}}\right] = P\left[Z_n \le \frac{x}{\sigma/\sqrt{n}}\right].$$

Para ver que pasa cuando $n \to \infty$ *con* F_{X_n} *veamos que pasa con* $\frac{x}{\sigma/\sqrt{n}}$

$$\lim_{n\to\infty} \frac{x}{\sigma/\sqrt{n}} = \begin{cases} -\infty & si\ x < 0, \\ 0 & si\ x = 0, \\ \infty & si\ x > 0, \end{cases}$$

entonces,

$$\lim_{n\to\infty} F_{X_n}(x) = \begin{cases} 0 & si\ x < 0, \\ 1/2 & si\ x = 0, \\ 1 & si\ x > 0, \end{cases}$$

y

$$F_X(x) = \begin{cases} 0 & si\ x < 0, \\ 1 & si\ x \geq 0. \end{cases}$$

Por lo que se concluye que $X_n \xrightarrow{D} X.$

Ejemplo 1.4.3 *Sea* X *una v.a. con distribución uniforme en el intervalo* $(0, 1)$ *y* $\{X_n\}$ *una sucesión de v.a's discretas cada una de las cuales tiene distribución uniforme en el conjunto* $\left\{\frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}, \frac{n}{n}\right\}$ *veremos que*

$$X_n \xrightarrow{D} X.$$

Que X *tenga distribución uniforme en el intervalo* $(0, 1)$, *quiere decir que*

$$F_X(x) = \begin{cases} 0 & si & x \leq 0, \\ x & si & 0 < x < 1, \\ 1 & si & 1 \leq x. \end{cases}$$

Y que X_n *se distribuya como uniforme en el conjunto* $\left\{\frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}, \frac{n}{n}\right\}$ *implica que*

$$F_{X_n}(x) = \begin{cases} 0 & si \quad x < 1/n, \\ 1/n & si \quad 1/n \leq x < 2/n, \\ 2/n & si \quad 2/n \leq x < 3/n, \\ \vdots & \\ i/n & si \quad i/n \leq x < (i+1)/n, \\ \vdots & \\ (n-1)/n & si \quad (n-1)/n \leq x < 1, \\ 1 & si \quad 1 \leq x, \end{cases}$$

reescribiendo a F_{X_n} *como*

$$F_{X_n}(x) = \begin{cases} 0 & si\ x < 1/n, \\ x - (x - 1/n) & si\ 1/n \leq x < 2/n, \\ x - (x - 2/n) & si\ 2/n \leq x < 3/n, \\ \vdots & \\ x - (x - i/n) & si\ i/n \leq x < (i+1)/n, \\ \vdots & \\ x - (x - (n-1)/n) & si\ (n-1)/n \leq x < 1, \\ 1 & si\ 1 \leq x. \end{cases}$$

Entonces para $i/n \leq x < (i+1)/n$, $\lim_{n\to\infty}(x - i/n) = 0$, *por lo que se obtiene* *que*

$$\lim_{n\to\infty} F_{X_n}(x) = \begin{cases} 0 & si \quad x < 0, \\ x & si \quad 0 \leq x \quad < 1, \\ 1 & si \quad 1 \leq x. \end{cases}$$

Que corresponde a la función de distribución continua, uniforme en el $[0,1)$. *Por lo que* $X_n \xrightarrow{D} X$.

El siguente Teorema (Teorema de Poisson) también surge de la convergencia de distribución, una variable aleatoria con distribución binomial converge en distribución a una variable aleatoria con distribución poisson. Muchos autores lo introducen hasta la parte de convergencia de series ya que la variable aleatoria con distribución binomial no es más que una suma de variables aleatorias, independientes e idénticamente distribuidas como bernoullis.

Pero la demostración es muy sencilla, ya que se considera que np siempre es constante!!, es decir que para n creciente se tiene una p decreciente (la probabilidad de éxito de un evento bernoulli depende del número de repeticiones del experimento).

Teorema 1.4.4 (de Poisson) *Sea* $\{X_n\}$ *una sucesión de variables tales que* X_n *se distribuye como una binomial de parámetros* n *y* $p \in (0,1)$ *y sea* X *una variable aleatoria con distribución poisson de parámetro* $\lambda = np$ *constante. Entonces*

$$X_n \xrightarrow{D} X.$$

Demostración.

$$\begin{aligned}
P\left[X_n \leq x\right] &= \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{x} \frac{n(n-1)(n-2).....(n-k+2)(n-k+1)}{k!} \frac{p^k n^k}{n^k} \left(1-\frac{np}{n}\right)^{n-k} \\
&= \sum_{k=0}^{x} \frac{(np)^k}{k!} \left(1-\frac{np}{n}\right)^{n} \left(1-\frac{1}{n}\right) ... \left(1-\frac{k-1}{n}\right) \left(1-\frac{np}{n}\right)^{-k} \\
&= \sum_{k=0}^{x} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^{n} \left(1-\frac{1}{n}\right) ... \left(1-\frac{k-1}{n}\right) \left(1-\frac{\lambda}{n}\right)^{-k}.
\end{aligned}$$

Como

$$\lim_{n\to\infty} \left(1-\frac{i}{n}\right) = 1 \text{ para cada } i < n,$$

tomemos el limite cuando $n \to \infty$ *de* $F_{X_n}(x)$

$$\lim_{n\to\infty} F_{X_n}(x) = \sum_{k=0}^{x} \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n = \sum_{k=0}^{x} \frac{\lambda^k}{k!} e^{-\lambda} = F_X(x). \quad \blacksquare$$

Debido a la importancia de este resultado lo ejemplificaremos.

Ejemplo 1.4.5 *En una fábrica se empacan cajas de cereal con pasas mezclando primero, en un gran recipiente, el cereal con las pasas y de tal manera que cada caja cantenga, en promedio, 12 pasas. Estime la probabilidad de que cada caja de cereal, seleccionada al azar, contenga menos de 7 pasas.*

Definamos X *como el número de pasas que contiene una caja,* $x = \{0, 1, 2, ...\}$. *Debido a que* n *no está acotada y por lo mismo no tenemos la probabilidad de encontrar una pasa en la caja (la probabilidad de éxito), pero como si contamos con el número esperado de pasas que hay por caja* $np = \lambda = 12$ *entonces podemos* *utilizar una aproximación, que por la información obtenida puede ser por medio de una poisson. Es decir, para*

$$Y \sim Poisson\,(12)\,,$$

$$\begin{aligned} P\,(X<7) &= \sum_{x=0}^{6} P\,(X=x) \approx \sum_{x=0}^{6} P\,(Y=x) \\ &= \sum_{x=0}^{6} \frac{12^x}{x!} e^{-12} = e^{-12} \sum_{x=0}^{6} \frac{12^x}{x!} \approx 0.42. \end{aligned}$$

Proposición 1.4.6 *Sean $\{X_n\}$ y $\{Y_n\}$ dos sucesiones de variables aleatorias y sea c una constante:*

1. *Si $X_n \xrightarrow{D} X$ y $X_n \xrightarrow{D} Y$, entonces $F_X = F_Y$.*

2. *Si $X_n \xrightarrow{D} X$, entonces $cX_n \xrightarrow{D} cX$.*

3. *Si $X_n \xrightarrow{D} X$, entonces $X_n + c \xrightarrow{D} X + c$.*

Demostración.

1. *De la definición de convergencia en distribución, se sigue inmediatamente que $F_X\,(z) = F_Y\,(z)$ para cualquier real z en el que F_X y F_Y sean continuas. Pero como el conjunto de discontinuidades de F_X y de F_Y es numerable, entonces el conjunto de reales en los que tanto F_X como F_Y son continuas, es denso en $\mathbb{R}$. La conclusión se sigue entonces de la continuidad por la derecha de F_X y de F_Y.*

2. *Si* $X_n \xrightarrow{D} X$ *, entonces*

$$\begin{aligned}\lim_{n\to\infty} F_{X_n}(x) &= F_X(x)\\ \lim_{n\to\infty} P(X_n \leq x) &= P(X \leq x),\end{aligned}$$

si $c=0$ *es inmediato, veamos para* $c \neq 0$

$$\begin{aligned}P(cX_n \leq x) &= \begin{cases} P\left(X_n \leq \frac{x}{c}\right) & si\ c>0,\\ P\left(X_n \geq \frac{x}{c}\right) & si\ c<0,\end{cases}\\ &= \begin{cases} P\left(X_n \leq \frac{x}{c}\right) & si\ c>0,\\ 1-P\left(X_n < \frac{x}{c}\right) & si\ c<0,\end{cases}\end{aligned}$$

tomando el límite cuando $n \to \infty$ *tenemos*

$$\begin{aligned}\lim_{n\to\infty} F_{cX_n}(x) &= \begin{cases} P\left(X \leq \frac{x}{c}\right) & si\ c>0,\\ 1-P\left(X < \frac{x}{c}\right) & si\ c<0,\end{cases}\\ &= \begin{cases} P(cX_n \leq x) & si\ c>0,\\ 1-P(cX_n > x) & si\ c<0,\end{cases}\\ &= F_{cX}(x).\end{aligned}$$

3. *Para el caso en que* $c=0$ *se trata de la misma función por lo que sólo haremos el caso en que* $c \neq 0$

$$P(X_n + c \leq x) = P(X_n \leq x-c),$$

tomando el límite

$$\begin{aligned}
\lim_{n\to\infty} F_{X_n+c}(x) &= P(X \leq x - c) \\
&= P(X + c \leq x) \\
&= F_{X+c}(x).
\end{aligned}$$

■

1.5 Convergencia en L^p

Definición 1.5.1 *Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $\{X_n\}$ una sucesión de variables aleatorias,con esperanza finita, se dice que X_n converge en L^p a la variable aleatoria X, si y sólo si*

$$\lim_{n\to\infty} E[|X_n - X|^p] = 0,$$

lo que denotaremos como $X_n \xrightarrow{L^p} X$. En donde si $p = 1$ se le llamará convergencia en media y si $p = 2$ convergancia en media cuadrática.

Ejemplo 1.5.2 *Considere el espacio $((0,1], \mathfrak{B}(0,1), P)$,donde P es la medida* *de Lebesgue sobre* $(0,1]$*. Sean*

$$
\begin{aligned}
X_1 &= I_{(0,1]}, \\
X_2 &= I_{(0,1/2]}, \quad X_3 = I_{(1/2,1]}, \\
X_4 &= I_{(0,1/3],} \quad X_5 = I_{(1/3,2/3]}, \quad X_6 = I_{(2/3,1]}, \\
X_7 &= I_{(0,1/4]}, \quad X_8 = I_{(1/4,2/4],} \quad X_9 = I_{(2/4,3/4],} \quad X_{10} = I_{(3/4,1],} \\
&\ldots\ldots\ldots\ldots \\
X_n &= I_{\left(\frac{j_n-1}{k_n},\frac{j_n}{k_n}\right]}, \qquad\qquad \textit{donde } k_n \textit{ se incrementa con } n,
\end{aligned}
$$

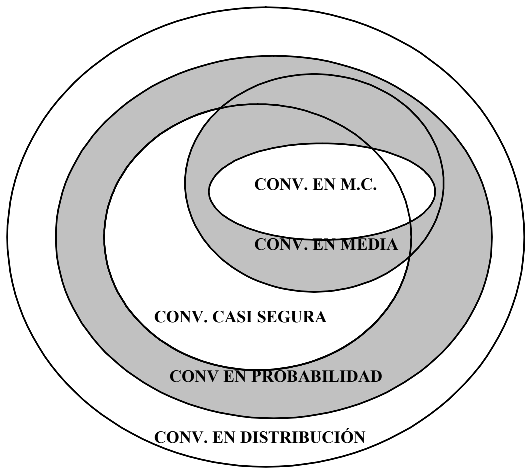
y sea X *la constante cero, veamos que* $X_n \xrightarrow{L^p} X$

$$
\begin{aligned}
\lim_{n\to\infty} E\left[|X_n|^p\right] &= \lim_{n\to\infty} \int_{(j_n-1)/k_n}^{j_n/k_n} 1^p \; dx \\
&= \lim_{n\to\infty} \frac{1}{k_n} = 0,
\end{aligned}
$$

por lo que concluimos que en efecto $X_n \xrightarrow{L^p} X$.

1.6 Relación entre modos de convergencia de variables aleatorias

En realidad lo que vamos a ver en esta sección se muestra en la gráfica



Veamos unos ejemplos que muestren la no relación y despúes demostraremos las relaciones que existen.

Ejemplo 1.6.1 *Sea X una variable aleatoria con distribución normal estándar y, para cada $n \in \mathbb{N}$, definamos*

$$X_n = \begin{cases} X & \text{si } n \text{ es impar,} \\ -X & \text{si } n \text{ es par.} \end{cases}$$

Como la normal estándar es simétrica alrededor del cero, las variables aleatorias X y $-X$ tienen ambas distribución normal estándar, entonces $F_{X_n}(x) = F_X(x)$ para toda $x \in \mathbb{R}$,es decir que $X_n \xrightarrow{D} X$. Por otro lado, $|X_{2n} - X| = 2|X|$ para cualquier $n \in \mathbb{N}$, así que

$$P(|X_{2n} - X| > \varepsilon) = P(2|X| > \varepsilon) = P\left(|X| > \frac{\varepsilon}{2}\right),$$

para cualquier $\varepsilon > 0$. *Por lo tanto*

$$\lim_{n\to\infty} P\left(|X_{2n} - X| > \varepsilon\right) = P\left(|X| > \frac{\varepsilon}{2}\right) > 0,$$

es decir, que $\{X_n\}$ ***converge en distribución pero no converge en probabilidad a*** X.

Ejemplo 1.6.2 *Usando el mismo ejemplo anterior,*

$$P\left(\lim_{n\to\infty} X_n = X\right) = 0 \quad \textit{si } n \textit{ es par,}$$

por lo que $\{X_n\}$ ***converge en distribución pero no converge casi seguramente a*** X.

Ejemplo 1.6.3 *Análogamente, con el mismo ejemplo,*

$$E\left[|X_n - X|^p\right] = E\left[2^p\,|X|^p\right] = 2^p E\left[|X|^p\right], \qquad\qquad \textit{si } n \textit{ es impar,}$$

calculemos $E\left[|X|^p\right]$

$$E\left[|X|^p\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^p\, e^{-\frac{1}{2}x^2} dx = \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} |x|^p\, e^{-\frac{1}{2}x^2} dx,$$

haciendo $y = \frac{1}{2}x^2$

$$E\left[|X|^p\right] = \frac{2^{p/2}}{\sqrt{\pi}} \int_0^{\infty} y^{\frac{p-1}{2}} e^{-y} dx = \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right) > 0,$$

por lo que $\{X_n\}$ ***converge en distribución pero no converge en*** $\mathbf{L}^p$ ***a*** X.

Ejemplo 1.6.4 *Considere el espacio* $((0,1], \beta(0,1), P)$ *,donde* P *es la medida de Lebesgue sobre* Ω*. Sean*

$$\begin{aligned}
X_1 &= I_{(0,1]}, \\
X_2 &= I_{(0,1/2]}, \quad X_3 = I_{(1/2,1]}, \\
X_4 &= I_{(0,1/3],} \quad X_5 = I_{(1/3,2/3]}, \quad X_6 = I_{(2/3,1]}, \\
X_7 &= I_{(0,1/4]}, \quad X_8 = I_{(1/4,2/4]}, \quad X_9 = I_{(2/4,3/4]}, \quad X_{10} = I_{(3/4,1]}, \\
&\ldots\ldots\ldots\ldots \\
X_n &= I_{\left(\frac{j_n - 1}{k_n}, \frac{j_n}{k_n}\right]}, \qquad \text{donde } k_n \text{ se incrementa con } n.
\end{aligned}$$

El

$$\lim_{n\to\infty} P\left[|X_n| > \varepsilon\right] = \lim_{n\to\infty} \frac{1}{k_n} = 0,$$

sin embargo

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) \ \textit{exista}\right\} = \emptyset,$$

es decir que $\{X_n\}$ ***converge en probabilidad pero no converge casi seguramente.***

Ejemplo 1.6.5 *Usando el ejemplo anterior, se observa que* ***converge en*** L^p ***pero no converge casi seguramente.***

Ejemplo 1.6.6 *Considere el espacio* $((0,1], \mathfrak{B}(0,1), P)$ *,donde* P *es la medida* *de Lebesgue sobre* Ω. *Sean*

$$\begin{aligned}
X_1 &= I_{(0,1]}, \\
X_2 &= 2^{1/p} I_{(0,1/2]}, \qquad X_3 = 2^{1/p} I_{(1/2,1]}, \\
X_4 &= 3^{1/p} I_{(0,1/3]}, \qquad X_5 = 3^{1/p} I_{(1/3,2/3]}, \qquad X_6 = 3^{1/p} I_{(2/3,1]}, \\
X_7 &= 4^{1/p} I_{(0,1/4]}, \quad X_8 = 4^{1/p} I_{(1/4,2/4]}, \quad X_9 = 4^{1/p} I_{(2/4,3/4]}, \quad X_{10} = 4^{1/p} I_{(3/4,1]}, \\
&\qquad \ldots\ldots\ldots \\
X_n &= k_n^{1/p} I_{\left(\frac{j_n - 1}{k_n}, \frac{j_n}{k_n}\right]}, \qquad \text{donde } k_n \text{ se incrementa con } n \text{ y sea } p > 0,
\end{aligned}$$

entonces

$$\lim_{n\to\infty} P\left[|X_n| > \varepsilon\right] = \lim_{n\to\infty} \frac{1}{k_n} = 0,$$

sin embargo

$$E\left[|X_n|^p\right] = \int_{(j_n-1)/k_n}^{j_n/k_n} k_n \, dx = 1,$$

es decir que, $\{X_n\}$ ***converge en probabilidad pero no converge en*** L^p.

Ejemplo 1.6.7 *Considere el espacio* $((0,1], \mathfrak{B}(0,1), P)$ *con* P *la medida de probabilidad uniforme, definimos*

$$X_n(\omega) = \begin{cases} 2^n & si\ \omega \in \left(0, \frac{1}{n}\right), \\ 0 & e.o.c., \end{cases}$$

y

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = 0\right\} = (0,1] = \Omega,$$

entonces

$$P\left[\lim_{n\to\infty} X_n = 0\right] = 1.$$

Pero

$$E\left[|X_n|^p\right] = \frac{2^{np}}{n},$$

calculando el límite cuando $n \to \infty$ *entonces*

$$E\left[|X_n|^p\right] \to +\infty,$$

es decir que $\{X_n\}$ ***converge casi seguramente pero no converge en*** L^p.

Ahora vamos a demostrar lo que se cumple que es

$$X_n \stackrel{cs}{\to} X \quad \Rightarrow \quad X_n \stackrel{P}{\to} X,$$

$$X_n \stackrel{P}{\to} X \quad \Rightarrow \quad X_n \stackrel{D}{\to} X,$$

$$X_n \stackrel{L}{\to} X \quad \Rightarrow \quad X_n \stackrel{P}{\to} X,$$

$$X_n \stackrel{L^2}{\to} X \quad \Rightarrow \quad X_n \stackrel{L}{\to} X.$$

Proposición 1.6.8 *Sea* $\{X_n\}$ *una sucesión de variables aleatorias tal que* $X_n \stackrel{cs}{\to} X$. *Entonces* $X_n \stackrel{P}{\to} X$.

Demostración. *Que* $X_n \stackrel{cs}{\to} X$ *implica que* $P\left[\lim_n \sup\left(|X_n - X| > \varepsilon\right)\right] = 0,$

$$P\left[\limsup_n \left(|X_n - X| > \varepsilon\right)\right] = \lim_{m\to\infty} P\left[\bigcup_{n=m}^{\infty} \left(|X_n - X| > \varepsilon\right)\right],$$

Por otro lado sabemos que $\{\omega \in \Omega \mid |X_m - X| > \varepsilon\} \subseteq \bigcup_{n=m}^{\infty} \{\omega \in \Omega \mid |X_n - X| > \varepsilon\},$

por lo que

$$\lim_{n\to\infty} P\left[|X_n - X| > \varepsilon\right] \leq \lim_{m\to\infty} P\left[\bigcup_{n=m}^{\infty} \left(|X_n - X| > \varepsilon\right)\right] = 0. \ \blacksquare$$

Proposición 1.6.9 *Sea $\{X_n\}$ una sucesión de variables aleatorias tal que $X_n \overset{P}{\to} X$. Entonces $X_n \overset{D}{\to} X$.*

Demostración. *Para* $\varepsilon > 0$, $n \in \mathbb{N}$ *y* $t \in \mathbb{R}$, *se tiene*

$$\begin{aligned}
F_X(t-\varepsilon) &= P(X \le t-\varepsilon) \\
&= P(X \le t-\varepsilon, |X - X_n| > \varepsilon) + P(X \le t-\varepsilon, |X - X_n| \le \varepsilon) \\
&= P(X \le t-\varepsilon, |X - X_n| > \varepsilon) \\
&\quad + P(X \le t-\varepsilon, X - \varepsilon \le X_n \le X + \varepsilon) \\
&\le P(|X - X_n| > \varepsilon) + P(X \le t) = P(|X - X_n| > \varepsilon) + F_{X_n}(t).
\end{aligned}$$

Por otro lado,

$$\begin{aligned}
F_{X_n}(t) &= P(X_n \le t) \\
&= P(X_n \le t, |X - X_n| > \varepsilon) + P(X_n \le t, |X - X_n| \le \varepsilon) \\
&= P(X_n \le t, |X - X_n| > \varepsilon) + P(X_n \le t, X_n - \varepsilon \le X \le X_n + \varepsilon) \\
&\le P(|X - X_n| > \varepsilon) + P(X \le t+\varepsilon) = P(|X - X_n| > \varepsilon) + F_X(t+\varepsilon).
\end{aligned}$$

De las dos desigualdades anteriores, se tiene que

$$F_X(t-\varepsilon) - P(|X - X_n| > \varepsilon) \le F_{X_n}(t) \le F_X(t+\varepsilon) + P(|X - X_n| > \varepsilon).$$

Tendiendo n a ∞ y utilizando el hecho de que $X_n \overset{P}{\to} X$, se obtiene que para cualquier $\varepsilon > 0$, $n \in \mathbb{N}$ y $t \in \mathbb{R}$,

$$F_X(t-\varepsilon) \le \lim \inf_{n\to\infty} F_{X_n}(t) \le \lim \sup_{n\to\infty} F_{X_n}(t) \le F_X(t+\varepsilon).$$

Ahora, si t es un punto de continuidad de F_X, entonces, tomando límites cuando $\varepsilon \to 0$, se obtiene

$$F_X(t) \leq \liminf_{n\to\infty} F_{X_n}(t) \leq \limsup_{n\to\infty} F_{X_n}(t) \leq F_X(t),$$

así que $\lim_{n\to\infty} F_{X_n}(t) = F_X(t)$. ■

Hay un caso particular en el que la convergencia en distribución sí implica la convergencia en probabilidad y es el siguiente:

Proposición 1.6.10 *Sea $\{X_n\}$ una sucesión de variables aleatorias tal que $X_n \xrightarrow{D} 0$. Entonces $X_n \xrightarrow{P} 0$.*

Demostración. *Por hipótesis sabemos que*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) = \begin{cases} 0 & si\ x < 0, \\ 1 & si\ x > 0. \end{cases}$$

Para $\varepsilon > 0$,

$$\begin{aligned} P(|X_n| > \varepsilon) &= P(X_n > \varepsilon) + P(X_n < -\varepsilon) \\ &\leq P(X_n > \varepsilon) + P(X_n \leq -\varepsilon) \\ &= 1 - F_{X_n}(\varepsilon) + F_{X_n}(-\varepsilon), \end{aligned}$$

tomando el límite se obtiene $\lim_{n\to\infty} P(|X_n| > \varepsilon) = 0$. ■

Corolario 1.6.11 *Sean $\{X_n\}$ y $\{Y_n\}$ dos sucesiones de variables aleatorias:*

1. *Si $X_n \xrightarrow{D} 0$ y $Y_n \xrightarrow{D} 0$, entonces $X_n + Y_n \xrightarrow{D} 0$.*

2. *Si $X_n \xrightarrow{D} 0$ y $Y_n \xrightarrow{D} 0$, entonces $X_n Y_n \xrightarrow{D} 0$.*

Demostración.

1. *Por la proposición anterior $X_n \xrightarrow{P} 0$ y $Y_n \xrightarrow{P} 0$, y por las propiedades de convergencia en probabilidad*

$$X_n + Y_n \xrightarrow{P} 0,$$

y como la convergencia en probabilidad implica la convergencia en distribución, entonces

$$X_n + Y_n \xrightarrow{D} 0.$$

2. *Análogamente $X_n Y_n \xrightarrow{D} 0$.*

■

Proposición 1.6.12 *Sea $\{X_n\}$ una sucesión de v.a. tal que $X_n \xrightarrow{L^1} 0$.*

Entonces $X_n \xrightarrow{P} 0$.

Demostración. *Por hipótesis tenemos que*

$$\lim_{n\to\infty} E\left[|X_n - X|\right] = 0,$$

y sabemos que para $\varepsilon > 0$

$$\begin{aligned} E\left[|X_n - X|\right] &= E\left[|X_n - X| \cdot I_{|X_n - X| > \varepsilon}\right] + E\left[|X_n - X| \cdot I_{|X_n - X| \leq \varepsilon}\right] \\ &\geq E\left[|X_n - X| \cdot I_{|X_n - X| > \varepsilon}\right] = \int_0^{\varepsilon} P\left(|X_n - X| > \varepsilon\right)\, dx \\ &\geq \varepsilon P\left(|X_n - X| > \varepsilon\right) \geq 0, \end{aligned}$$

dividiendo entre ε *y tomando límites obtenemos*

$$0 \geq \lim_{n \to \infty} P\left(|X_n - X| > \varepsilon\right) \geq 0,$$

por lo tanto

$$\lim_{n \to \infty} P\left(|X_n - X| > \varepsilon\right) = 0. \qquad \blacksquare$$

Proposición 1.6.13 *Sea* $\{X_n\}$ *una sucesión de v.a. tal que* $X_n \overset{L^2}{\to} 0$.

Entonces $X_n \overset{L^1}{\to} 0$.

Demostración. *La desigualdad de Cauchy -Schwarz nos dice que*

$$E\left[|YX|\right] \leq \sqrt{E\left[X^2\right]}\sqrt{E\left[Y^2\right]},$$

sustituyendo a $X = X_n - X$ *y* $Y \equiv 1$

$$E\left[|X_n - X|\right] \leq \sqrt{E\left[|X_n - X|^2\right]},$$

como $|X_n - X|$ *es una variable aleatoria que toma valores en* $\mathbb{R}^+ \cup \{0\}$, *entonces*

$$0 \leq E\left[|X_n - X|\right] \leq \sqrt{E\left[|X_n - X|^2\right]},$$

tomando límite

$$0 \leq \lim_{n \to \infty} E\left[|X_n - X|\right] \leq 0,$$

es decir

$$\lim_{n\to\infty} E\left[|X_n - X|\right] = 0. \qquad \blacksquare$$

CAPÍTULO 2

CONVERGENCIA DE SERIES

2.1 Introducción

En el capítulo anterior se vieron los principales tipos de convergencias en variables aleatorias. En este capítulo habléremos de la variable aleatoria compuesta por la serie $\sum_n X_n$, y amerita un capítulo aparte pues es apartir de esta nueva variable aleatoria que se desprenden los Teoremas límite más importantes en la probabilidad y la estadística como son La ley de los grandes números y El Teorema Central del Límite.

Tomando en cuenta estas consideraciones primero veamos algunos conceptos que nos ayudarán a probar estos teoremas.

Los libros usados para este tema fueron [1, K.L. Chung], [4, García, Miguel, I], [5, García, Miguel, II], [10, Carrasco, Guadalupe], [11, Rincón, Luis].

Proposición 2.1.1 (Desigualdad de Chebyshev) *Sea X cualquier variable aleatoria y ε cualquier número real positivo, entonces*

$$P\left[|X| \geq \varepsilon\right] \leq \frac{1}{\varepsilon} E\left[|X|\right].$$

Demostración.

$$\begin{aligned} E\left[|X|\right] &= E\left[|X|\cdot I_{[(|X|\geq\varepsilon)]}\right]+E\left[|X|\cdot I_{[(|X|<\varepsilon)]}\right] \\ &\geq E\left[|X|\cdot I_{[(|X|\geq\varepsilon)]}\right]=\int_0^{\varepsilon}P\left(|X|\geq x\right)dx \\ &\geq \int_0^{\varepsilon}P\left(|X|\geq\varepsilon\right)dx=\varepsilon P\left(|X|\geq\varepsilon\right). \end{aligned}$$

■

Corolario 2.1.2 *Sea X cualquier variable aleatoria, sean ε un número real positivo y r cualquier natural, entonces*

$$P\left[|X|\geq\varepsilon\right]\leq\frac{1}{\varepsilon^r}E\left[|X|^r\right].$$

Demostración.

$$P\left[|X|\geq\varepsilon\right]=P\left[|X|^r\geq\varepsilon^r\right]$$

aplicando la desigualdad para $|X|^r$ y usando ε^r,

$$P\left[|X|\geq\varepsilon\right]=P\left[|X|^r\geq\varepsilon^r\right]\leq\tfrac{1}{\varepsilon^r}E\left[|X|^r\right]. \qquad ■$$

Proposición 2.1.3 (Desigualdad de Chebyshev) *Sea X variable aleatoria de esperanza finita, y $\varepsilon>0$, entonces*

$$P\left(|X-E\left(X\right)|\geq\varepsilon\right)\leq\frac{1}{\varepsilon^2}var\left(X\right).$$

Demostración.

$$\begin{aligned} var\left(X\right) &= E\left[\left(X-E\left(X\right)\right)^2\right] \\ &\geq E\left[\left(X-E\left(X\right)\right)^2 I_{(|X-E(X)|\geq\varepsilon)}\right]=\int_0^{\varepsilon^2}P\left(|X-E\left(X\right)|\geq x\right)dx \end{aligned}$$

$$\geq \textstyle\int_0^{\varepsilon^2}P\left(|X-E\left(X\right)|\geq\varepsilon\right)dx=\varepsilon^2P\left(|X-E\left(X\right)|\geq\varepsilon\right). \qquad ■$$

Definición 2.1.4 *Dos sucesiones de variables aleatoria* $\{X_n\}$ *y* $\{Y_n\}$ *son llamadas equivalentes si*

$$\sum_n P(X_n \neq Y_n) < \infty.$$

Proposición 2.1.5 *Si* $\{X_n\}$ *y* $\{Y_n\}$ *son equivalentes, entonces* $\sum_n (X_n - Y_n)$ *converge casi seguramente*

Demostración. *Por el lema de Borel-Cantelli si* $\sum_n P(X_n \neq Y_n) < \infty$ *entonces*

$$P(X_n \neq Y_n \text{ en una infinidad de valores de } n) = 0,$$

por lo que existe un conjunto N *de probabilidad cero tal que si* $\omega \in \Omega \setminus N$ *entonces existe* $n_0(\omega)$ *tal que*

$$\begin{aligned} n &\geq n_0(\omega) \Rightarrow X_n(\omega) = Y_n(\omega), \\ &\Rightarrow X_n(\omega) - Y_n(\omega) = 0, \end{aligned}$$

es decir que para todo ω *las sucesiones* $\{X_n(\omega)\}$ *y* $\{Y_n(\omega)\}$ *difieren en un conjunto de probabilidad cero llamémoslo* N_ω*, entonces*

$$\left\{\omega \mid \lim_{n\to\infty} \sum_{j=1}^{n} |X_j(\omega) - Y_j(\omega)| > 0\right\} = \bigcup_{\omega\in\Omega} N_\omega,$$

y

$$P\left(\bigcup_{\omega\in\Omega} N_\omega\right) \leq \sum P(N_\omega) = 0,$$

por lo tanto

$$\textstyle\sum_{j=1}^{n} (X_j(\omega) - Y_j(\omega)) \xrightarrow{c.s.} 0. \qquad \blacksquare$$

Proposición 2.1.6 (Desigualdad de Kolmogorov) *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes de varianza finita y ε cualquier número real positivo, entonces*

$$P\left[\max_{1\leq j\leq n}|S_j - E[S_j]| > \varepsilon\right] \leq \frac{\sigma^2(S_n)}{\varepsilon^2},$$

para $j \in \{1,2,3,..,n\}$ y $S_j = \sum_{i=1}^{j} X_i$

Demostración. *Sea*

$$Y_i = X_i - E[X_i], \tag{2.1.1}$$

y sea

$$\breve{S}_j = \sum_{i=1}^{j} Y_i = S_j - E[S_j], \tag{2.1.2}$$

entonces

$$\begin{aligned} E\left[\breve{S}_j\right] &= E[S_j - E[S_j]] = 0, \\ \sigma^2\left(\breve{S}_j\right) &= \sigma^2(S_j - E[S_j]) = \sigma^2(S_j), \end{aligned}$$

es decir, probaremos que

$$P\left[\max_{1\leq j\leq n}\left|\breve{S}_j\right| > \varepsilon\right] \leq \frac{\sigma^2\left(\breve{S}_n\right)}{\varepsilon^2}.$$

Sea $\varepsilon > 0$. Para cualquier $\omega \in \Omega$

$$M = \left\{\omega \mid \max_{1\leq j\leq n}\left|\breve{S}_j\right| > \varepsilon\right\}, \tag{2.1.3}$$

$$m(\omega) = \left\{j \mid 1 \leq j \leq n,\ \left|\breve{S}_j(\omega)\right| > \varepsilon\right\},$$

$m(\omega)$ indica la j para que $\breve{S}_j$ sea la más pequeña del conjunto de $\left\{\breve{S}_n\right\}$ y además exceda a ε

$$\begin{aligned} M_k &= \{\omega \mid m(\omega)=k\} = \left\{\omega \mid \max_{1\le j\le k-1}\left|\breve{S}_j\right| \le \varepsilon, \quad \left|\breve{S}_k\right| > \varepsilon\right\} \\ \text{para } k &= 1 \quad M_1 = \left\{\omega \mid \left|\breve{S}_1\right| > \varepsilon\right\}, \end{aligned} \tag{2.1.4}$$

M_k es el conjunto de ω que hacen cumplir que $\breve{S}_k$ sea la más pequeña y también que exceda a ε. Observese que

$$M = \bigcup_{k=1}^{n} M_k,$$

El conjunto de M_k son mutuamente excluyentes ya que si $\omega \in M_s$ y $\omega \in M_t$ entonces $s = m(\omega) = t$

Como $E\left[\breve{S}_n\right] = 0 \Rightarrow \sigma^2\left(\breve{S}_n\right) = E\left[\breve{S}_n^2\right]$ entonces

$$\begin{aligned} \sigma^2\left(\breve{S}_n\right) &= E\left[\breve{S}_n^2\right] \\ &\ge E\left[\breve{S}_n^2 I_M\right] = \sum_{k=1}^{n} E\left[\breve{S}_n^2 I_{M_k}\right] \\ &= \sum_{k=1}^{n} E\left[\left(\breve{S}_k + \breve{S}_n - \breve{S}_k\right)^2 I_{M_k}\right] \\ &= \sum_{k=1}^{n} E\left[\left(\breve{S}_k^2 + 2\breve{S}_k\left(\breve{S}_n - \breve{S}_k\right) + \left(\breve{S}_n - \breve{S}_k\right)^2\right) I_{M_k}\right], \end{aligned}$$

y como el conjunto de Y_i esta conformado de v.a. independientes, se tiene que

$$\begin{aligned} E\left[\breve{S}_k\left(\breve{S}_n - \breve{S}_k\right) I_{M_k}\right] &= E\left[\left(\sum_{i=1}^{k} Y_i\right)\left(\sum_{i=k+1}^{n} Y_i\right) I_{M_k}\right] \\ &= E\left[\left(\sum_{i=1}^{k} Y_i\right) I_{M_k}\right] E\left[\left(\sum_{i=k+1}^{n} Y_i\right)\right] = 0, \end{aligned}$$

por lo tanto

$$\begin{aligned}\sigma^2\left(\breve{S}_n\right) &\geq \sum_{k=1}^{n} E\left[\breve{S}_k^2\ I_{M_k}\right] \geq \sum_{k=1}^{n} E\left[\varepsilon^2 I_{M_k}\right]\\ &= \varepsilon^2 \sum_{k=1}^{n} P\left(M_k\right) = \varepsilon^2 P\left(M\right),\end{aligned}$$

o bien

$$P\left[\max\nolimits_{1\leq j\leq n}\left|\breve{S}_j\right| > \varepsilon\right] \leq \frac{\sigma^2\left(\breve{S}_n\right)}{\varepsilon^2}. \qquad \blacksquare$$

Proposición 2.1.7 *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes de esperanza finita. Suponga que existe* A *tal que*

$$|X_i - E\left[X_i\right]| \leq A < \infty,$$

entonces para cualquier $\varepsilon > 0$,

$$P\left[\max_{1\leq j\leq n}|S_j| \leq \varepsilon\right] \leq \frac{(4\varepsilon + 2A)^2}{\sigma^2\left(S_n\right)}.$$

Demostración. *Definamos*

$$W_k = \left\{\omega \mid \max_{1\leq j\leq k}|S_j| \leq \varepsilon\right\},$$

y definiendo M *como en* (2.1.3) *y* M_k *como en* (2.1.4) *para* $k \geq 1$ *entonces*

$$\begin{aligned}M_k &= \left\{\omega \mid \max_{1\leq j\leq k-1}|S_j| \leq \varepsilon,\quad |S_k| > \varepsilon\right\}\\ &= \left\{\omega \mid \max_{1\leq j\leq k-1}|S_j| \leq \varepsilon,\quad \max_{1\leq j\leq k}|S_k| > \varepsilon\right\}\\ &= \left\{\omega \mid \max_{1\leq j\leq k-1}|S_j| \leq \varepsilon\right\} \cap \left\{\omega \mid \max_{1\leq j\leq k}|S_k| > \varepsilon\right\}\\ &= \left\{\omega \mid \max_{1\leq j\leq k-1}|S_j| \leq \varepsilon\right\} \setminus \left\{\omega \mid \max_{1\leq j\leq k}|S_k| \leq \varepsilon\right\}\\ &= W_{k-1} \setminus W_k. \qquad (2.1.5)\end{aligned}$$

Si $P(W_k) = 0$ *es inmediato, entonces suponganos* $P(W_k) > 0$. *Sea* Y_i *como en* (2.1.1) *y* $\breve{S}_j$ *como en* (2.1.2) *para* $k = 0$ *entonces* $\breve{S}_k = 0$. *Para* $0 \leq k \leq n$ *definamos*

$$\begin{aligned} a_k &= E\left[\breve{S}_k \mid W_k\right] \\ &= E\left[(S_k - E[S_k]) \mid W_k\right] \\ &= E[S_k \mid W_k] - E[S_k], \end{aligned} \tag{2.1.6}$$

así

$$\begin{aligned} E\left[\left(\breve{S}_k - a_k\right) \ I_{W_k}\right] &= E\left[\breve{S}_k I_{W_k}\right] - a_k E\left[I_{W_k}\right] \\ &= E\left[\breve{S}_k I_{W_k}\right] - a_k P[W_k] \\ &= E\left[\breve{S}_k I_{W_k}\right] - E\left[\breve{S}_k \mid W_k\right] P[W_k] \\ &= E\left[\breve{S}_k I_{W_k}\right] - E\left[\breve{S}_k I_{W_k}\right] = 0, \end{aligned} \tag{2.1.7}$$

por otra parte, por (2.1.2) *y* (2.1.6) *y por cómo definimos* W_k

$$\begin{aligned} \left|\breve{S}_k - a_k\right| &= \left|S_k - E[S_k] - E[S_k \mid W_k] + E[S_k]\right| \\ &\leq |S_k| + \varepsilon, \end{aligned} \tag{2.1.8}$$

$$\begin{aligned}
|a_k - a_{k+1}| &= |E[S_k \mid W_k] - E[S_k] - E[S_{k+1} \mid W_{k+1}] + E[S_{k+1}]| \\
&= |E[S_k \mid W_k] - E[S_{k+1} \mid W_{k+1}] + E[X_{k+1}]| \\
&= |E[S_k \mid W_k] - E[S_k \mid W_{k+1}] - E[X_{k+1} \mid W_{k+1}] + E[X_{k+1}]| \\
&= |E[S_k \mid W_k] - E[S_k \mid W_{k+1}] - E[(X_{k+1} - E[X_{k+1}]) \mid W_{k+1}]| \\
&\le |E[S_k \mid W_k]| + |E[S_k \mid W_{k+1}]| + |E[(X_{k+1} - E[X_{k+1}]) \mid W_{k+1}]| \\
&\le \varepsilon + \varepsilon + A = 2\varepsilon + A, \qquad (2.1.9)
\end{aligned}$$

y por (2.1.5)

$$\begin{aligned}
E\left[\left(\breve{S}_{k+1} - a_{k+1}\right)^2 I_{W_{k+1}}\right] &= E\left[\left(\breve{S}_k + Y_{k+1} - a_{k+1} + a_k - a_k\right)^2 I_{W_k}\right] \qquad (2.1.10) \\
&\quad - E\left[\left(\breve{S}_k + Y_{k+1} - a_{k+1} + a_k - a_k\right)^2 I_{M_{k+1}}\right] \\
&\qquad (2.1.11)
\end{aligned}$$

y

$$\begin{aligned}
(\mathit{2.1.10}) &= E\left[\left[\begin{array}{c} \left(\breve{S}_k - a_k\right)^2 + (a_k - a_{k+1})^2 + Y_{k+1}^2 + 2\left(\breve{S}_k - a_k\right)(a_k - a_{k+1}) \\ +2\left(\breve{S}_k - a_k\right) Y_{k+1} + 2(a_k - a_{k+1}) Y_{k+1} \end{array}\right] I_{W_k}\right] \\
&\ge E\left[\left(\breve{S}_k - a_k\right)^2 \ I_{W_k}\right] + E\left[Y_{k+1}^2 \ I_{W_k}\right] \\
&= E\left[\left(\breve{S}_k - a_k\right)^2 \ I_{W_k}\right] + \sigma^2(X_{k+1}) P(W_k),
\end{aligned}$$

y por (2.1.8) *y* (2.1.9) *y la definición de* M_{k+1}

$$\begin{aligned} -\,(2.1.11) &= E\left[\left[\left(\breve{S}_k - a_k\right) + (a_k - a_{k+1}) + Y_{k+1}\right]^2 \ I_{M_{k+1}}\right] \\ &\leq E\left[\left[\left|\breve{S}_k - a_k\right| + |a_k - a_{k+1}| + |Y_{k+1}|\right]^2 \ I_{M_{k+1}}\right] \\ &\leq E\left[\left[\left|\breve{S}_k - a_k\right| + |a_k - a_{k+1}| + |Y_{k+1}|\right]^2 \ I_{M_{k+1}}\right] \\ &\leq E\left[\left[|S_k| + \varepsilon + 2\varepsilon + A + A\right]^2 \ I_{M_{k+1}}\right] \\ &\leq (4\varepsilon + 2A)^2 P\left(M_{k+1}\right), \end{aligned}$$

entonces

$$\begin{aligned} E\left[\left(\breve{S}_{k+1} - a_{k+1}\right)^2 \ I_{W_{k+1}}\right] - E\left[\left(\breve{S}_k - a_k\right)^2 \ I_{W_k}\right] &\geq \sigma^2\left(X_{k+1}\right) P\left(W_k\right) \\ &\quad - (4\varepsilon + 2A)^2 P\left(M_{k+1}\right), \end{aligned}$$

claramente $W_n \subset W_k$ *para* $0 \leq k \leq n-1$ *ya que*

$$\begin{aligned} W_n &= \left\{\omega \mid \max_{1\leq j\leq n} |S_j| \leq \varepsilon\right\} \\ &= \left\{\omega \mid \max_{1\leq j\leq k} |S_j| \leq \varepsilon, \max_{k+1\leq j\leq n} |S_j| \leq \varepsilon\right\}, \end{aligned}$$

entonces $P\left(W_n\right) < P\left(W_k\right)$ *y para* $k = n$ $P\left(W_k\right) = P\left(W_n\right)$ *entonces*

$$\begin{aligned} E\left[\left(\breve{S}_{k+1} - a_{k+1}\right)^2 \ I_{W_{k+1}}\right] - E\left[\left(\breve{S}_k - a_k\right)^2 \ I_{W_k}\right] &\geq \sigma^2\left(X_{k+1}\right) P\left(W_n\right) \\ &\quad - (4\varepsilon + 2A)^2 P\left(M_{k+1}\right), \end{aligned}$$

sumando para toda k *en donde nuestras esperanzas y probabilidades esten definidas*

$$\begin{aligned} &\sum_{k=0}^{n-1}\left(E\left[\left(\breve{S}_{k+1} - a_{k+1}\right)^2 \ I_{W_{k+1}}\right] - E\left[\left(\breve{S}_k - a_k\right)^2 \ I_{W_k}\right]\right) \\ \geq\ &\sum_{k=0}^{n-1}\sigma^2\left(X_{k+1}\right) P\left(W_n\right) - \sum_{k=0}^{n-1}(4\varepsilon + 2A)^2 P\left(M_{k+1}\right). \end{aligned}$$

Por lo tanto

$$\begin{aligned} E\left[\left(\breve{S}_n - a_n\right)^2 \; I_{W_n}\right] - E\left[\left(\breve{S}_0 - a_0\right)^2 \; I_{W_0}\right] &\geq P\left(W_n\right)\sum_{k=0}^{n-1}\sigma^2\left(X_{k+1}\right) \\ &\quad - \left(4\varepsilon + 2A\right)^2 \sum_{k=0}^{n-1} P\left(M_{k+1}\right), \end{aligned}$$

es decir que

$$\begin{aligned} E\left[\left(\breve{S}_n - a_n\right)^2 \; I_{W_n}\right] &\geq P\left(W_n\right)\sigma^2\left(S_n\right) - \left(4\varepsilon + 2A\right)^2 P\left(M\right) \\ &= P\left(W_n\right)\sigma^2\left(S_n\right) - \left(4\varepsilon + 2A\right)^2 P\left(W_n^c\right), \end{aligned}$$

y

$$E\left[\left(\breve{S}_n - a_n\right)^2 \; I_{W_n}\right] \leq E\left[\left(\left|S_n\right| + \varepsilon\right)^2 \; I_{W_n}\right] \leq 4\varepsilon^2 P\left(W_n\right),$$

juntando estos dos últimos resultados

$$\begin{aligned} &\Rightarrow \; 4\varepsilon^2 P\left(W_n\right) \geq P\left(W_n\right)\sigma^2\left(S_n\right) - \left(4\varepsilon + 2A\right)^2\left[1 - P\left(W_n\right)\right], \\ \left(4\varepsilon + 2A\right)^2 &\geq \left(4\varepsilon + 2A\right)^2 - P\left(W_n\right)\left[\left(4\varepsilon + 2A\right)^2 - 4\varepsilon^2\right] \geq P\left(W_n\right)\sigma^2\left(S_n\right), \end{aligned}$$

por lo que obtenemos el resultado, es decir

$$P\left(W_n\right) \leq \frac{\left(4\varepsilon + 2A\right)^2}{\sigma^2\left(S_n\right)}. \qquad \blacksquare$$

2.2 Ley de los Grandes Números

La frase *ley de los grandes números* es usada ocasionalmente para referirse al principio de que la probabilidad, de cualquier evento posible (incluso uno improbable), incrementa con el número de eventos en la serie. Por ejemplo, la probabilidad de que un individuo gane la lotería es bastante baja; sin embargo, la probabilidad de que alguien gane la lotería es bastante alta, suponiendo, claro, que muchas personas compraron boletos de lotería.

Se dice que una sucesión de variables aleatorias definidas sobre un espacio de probabilidad común obedece la *ley de los grandes números* cuando la media de las muestras de variables $\left(\frac{\sum_{j=1}^{n} X_i}{n}\right)$ tiende a la media de las esperanzas de las variables aleatorias de la sucesión$\left(\frac{\sum_{j=1}^{n} E\left[X_i\right]}{n}\right)$, cuando el número total de variables aumenta.

En **Estadística**, las leyes de los grandes números implican que el promedio de una muestra al azar de una población de *gran tamaño* tenderá a estar cerca de la mediana de la población. En el contexto de **Teoría de Probabilidad**, la interpretación no debe tomarse así, las leyes de los grandes números dicen que el promedio de una secuencia de variables elegidas al azar con una distribución de probabilidad común, converge en probabilidad o casi seguramente (en la ley débil o en la ley fuerte, respectivamente) a su valor esperado.

2.2.1 Ley Débil de los Grandes Números

La *ley débil de los grandes números* fue demostrada por primera vez por J. Bernoulli, el cual establece que si $\mathcal{E}$ es un experimento y A un evento relativo a este experimento de probabilidad p, y consideramos un nuevo experimento consistente en la repetición indefinida del experimento $\mathcal{E}$, de tal manera que cada repetición es independiente de las otras. Entonces, llamando X_n al número de veces que ocurre el evento A en n repeticiones, se tiene que

$$\frac{X_n}{n} \xrightarrow{P} p.$$

Este enunciado nos lleva, inmediatamente a pensar en la distribución binomial. Así que antes de ver las demostraciones que se desarrollaron al respecto de esta ley, veamos ese caso particular.

Ejemplo 2.2.1 *Sea X_n una v.a. que se distribuye como una binomial de parámetros n y p.*

$$P\left[\left|\frac{X_n}{n} - p\right| > \varepsilon\right] = P\left[|X_n - np| > n\varepsilon\right],$$

y como $E[X_n] = np$, usando la desigualdad de Chebyshev

$$P\left[|X_n - np| > n\varepsilon\right] \leq \frac{np(1-p)}{(n\varepsilon)^2} = \frac{1-p}{n\varepsilon^2},$$

calculando el límite se tiene que

$$\lim_{n\to\infty} P\left[|X_n - np| > n\varepsilon\right] = 0,$$

es decir que $\dfrac{X_n}{n} \xrightarrow{p} p.$

Más adelante veremos que para esta X_n se cumple también la convergencia casi segura.

En el caso en que las variables aleatorias no necesariamente están idénticamente distribuidas, pero las varianzas de éstas sean finitas, Markov demostró lo siguiente.

Teorema 2.2.2 (Ley débil de los grandes números-Markov) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes de varianza finita tales que*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{j=1}^{n} \sigma_j^2 = 0.$$

Entonces

$$\frac{1}{n} \sum_{j=1}^{n} X_j - \frac{1}{n} \sum_{j=1}^{n} \mu_j \xrightarrow{P} 0,$$

donde μ_j *es la esperanza de* X_j.

Demostración. *Para cada* $n \in \mathbb{N}$, *tomemos otra vez la variable aleatoria* $S_n = \dfrac{X_1 + \cdots + X_n}{n}$. *Su esperanza es*

$$E(S_n) = E\left(\frac{1}{n} \sum_{j=1}^{n} X_j\right) = \frac{1}{n} \sum_{j=1}^{n} E(X_j) = \frac{1}{n} \sum_{j=1}^{n} \mu_j,$$

y su varianza

$$var(S_n) = var\left(\frac{1}{n} \sum_{j=1}^{n} X_i\right) = \frac{1}{n^2} \sum_{j=1}^{n} \sigma_j^2.$$

Aplicando la desigualdad de Chebyshev se tiene

$$P(|S_n - E(S_n)| > \varepsilon) \leq \frac{1}{\varepsilon^2} var(Y_n) = \frac{\sigma^2}{n^2 \varepsilon^2} \sum_{j=1}^{n} \sigma_j^2,$$

entonces

$$\lim_{n\to\infty} P(|S_n - E(S_n)| > \varepsilon) = 0. \qquad \blacksquare$$

Más tarde (1867) el matemático sovietico Pafnuty L. Chebyshev publicó la forma general de este resultado.

Teorema 2.2.3 (Ley débil de los grandes números-Chebyshev) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes e identicamente distribuidas, de varianza finita. Entonces*

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{P} \mu,$$

donde μ *es la esperanza común de las variables aleatorias* $X_1, X_2, \ldots$

Demostración. *Sea* σ^2 *la varianza común de cada* X_n *y sea* $S_n = \dfrac{X_1 + \cdots + X_n}{n}$ *, entonces para cada* $n \in \mathbb{N}$

$$E\left[S_n\right] = \mu,$$

y

$$var\left(S_n\right) = var\left(\frac{1}{n}\sum_{j=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{j=1}^{n}\sigma^2 = \frac{\sigma^2}{n}.$$

Por la desigualdad de Chebyshev se tiene

$$P\left(|S_n - \mu| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} var\left(S_n\right) = \frac{\sigma^2}{n\varepsilon^2},$$

de donde

$$\lim_{n\to\infty} P\left(|S_n - \mu| > \varepsilon\right) = 0,$$

por lo tanto

$$S_n \xrightarrow{P} \mu. \qquad \blacksquare$$

2.2.2 Ley Fuerte de los Grandes Números

Andrey Nikolaevich Kolmogorov mostró que la versión de la Ley de los grandes números que hemos visto hasta aquí, se puede mejorar demostrando que la convergencia puede ser casi segura, y no sólo en probabilidad (por eso en algunas ocasiones a la convergencia casi segura se le llega a decir convergencia fuerte).

Pero antes de ver el resultado al que llegó Kolmogorov, veamos otro resultado más sensillo que es el de Borel, en el que se basó Kolmogorov para llegar al suyo.

Teorema 2.2.4 (Ley fuerte de los grandes números-Borel) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes, todas con distribución bernoulli de parámetro* $p \in (0,1)$. *Entonces*

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{cs} p.$$

Demostración. *Sea* $S_n = X_1 + X_2 + \cdots + X_n$. *Entonces* S_n *tiene distribución binomial con parámetros* n *y* p, *así que*

$E(S_n) = np,$

$E(S_n^2) = np + n(n-1)p^2,$

$E(S_n^3) = np + 3n(n-1)p^2 + n(n-1)(n-2)p^3,$

$E(S_n^4) = np + 7n(n-1)p^2 + 6n(n-1)(n-2)p^3 + n(n-1)(n-2)(n-3)p^4,$

por tanto

$$
\begin{aligned}
E\left[\left(\frac{S_n}{n}-p\right)^4\right] &= \frac{E\left(S_n^4\right)}{n^4}-4\frac{E\left(S_n^3\right)}{n^3}p+6\frac{E\left(S_n^2\right)}{n^2}p^2-4\frac{E\left(S_n\right)}{n}p^3+p^4 \\
&= \frac{np+7n\left(n-1\right)p^2+6n\left(n-1\right)\left(n-2\right)p^3}{n4} \\
&\quad +\frac{n\left(n-1\right)\left(n-2\right)\left(n-3\right)p^4}{n^4} \\
&\quad -4\frac{np^2+3n\left(n-1\right)p^3+n\left(n-1\right)\left(n-2\right)p^4}{n^3} \\
&\quad +6\frac{np^3+n\left(n-1\right)p^4}{n^2}-4\frac{np^4}{n}+p^4 \\
&= \frac{3np^2\left(1-p\right)^2}{n^3}-\frac{6p^2\left(1-p\right)^2}{n^3}+\frac{p\left(1-p\right)}{n^3} \\
&= \frac{1}{n^3}p\left(1-p\right)\left[3np\left(1-p\right)-6p\left(1-p\right)+1\right] \\
&\leq \frac{1}{4n^3}\left(\frac{3n}{4}+n\right)<\frac{1}{n^2}.
\end{aligned}
$$

Para $\varepsilon>0$ *y* $r\in\mathbb{N}$ *sabemos que*

$$P\left[|X|\geq\varepsilon\right]\leq\frac{1}{\varepsilon^r}E\left[|X|^r\right],$$

por lo que

$$P\left(\left|\frac{S_n}{n}-p\right|>\varepsilon\right)\leq\frac{E\left(\left(\frac{S_n}{n}-p\right)^4\right)}{\varepsilon^4}<\frac{1}{n^2\varepsilon^4}.$$

Como no podemos calcular inmediatamenete el límite cuando $n\to\infty$*, entonces usaremos* (1.3.10) *que dice que si*

$$\sum_{n=1}^{\infty}P\left(\left|\frac{S_n}{n}-p\right|>\varepsilon\right)<\infty \quad \textit{entonces} \quad \frac{S_n}{n}\xrightarrow{cs}p,$$

y en efecto, basta con ver que la serie $\sum_{n=1}^{\infty}\frac{1}{n^2}$ *converge. Por el criterio de la* *integral*

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converge si y sólo si } \int_1^{\infty} \frac{1}{x^2} dx \text{ converge,}$$

$$\int_1^{\infty} \frac{1}{x^2} dx = \lim_{b\to\infty} \int_1^{b} \frac{1}{x^2} = \lim_{b\to\infty} \left[\frac{b^{-1}-1}{-1}\right] = \lim_{b\to\infty} \left(1 - \frac{1}{b}\right) = 1,$$

por lo tanto

$$\frac{S_n}{n} \xrightarrow{cs} p. \qquad \blacksquare$$

Ahora veamos el resultado al que llegó Rajchman en 1932, usando subsucesiones.

Teorema 2.2.5 (Ley fuerte de los grandes números-Rajchman) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes e idénticamente distribuidas, de varianza finita. Entonces, para cualquier* $\varepsilon > 0$.

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{cs} \mu,$$

donde μ *es la esperanza común de* $X_1, X_2, X_3, ...$

Demostración. *Sea* $S_n = X_1 + X_2 + \cdots + X_n$ *y sea* $Y_n = \frac{S_n}{n} - \mu$ *. Entonces queremos probar que*

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = P\left(\lim_{n\to\infty} Y_n = 0\right) = 1.$$

Definamos

$$Z_n = \max_{\left\{k: n^2 \le k < (n+1)^2\right\}} \left|S_k - S_{n^2} - \mu\left(k - n^2\right)\right|,$$

entonces, para $n \in \mathbb{N}$ *y* $n^2 \leq k < (n+1)^2$

$$\begin{aligned} |Y_k| &= \frac{|S_k - k\mu|}{k} \leq \frac{|S_k - k\mu|}{n^2} \\ &\leq \frac{|S_{n^2} - n^2\mu|}{n^2} + \frac{|S_k - S_{n^2} - \mu(k-n^2)|}{n^2} \\ &= |Y_{n^2}| + \frac{|S_k - S_{n^2} - \mu(k-n^2)|}{n^2} \\ &\leq |Y_{n^2}| + \frac{Z_n}{n^2}, \end{aligned} \tag{2.2.12}$$

Como la sucesión de$\{X_n\}$ *es de v.a. independientes e idénticamente distribuidas, si* σ^2 *es su varianza común, entonces*

$$\begin{aligned} E[Y_n] &= E\left(\frac{1}{n}\sum_{j=1}^{n} X_i - \mu\right) = \frac{1}{n}\sum_{j=1}^{n}\mu - \mu = 0, \\ E\left[Y_n^2\right] &= E\left[\left(\frac{1}{n}\sum_{j=1}^{n} X_i - \mu\right)^2\right] = E\left[\left(\frac{1}{n}\sum_{j=1}^{n}(X_i - \mu)\right)^2\right] \\ &= \frac{1}{n^2}\sum_{j=1}^{n}\sigma^2. \end{aligned}$$

Ahora usando Chebyshev,

$$\sum_{n=1}^{\infty} P[|Y_{n^2}| > \varepsilon] \leq \sum_{n=1}^{\infty}\frac{1}{\varepsilon^2}E\left[|Y_{n^2}|^2\right] = \sum_{n=1}^{\infty}\frac{\sigma^2}{n^2\varepsilon^2} < \infty.$$

Y por (1.3.10)

$$P\left[\lim_{n\to\infty} Y_{n^2} = 0\right] = 1,$$

o lo que es lo mismo por(1.3.9)

$$P[\{\omega \in \Omega \mid (|Y_{n^2}(\omega)| > \varepsilon) \text{ para una infinidad de valores de } n\}] = 0, \tag{2.2.13}$$

es decir que si $\Omega^0 = \{\omega \in \Omega \mid (|Y_{n^2}(\omega)| \leq \varepsilon) \text{ para una infinidad de valores de } n\}$, *entonces*

$$P\left(\Omega^0\right) = 1.$$

Ahora falta comprobar que $\frac{Z_n}{n^2}$ *esta acotado, para ello analicemos*

$$Z_n^2 = \max_{\left\{k:n^2\leq k<(n+1)^2\right\}} \left|S_k - S_{n^2} - \mu\left(k-n^2\right)\right|^2 \leq \sum_{k=n^2}^{(n+1)^2-1} \left|S_k - S_{n^2} - \mu\left(k-n^2\right)\right|^2,$$

y como

$$\begin{aligned} E\left[\left|S_k - S_{n^2} - \mu\left(k-n^2\right)\right|^2\right] &= E\left[\left|\sum_{i=n^2+1}^{k}(X_i-\mu)\right|^2\right] = \sum_{i=n^2+1}^{k} E\left[(X_i-\mu)^2\right] \\ &\leq \sum_{i=n^2+1}^{(n+1)^2-1} E\left[(X_i-\mu)^2\right] = 2n\sigma^2, \end{aligned}$$

entonces,

$$\begin{aligned} E\left[Z_n^2\right] &\leq \sum_{k=n^2}^{(n+1)^2-1} E\left[\left|S_k - S_{n^2} - \mu\left(k-n^2\right)\right|^2\right] \leq (2n+1)\,2n\sigma^2 = \sigma^2\left(4n^2+2n\right) \\ &\leq 6\sigma^2 n^2. \end{aligned}$$

Usando, la desigualdad de Chebyshev

$$\sum_{n=1}^{\infty} P\left[\left|\frac{Z_n}{n^2}\right| > \varepsilon\right] \leq \sum_{n=1}^{\infty}\frac{1}{\varepsilon^2}E\left[\frac{Z_n^2}{n^4}\right] \leq \sum_{n=1}^{\infty}\frac{6\sigma^2 n^2}{\varepsilon^2 n^4} = \sum_{n=1}^{\infty}\frac{6\sigma^2}{\varepsilon^2 n^2} < \infty,$$

y por (1.3.10)

$$P\left[\lim_{n\to\infty}\frac{Z_n}{n^2} = 0\right] = 1,$$

o lo que es lo mismo

$$P\left[\left\{\omega\in\Omega \mid \left(\left|\frac{Z_n}{n^2}(\omega)\right| > \varepsilon\right) \text{ para una infinidad de valores de } n\right\}\right] = 0, \tag{2.2.14}$$

es decir que si $\Omega^1 = \left\{ \omega \in \Omega \mid \left(\left| \frac{Z_n}{n^2} (\omega) \right| \leq \varepsilon \right) \text{ para una infinidad de valores de } n \right\}$, *entonces*

$$P\left(\Omega^1\right) = 1.$$

Por (2.2.13) *y* (2.2.14) *podemos decir que para* $\omega \in (\Omega^0 \cap \Omega^1)$ *y dada* $\varepsilon > 0$, *existe* $N(\omega) \in \mathbb{N}$ *tal que*

$$\left| \frac{Z_n}{n^2} (\omega) \right| < \frac{\varepsilon}{2} \quad y \quad |Y_{n^2}(\omega)| < \frac{\varepsilon}{2}$$

para $n \geq N(\omega)$. *Y por* (2.2.12) $|Y_k(\omega)| < \varepsilon$. *Así que*

$$\lim_{k \to \infty} Y_k = 0,$$

por lo tanto

$$P\left(\lim_{n\to\infty} Y_n = 0\right) \geq P\left(\Omega^0 \cap \Omega^1\right) = 1 - P\left(\Omega^0 \cap \Omega^1\right)^C = 1. \qquad \blacksquare$$

Antes de ver el resultado al que llegó Kolmogorov veamos otro resultado que nos será de gran ayuda.

Proposición 2.2.6 (Condición de Kolmogorov) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes de esperanza nula, y*

$$\sum_n \frac{E\left[X_n^2\right]}{n^2} < \infty,$$

entonces

$$\frac{1}{n} \sum_{j=1}^{n} X_j \xrightarrow{cs} 0.$$

Demostración. *Sea* $S_n = \sum_{i=1}^n X_i$ *entonces*

$$E\left[S_n\right] = 0,$$

$$var\left(S_n\right) = \frac{1}{n^2}\sum_{j=1}^{n}\sigma_j^2 = \frac{1}{n^2}\sum_{j=1}^{n} E\left[X_j^2\right],$$

y para cada $\varepsilon > 0$, *sea*

$$A_\varepsilon = \left\{\omega \in \Omega : \left|\frac{S_n(\omega)}{n}\right| > \varepsilon \text{ para una infinidad de valores de } n\right\},$$

entonces por (1.3.9) *bastará probar que* $P\left(A_\varepsilon\right) = 0$, *para lo cual definamos*

$$B_{n,\varepsilon} = \left\{\omega \in \Omega : \left|\frac{S_k(\omega)}{k}\right| > \varepsilon \text{ para alguna } k \in \mathbb{N} \text{ tal que } 2^{n-1} < k \leq 2^n\right\},$$

así

$$A_\varepsilon = \left\{\omega \in \Omega : \omega \in B_{n,\varepsilon} \text{ para una infinidad de valores de } n\right\}.$$

Entonces para probar que $P\left(A_\varepsilon\right) = 0$, *por* $(1.3.8)$, *basta con probar que*

$\sum_{n=1}^{\infty} P\left(B_{n,\varepsilon}\right) < \infty$,

$$\begin{aligned}
P\left(B_{n,\varepsilon}\right) &= P\left[\max_{2^{n-1}<k\leq 2^n}\left|\frac{S_k(\omega)}{k}\right| > \varepsilon\right] = P\left[\max_{2^{n-1}<k\leq 2^n}\left|S_k(\omega)\right| > \varepsilon k\right] \\
&\leq P\left[\max_{2^{n-1}<k\leq 2^n}\left|S_k(\omega)\right| > \varepsilon 2^{n-1}\right] \leq \frac{Var\left(S_{2^n}\right)}{\varepsilon^2 2^{2n-2}} \\
&= \frac{4}{\varepsilon^2 2^{2n}}\sum_{i=1}^{2^n} E\left[X_i^2\right],
\end{aligned}$$

por lo tanto

$$\begin{aligned}
\sum_{n=1}^{\infty} P\left(B_{n,\varepsilon}\right) &\leq \sum_{n=1}^{\infty} \frac{4}{\varepsilon^2 2^{2n}} \sum_{i=1}^{2^n} E\left[X_i^2\right] = \frac{4}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{2^{2n}} \sum_{i=1}^{2^n} E\left[X_i^2\right] \\
&= \frac{4}{\varepsilon^2} \left[\begin{array}{c}
\frac{1}{2^2} E\left[X_1^2\right] + \frac{1}{2^2} E\left[X_2^2\right] \\
+\frac{1}{2^4} E\left[X_1^2\right] + \frac{1}{2^4} E\left[X_2^2\right] + \frac{1}{2^4} E\left[X_3^2\right] + \frac{1}{2^4} E\left[X_4^2\right] \\
+\frac{1}{2^8} E\left[X_1^2\right] + \frac{1}{2^8} E\left[X_2^2\right] + \frac{1}{2^8} E\left[X_3^2\right] + \frac{1}{2^8} E\left[X_4^2\right] + ... \\
+........
\end{array} \right] \\
&= \frac{4}{\varepsilon^2} \left[\begin{array}{c}
E\left[X_1^2\right] \left(\frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^8} + ...\right) \\
+E\left[X_2^2\right] \left(\frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^8} + ...\right) \\
+E\left[X_3^2\right] \left(\frac{1}{2^4} + \frac{1}{2^8} + \frac{1}{2^{16}} + ...\right) \\
+E\left[X_4^2\right] \left(\frac{1}{2^4} + \frac{1}{2^8} + \frac{1}{2^{16}} + ...\right) \\
+.......
\end{array} \right] \\
&= \frac{4}{\varepsilon^2} \sum_{i=1}^{\infty} E\left[X_i^2\right] \cdot \sum_{\{n \in \mathbb{N}: 2^n \geq i\}} \frac{1}{2^{2n}}.
\end{aligned}$$

Sea ahora n_0 *el más pequeño número natural tal que* $i \leq 2^{n_0}$, *entonces*

$$\sum_{\{n \in \mathbb{N}: 2^n \geq i\}} \frac{1}{2^{2n}} = \sum_{n=n_0}^{\infty} \frac{1}{2^{2n}} = \frac{4}{2^{2n_0}} \leq \frac{4}{i^2},$$

lo que implica que

$$\sum_{n=1}^{\infty} P\left(B_{n,\varepsilon}\right) \leq \frac{4}{\varepsilon^2} \sum_{i=1}^{\infty} E\left[X_i^2\right] \frac{4}{i^2} = \frac{16}{\varepsilon^2} \sum_{i=1}^{\infty} \frac{E\left[X_i^2\right]}{i^2} < \infty,$$

por lo que

$$\frac{1}{n} \sum_{j=1}^{n} X_j \xrightarrow{cs} 0. \qquad \blacksquare$$

Este resultado es válido con la hipótesis $\sum_n \dfrac{E\left[X_n^p\right]}{n^p} < \infty$ para $p \in \mathbb{N} \setminus \{1\}$, [1, *Chung Cap.*5.4]

Teorema 2.2.7 (Ley fuerte de los grandes números-Kolmogorov) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes e idénticamente distribuidas*

$$\textit{Si } E\left[|X_1|\right] < \infty, \; \textit{entonces} \qquad \frac{X_1 + X_2 + \cdots + X_n}{n} \overset{cs}{\rightarrow} \mu,$$

donde μ *es la esperanza común de* $X_1, X_2, X_3, \ldots\ldots$

Demostración. *En el caso en que* $\mu < \infty$. *Definamos para* $n \in \mathbb{N}$

$$Y_n = \begin{cases} X_n & si \;\; |X_n| \leq n, \\ 0 & si \;\; |X_n| > n, \end{cases}$$

$$\sum_n P\left(X_n \neq Y_n\right) = \sum_n P\left(|X_n| > n\right) = \sum_n P\left(|X_1| > n\right) \leq E\left[|X_1|\right] < \infty,$$

usando (2.2.6) *para* $\{Y_n - E\left(Y_n\right)\}$

$$\begin{aligned} \sum_n \frac{E\left[|Y_n - E\left(Y_n\right)|^2\right]}{n^2} &= \sum_n \frac{\sigma^2\left(Y_n\right)}{n^2} \leq \sum_n \frac{E\left[Y_n^2\right]}{n^2} = \sum_n \frac{1}{n^2} E\left[X_n^2 I_{[|X_n| \leq n]}\right] \\ &= \sum_n \frac{1}{n^2} \sum_{j=1}^{n} E\left[X_n^2 I_{[j-1 < |X_n| \leq j]}\right] \\ &= E\left[X_1^2 I_{[0 < |X_1| \leq 1]}\right] \qquad (2.2.15) \\ &\quad + \frac{1}{2^2} E\left[X_2^2 I_{[0 < |X_2| \leq 1]}\right] + \frac{1}{2^2} E\left[X_2^2 I_{[1 < |X_2| \leq 2]}\right] \\ &\quad + \frac{1}{3^2} E\left[X_3^2 I_{[0 < |X_3| \leq 1]}\right] + \frac{1}{3^2} E\left[X_2^2 I_{[1 < |X_3| \leq 2]}\right] + \ldots\ldots, \end{aligned}$$

Como las v.a. de la sucesión $\{X_n\}$ son idénticamente distribuidas, entonces

$$\begin{aligned} 2.2.15 &= E\left[X_1^2 I_{[0<|X_1|\leq 1]}\right]\left(1+\frac{1}{2^2}+\frac{1}{3^3}+....\right) \\ &\quad +E\left[X_2^2 I_{[1<|X_1|\leq 2]}\right]\left(\frac{1}{2^2}+\frac{1}{3^3}+\frac{1}{4^2}+....\right)+...... \\ &= \sum_{j=1}^{\infty} E\left[X_j^2 I_{[j-1<|X_j|\leq j]}\right]\cdot\sum_{n=j}^{\infty}\frac{1}{n^2}. \end{aligned} \tag{2.2.16}$$

Ahora bien

$$\sum_{n=j}^{\infty}\frac{1}{n^2}\leq\int_{j-1}^{\infty}\frac{1}{x^2}dx=\frac{1}{j-1}\leq\frac{2}{j}\qquad \textit{para}\ \ j\geq 2,$$

y como

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=1+\sum_{n=2}^{\infty}\frac{1}{n^2}\leq 1+\frac{2}{2}=2,$$

entonces

$$\sum_{n=j}^{\infty}\frac{1}{n^2}\leq 2\qquad \textit{para toda } j.$$

Por otro lado

$$E\left[X_j^2 I_{[j-1<|X_j|\leq j]}\right]\leq jE\left[|X_j|\, I_{[j-1<|X_j|\leq j]}\right]=jE\left[|X_i|\, I_{[j-1<|X_i|\leq j]}\right],$$

para toda $i=\{1,2,3,....\}$, en particular para $i=1$. Por lo que

$$2.2.16\leq\sum_{j=1}^{\infty} jE\left[|X_1|\, I_{[j-1<|X_1|\leq j]}\right]\frac{2}{j}=2\sum_{j=1}^{\infty}E\left[|X_1|\, I_{[j-1<|X_1|\leq j]}\right]=2E\left[X_1\right]<\infty,$$

por lo tanto

$$\frac{1}{n}\sum_{j=1}^{n}\left(Y_j-E\left[Y_j\right]\right)\xrightarrow{c.s}0,$$

o bien

$$\frac{1}{n}\sum_{j=1}^{n}Y_j=\frac{1}{n}\sum_{j=1}^{n}X_j\xrightarrow{c.s}E\left[Y_j\right]. \tag{2.2.17}$$

Sólo falta ver $E[Y_j] \xrightarrow{c.s} \mu$ *o lo que es lo mismo* $\lim_{n\to\infty} E[Y_n] = \mu$

$$\begin{aligned} E[Y_n] &= E\left[Y_n I_{[|X_n|\leq n]}\right] + E\left[Y_n I_{[|X_n|>n]}\right] \\ &= E\left[X_n I_{[|X_n|\leq n]}\right] + E\left[0 I_{[|X_n|>n]}\right] \\ &= E\left[X_1 I_{[|X_n|\leq n]}\right] = E\left[X_n I_{[-n\leq X_n\leq n]}\right] \\ &= \int_{-n}^{n} x f_{X_1}(x)\, dx, \end{aligned}$$

así

$$\lim_{n\to\infty} E[Y_n] = \lim_{n\to\infty} \int_{-n}^{n} x f_{X_1}(x)\, dx = E[X_1] = \mu,$$

de donde se obtiene

$$\frac{1}{n}\sum_{j=1}^{n} X_j \xrightarrow{c.s} E[Y_j] \xrightarrow{c.s} \mu,$$

por (1.3.4) *obtenemos que*

$$\frac{1}{n}\textstyle\sum_{j=1}^{n} X_j \xrightarrow{c.s} \mu \;. \qquad \blacksquare$$

En general podemos decir que si se cumplela condición de Markov, es decir

$$\lim_{n\to\infty} \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2 = 0$$

entonces se cumple la Ley Débil de los Grandes Números (aunque el recíproco no es cierto). Y análogamente si se cumple la condición de Kolmogorov, es decir si

$$\sum_n \frac{E[X_n^p]}{n^p} < \infty \text{ para } p \in \mathbb{N}\setminus\{1\},$$

entonces se cumple la Ley Fuerte de los Grande Números.

Ejemplo 2.2.8 *Sea* $\{X_n\}$ *una sucesión de variables aleatorias tales que*

$$X_n = \begin{cases} \sqrt{n} & \\ & \text{con probabilidad } \frac{1}{n^2} \\ -\sqrt{n} & \\ & \\ 0 & \text{con probabilidad } 1 - \frac{2}{n^2} \end{cases}$$

En este caso $Var\left[X_n\right] = E\left[X_n^2\right] = \frac{2}{n}$ *y*

$$\sum_n \frac{E\left[X_n^2\right]}{n^2} = \sum_n \frac{2}{n^3} < \infty.$$

Entonces cumple la condición de Kolmogorov, por lo que se afirma que cumple la Ley de los Grande Números "si cumple la ley fuerte cumple la débil".

CAPÍTULO 3

EL TEOREMA CENTRAL DEL LÍMITE

3.1 Introducción

Cuando las variables aleatorios tienen una varianza finita, *el teorema central del límite* muestra que la distribución de la diferencia estandarizada entre la suma de variables aleatorias y el valor esperado de esta suma tiende a una distribución gaussiana cuando la cantidad de variables es muy grande.

La primer versión (el origen) del *Teorema central del límite* fue dada por De Moivre-Laplace, usando variables aleatorias independientes e idénticamente distribuidas como bernoullis.

Y el resultado actual se dabe gracias a Lindeberg, que generaliza para cualquier sucesión de variables aleatorias.

Los libros utilizados para el desarrollo de este capítulo fueron [1, Chung], [3, Feller], [5, García, Miguel, II], [6, Gnedenko], [9, Hernández, Fabian], [10, Carrasco, Guadalupe], [11, Rincón, Luis].

3.2 El Teorema Central del Límite De Moivre-Laplace

Teorema 3.2.1 (Teorema local de De Moivre–Laplace) *Sea $\{X_n\}$ una sucesión de variables aleatorias Bernoullis con parámetro p. Dados dos números reales a y b, $a < b$. Para cada entero k que cumple que $np + a\sqrt{npq} \leq k \leq np + b\sqrt{npq}$ definimos el número $y_{k,n} = \frac{k-np}{\sqrt{npq}}$. Entonces se tiene que*

$$\lim_{n\to\infty} \frac{P\left(X_1 + X_2 + \cdots + X_n = k\right)}{\frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-y_{n,k}^2/2}} = 1 \qquad \textit{uniformemente en } k.$$

Demostración. La fórmula de Stirling establece que, para $m \in \mathbb{N}$, se cumple que $m! = \sqrt{2\pi} m^m e^{-m} \sqrt{m} e^{\theta_m}$, donde $\frac{1}{12m+1} < \theta_m < \frac{1}{12m}$ (una demostración a esta fórmula la podemos encontrar el el libro de Feller 1965). $S_n = X_1 + X_2 + \cdots + X_n$, es una variable aleatoria con función de distribución binomial de parámetros n y p.

$$\begin{aligned}
P\left(S_n = k\right) &= \binom{n}{k} p^k \left(1-p\right)^{n-k} = \frac{n!}{k!\left(n-k\right)!} p^k \left(1-p\right)^{n-k} \\
&= \frac{\sqrt{2\pi} n^n e^{-n} \sqrt{n} e^{\theta_n}}{\sqrt{2\pi} k^k e^{-k} \sqrt{k} e^{\theta_k} \sqrt{2\pi}\left(n-k\right)^{n-k} e^{-(n-k)} \sqrt{n-k} e^{\theta_{n-k}}} p^k \left(1-p\right)^{n-k} \\
&= \frac{n^n \sqrt{n} e^{\theta_n - \theta_k - \theta_{n-k}}}{\sqrt{2\pi} k^k \sqrt{k} \left(n-k\right)^{n-k} \sqrt{n-k}} p^k \left(1-p\right)^{n-k} \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{npq}} \sqrt{\frac{np}{k}} \sqrt{\frac{nq}{n-k}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} e^{\theta_n - \theta_k - \theta_{n-k}}.
\end{aligned}$$

Ahora bien, si k satisface que $np + a\sqrt{npq} \leq k \leq np + b\sqrt{npq}$, entonces

$$\begin{aligned}
n - np - b\sqrt{npq} &\leq n - k \leq n - np - a\sqrt{npq}, \\
nq - b\sqrt{npq} &\leq n - k \leq nq - a\sqrt{npq}.
\end{aligned}$$

De las desigualdades anteriores se desprende que

$$\begin{aligned}
\frac{p}{p+b\sqrt{\frac{pq}{n}}} &= \frac{np}{np+b\sqrt{npq}} \leq \frac{np}{k} \leq \frac{np}{np+a\sqrt{npq}} = \frac{p}{p+a\sqrt{\frac{pq}{n}}},\\
\frac{p}{q-a\sqrt{\frac{pq}{n}}} &= \frac{np}{nq-a\sqrt{npq}} \leq \frac{np}{n-k} \leq \frac{np}{nq-b\sqrt{npq}} = \frac{p}{q-b\sqrt{\frac{pq}{n}}},
\end{aligned}$$

$$\begin{aligned}
\theta_n-\theta_k-\theta_{n-k} &\leq \theta_n+\theta_k+\theta_{n-k} \leq \frac{1}{12}\left(\frac{1}{n}+\frac{1}{k}+\frac{1}{n-k}\right)\\
&\leq \frac{1}{12}\left(\frac{1}{n}+\frac{1}{np+a\sqrt{npq}}+\frac{1}{nq-b\sqrt{npq}}\right)\\
\frac{1}{12n+1} &\leq \theta_n-\theta_k-\theta_{n-k},
\end{aligned}$$

por tanto,

$$\lim_{n\to\infty}\sqrt{\frac{np}{k}}\sqrt{\frac{nq}{n-k}}e^{\theta_n-\theta_k-\theta_{n-k}} = 1 \qquad \text{uniformemente en } k.$$

y

$$\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k} = \frac{1}{\left(\frac{k}{np}\right)^k\left(\frac{n-k}{nq}\right)^{n-k}}\ , \tag{3.2.1}$$

tomando $y_{k,n}=\frac{k-np}{\sqrt{npq}}$

$$\begin{aligned}
\left(\frac{k}{np}\right)^k &= \left(\frac{np+k-np\left(\frac{\sqrt{npq}}{\sqrt{npq}}\right)}{np}\right)^k\\
&= \left(\frac{np+y_{k,n}\sqrt{npq}}{np}\right)^k\\
&= \left(1+y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{np+y_{k,n}\sqrt{npq}},
\end{aligned}$$

y

$$
\begin{aligned}
\left(\frac{n-k}{nq}\right)^{n-k} &= \left(\frac{n-k+np-np\left(\frac{\sqrt{npq}}{\sqrt{npq}}\right)}{nq}\right)^{n-k} \\
&= \left(\frac{nq-(k-np)\left(\frac{\sqrt{npq}}{\sqrt{npq}}\right)}{nq}\right)^{n-k} \\
&= \left(\frac{nq-y_{k,n}\sqrt{npq}}{nq}\right)^{n-k} \\
&= \left(1-y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{nq-y_{k,n}\sqrt{npq}},
\end{aligned}
$$

por lo que

$$
\begin{aligned}
\log(3.2.1) &= -\left(np+y_{k,n}\sqrt{npq}\right)\log\left(1+y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right) \qquad (3.2.2)\\
&\quad -\left(nq-y_{k,n}\sqrt{npq}\right)\log\left(1-y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right),
\end{aligned}
$$

y si usamos la expansión de Taylor en el logaritmo, alrededor del 1 (el primer término se elimina, al ser $\log 1 = 0$), se tiene

$$\begin{aligned}
3.2.2 \quad &= -\left(np + y_{k,n}\sqrt{npq}\right)\left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}(k-1)!}{k!}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{k}\right) \\
&\quad -\left(nq - y_{k,n}\sqrt{npq}\right)\left(\sum_{k=1}^{\infty} \frac{(-1)(k-1)!}{k!}\left(-y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{k}\right) \\
&= -\left(np + y_{k,n}\sqrt{npq}\right)\left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{k}\right) \\
&\quad +\left(nq - y_{k,n}\sqrt{npq}\right)\left(\sum_{k=1}^{\infty} \frac{1}{k}\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{k}\right) \\
&= -np\left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{k}\right) - y_{k,n}\sqrt{npq}\left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{k}\right) \\
&\quad +nq\left(\sum_{k=1}^{\infty} \frac{1}{k}\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{k}\right) - y_{k,n}\sqrt{npq}\left(\sum_{k=1}^{\infty} \frac{1}{k}\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{k}\right) \\
&= -np\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right) + \frac{np}{2}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{2} - \frac{np}{3}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{3} + \frac{np}{4}\left(y_{k,n}\frac{\sqrt{q}}{\sqrt{np}}\right)^{4} - \ldots \\
&\quad -\sqrt{npq}\,y_{k,n}^{2}\left(\frac{\sqrt{q}}{\sqrt{np}}\right) + \frac{\sqrt{npq}\,y_{k,n}^{3}}{2}\left(\frac{\sqrt{q}}{\sqrt{np}}\right)^{2} - \frac{\sqrt{npq}\,y_{k,n}^{4}}{3}\left(\frac{\sqrt{q}}{\sqrt{np}}\right)^{3} + \ldots \\
&\quad +nq\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right) + \frac{nq}{2}\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{2} + \frac{nq}{3}\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{3} + \frac{nq}{4}\left(y_{k,n}\frac{\sqrt{p}}{\sqrt{nq}}\right)^{4} + \ldots \\
&\quad -\sqrt{npq}\,y_{k,n}^{2}\left(\frac{\sqrt{p}}{\sqrt{nq}}\right) - \frac{\sqrt{npq}\,y_{k,n}^{3}}{2}\left(\frac{\sqrt{p}}{\sqrt{nq}}\right)^{2} - \frac{\sqrt{npq}\,y_{k,n}^{4}}{3}\left(\frac{\sqrt{p}}{\sqrt{nq}}\right)^{3} - \ldots \\
&= y_{k,n}^{2}\left(\frac{q}{2} - q + \frac{p}{2} - p\right) \\
&\quad +y_{k,n}^{3}\left(-\frac{\sqrt{q}^{3}}{3\sqrt{np}} + \frac{\sqrt{q}^{3}}{2\sqrt{np}} + \frac{\sqrt{p}^{3}}{3\sqrt{nq}} - \frac{\sqrt{p}^{3}}{2\sqrt{nq}}\right) \\
&\quad +y_{k,n}^{4}\left(\frac{q^{2}}{4np} - \frac{q^{2}}{3np} + \frac{p^{2}}{4nq} - \frac{p^{2}}{3nq}\right) + \ldots\ldots \\
&= -\frac{y_{k,n}^{2}}{2} + \frac{y_{k,n}^{3}}{6\sqrt{npq}}\left(q^{2} - p^{2}\right) + \frac{y_{k,n}^{4}}{12npq}\left(-q^{3} - p^{3}\right) + \ldots.. \\
&= -\frac{y_{k,n}^{2}}{2} + A_{n}(y_{k,n}),
\end{aligned}$$

donde $A_n(y_{k,n}) = \frac{y_{k,n}^3}{6\sqrt{npq}}(q^2 - p^2) + \frac{y_{k,n}^4}{12npq}(-q^3 - p^3) + ...,$ por lo que

$$A_n(y_{k,n}) \underset{n\to\infty}{\longrightarrow} 0 \text{ uniformemente en } k,$$

entonces

$$\lim_{n\to\infty} \exp(A_n(y_{k,n})) = 1 \text{ uniformemente en } k.$$

así que

$$\begin{aligned}
&\frac{P(S_n = k)}{\frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-y_{n,k}^2/2}} \\
&= \frac{\sqrt{2\pi}\sqrt{npq} e^{y_{n,k}^2/2}}{\sqrt{2\pi}\sqrt{npq}} \sqrt{\frac{np}{k}} \sqrt{\frac{nq}{n-k}} e^{\theta_n - \theta_k - \theta_{n-k}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \\
&= e^{y_{n,k}^2/2} \sqrt{\frac{np}{k}} \sqrt{\frac{nq}{n-k}} e^{\theta_n - \theta_k - \theta_{n-k}} e^{-y_{n,k}^2/2} e^{A_n(y_{k,n})} \\
&= \sqrt{\frac{np}{k}} \sqrt{\frac{nq}{n-k}} e^{\theta_n - \theta_k - \theta_{n-k}} e^{A_n(y_{k,n})} \underset{n\to\infty}{\longrightarrow} 1 \qquad \text{uniformemente.}
\end{aligned}$$

Es decir, en el límite, la variable se distribuye como normal de parámetros $\mu = np$ y $\sigma^2 = npq$. ■

Ejemplo 3.2.2 *Un productor de focos sabe que el 15% de su producción esta defectuosa. En su promoción de ventas, asegura que en cada envío de 10,000 focos no habrá más de α defectuosos y se compromete a devolver el costo total de la entrega si esa cantidad es rebasada. ¿De qué tamaño debe ser α para que en promedio, no tenga que pagar más del 1% de los envíos?*

Se trata de una Binomial donde $n = 10,000$ y $p = 0.05$.

Buscamos α *tal que* $p(X > \alpha) \leq 0.01$. *Usando la aproximación normal, tenemos*

$$\begin{aligned} P(X > \alpha) &\approx P\left(\frac{X - np}{\sqrt{npq}} > \frac{\alpha - np}{\sqrt{npq}}\right) = P\left(Z > \frac{\alpha - 500}{\sqrt{475}}\right) \\ &= 1 - P\left(Z \leq \frac{\alpha - 500}{\sqrt{475}}\right) = 0.01 \end{aligned}$$

Así que buscamos α *tal que*

$$P\left(Z \leq \frac{a - 500}{\sqrt{475}}\right) = 0.99 \Rightarrow \frac{a - 500}{\sqrt{475}} \approx 2.33$$

De donde $\alpha \approx 550.78$. *De manera que si asegura que en 10,000 focos no habrá más de 551 defectuosos, en promedio, tendrá que pagar a lo más el 1% de los envíos.*

Para la demostración del Teorema Central del Límite Generalizado para cualquier sucesión de variables aleatorias independientes e idénticamente distribuidas necesitamos de la definición y algunas propiedades de la función característica. Por lo que antes de la demostración, desarrollaremos algunos conceptos al respecto

3.3 Función Característica

Las dos funciones generatrices más comunes son: la Función Generatriz de Momentos y la Función Característica. La primera nos ayuda a calcular momentos y la segunda nos ayuda a caracterizar de manera *única* a las distribuciones de probabilidad, y es esta caracterización única la que nos va a servir para demostrar el Teorema Central del Límite.

Definición 3.3.1 *Sea* X *una v.a. La Función Característica de* X *denotada por*

$\Phi_X(\cdot)$, *está definida para toda* $t \in \mathbb{R}$ *como*

$$\Phi_X(t) = E\left[e^{itX}\right].$$

Otra propiedad importante de esta función es que, a diferencia de la de momentos, siempre existe.

Lema 3.3.2 *Sea* X *v.a. discreta o absolutamente continua, entonces* $E\left[e^{itx}\right]$ *existe para toda* $t \in \mathbb{R}$.

Demostración. *Para que* $E\left[e^{itX}\right]$ *exista es necesario que* $\left|E\left[e^{itX}\right]\right| < \infty$ *y* $\left|e^{itx}\right| = \left(\cos^2 tx + sen^2 tx\right)^{1/2} = 1$. *Entonces*

$$\left|E\left[e^{itX}\right]\right| \leq E\left[\left|e^{itX}\right|\right] = 1. \quad \blacksquare$$

Proposición 3.3.3 *Sea* X *v.a. con función característica dada por* $\Phi_X(\cdot)$ *entonces si* $E\left[X^r\right]$ *existe para algún* $r \in \mathbb{N}$, *entonces*

$$\Phi_X(t) \underset{t \to 0}{\to} 1 + \sum_{k=1}^{r} E\left[X^k\right] \frac{(it)^k}{k!} + R\left(t^r\right),$$

donde $R\left(t^r\right)$ *es el residuo de la expanción y representa una función que converge a cero más rápido que* t^r *cuando* $t \to 0$ *esto es*

$$\lim_{t \to 0} \frac{R\left(t^r\right)}{t^r} = 0.$$

La idea de la demostración es utilizando la expanción de MacLaurin para e^x. Esta proposición se puede encontrar el el libro [9, Hérnández Fabian Cap. 5.2].

Teorema 3.3.4 (Inversión de Lévy) *Sea X v.a. con función de densidad $f_X(x)$ y función caranterítica $\Phi_X(t)$. Si $a<b$, entonces*

$$P[a<X<b]+\frac{1}{2}P[X=a]+\frac{1}{2}P[X=b]=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\Phi_X(t)\,dt.$$

Demostración.

$$\begin{aligned}
M_T &= \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\Phi_x(t)\,dt \\
&= \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}E\left[e^{itx}\right]dt \\
&= \frac{1}{2\pi}\int_{-T}^{T}\int_{X}\frac{e^{-ita}-e^{-itb}}{it}e^{itx}f(x)\,dxdt \\
&= \int_{-\infty}^{\infty}\left[\int_{-T}^{T}\frac{e^{it(x-a)}-e^{it(x-b)}}{2\pi it}dt\right]f(x)\,dx \\
&= \int_{-\infty}^{\infty}\int_{0}^{T}\frac{e^{it(x-a)}-e^{-it(x-a)}-\left(e^{it(x-b)}-e^{-it(x-b)}\right)}{2\pi it}dtf(x)\,dx \\
&= \int_{-\infty}^{\infty}\left[\int_{0}^{T}\frac{sen\ t(x-a)-sen\ t(x-b)}{\pi t}dt\right]f(x)\,dx.
\end{aligned}$$

Sea $I_{a,b,T}(x)=\left[\int_0^T\frac{sen\ t(x-a)-sen\ t(x-b)}{\pi t}dt\right]$. *Ahora bien*

$$\frac{1}{\pi}\int_0^T\frac{sent\alpha}{t}dt\underset{T\to\infty}{\to}\begin{cases}\frac{1}{2} & si\ \alpha>0,\\ -\frac{1}{2} & si\ \alpha<0,\end{cases}$$

por lo que

$$\lim_{T\to\infty}I_{a,b,T}(x)=\begin{cases}-\frac{1}{2}-\frac{1}{2}=0 & si\quad x<a,\\ 0-\left(-\frac{1}{2}\right)=\frac{1}{2} & si\quad x=a,\\ \frac{1}{2}-\left(-\frac{1}{2}\right)=1 & si\quad a<x\quad<b,\\ \frac{1}{2}-0=\frac{1}{2} & si\quad x=b,\\ \frac{1}{2}-\frac{1}{2}=0 & si\quad x>b,\end{cases}$$

entonces

$$\lim_{T\to\infty} M_T = 0 + \int_{\{x:x=a\}} \frac{1}{2}(x) f(x)\,dx + \int_{\{x:x\in(a,b)\}} f(x)\,dx + \int_{\{x:x=b\}} \frac{1}{2} f(x)\,dx + 0$$

$$= P[a < X < b] + \tfrac{1}{2} P[X = a] + \tfrac{1}{2} P[X = b]. \qquad \blacksquare$$

Observamos que en caso absolutamente continuo $\frac{1}{2}P[X=a]+\frac{1}{2}P[X=b]=0$.

Teorema 3.3.5 (Unicidad de Lévy) *Sean X, Y v.a. con funciones de distribución F_x, F_Y y funciones carecterísticas Φ_X y Φ_Y respectivamente. Entonces*

$$\Phi_X = \Phi_Y \Leftrightarrow F_x = F_Y.$$

Demostración. *Del teorema de inversión de Lévy se tiene que*

$$\begin{aligned}
P[a < X \le x] + \frac{1}{2}P[X=a] - \frac{1}{2}P[X=x] &= F_X(x) - F_X(a) + \frac{1}{2}P[X=a] \\
&\quad -\frac{1}{2}P[X=x] \\
&= F_Y(x) - F_Y(a) + \frac{1}{2}P[Y=a] \\
&\quad -\frac{1}{2}P[Y=x],
\end{aligned}$$

haciendo $a \to -\infty$ tenemos

$$F_X(x) = F_Y(x),$$

para todos los valores de continuidad de ambas funciones de distribución. De donde se obtiene $\Phi_X = \Phi_Y \Rightarrow F_x = F_Y$. Ahora bien si $F_x = F_Y$, entonces

$$\Phi_X(t) = E\left[e^{itx}\right] = \int e^{itx} f_X(x)\,dx = \int e^{itx} f_Y(x)\,dx = \Phi_Y(t),$$

por lo tanto

$$\Phi_X = \Phi_Y \Leftrightarrow F_x = F_Y. \quad \blacksquare$$

El siguiente teorema es en el que se basa la demostración del Teorema Central del Límite, aquí desarrollada, pero su resultado necesitaba de los anteriores.

Teorema 3.3.6 (Continuidad de Lévy-Crámer) *Sea $\Phi_{X_n}(\cdot)$ función característica de X_n para $n \in \mathbb{N}$. Si $X_n \xrightarrow{D} X$ entonces $\Phi_{X_n}(t) \to \Phi_X(t)$ para toda $t \in R$ donde $\Phi_X(\cdot)$ es la función característica de X. Recíprocamente si $\Phi_{X_n}(t) \to \Phi(t)$ y la función límite es continua en $t = 0$ entonces, $X_n \xrightarrow{D} X$ y $\Phi(\cdot)$ es la función característica de X.*

Demostración. $\Rightarrow$)

$$\begin{aligned}
\lim_{n\to\infty} \Phi_{X_n}(t) &= \lim_{n\to\infty} \left(E[\cos tX_n] + iE[sen\ tX_n]\right) \\
&= E[\cos tX] + iE[sen\ tX] = \Phi_X(t).
\end{aligned}$$

$\Leftarrow$)*Para dos puntos de continuidad $a < x$ comunes a cada F_{X_n} y F_X, el teorema de inversión de Lévy establece que*

$$\begin{aligned}
F_X(x) - F_X(a) &= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itx}}{it} \Phi_X(t)\, dt \\
&= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itx}}{it} \left(\lim_{n\to\infty} \Phi_{X_n}(t)\right) dt \\
&= \lim_{n\to\infty} \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itx}}{it} \Phi_{X_n}(t)\, dt \\
&= \lim_{n\to\infty} \left(F_{X_n}(x) - F_{X_n}(a)\right),
\end{aligned}$$

haciendo $a \to -\infty$ se obtiene $X_n \xrightarrow{D} X$. $\blacksquare$

Gracias a este resultado es muy sencillo demostrar el Teorema Central del Límite de De Moivre_Laplace. Para verlo basta seguir la siguiente demostración, generalizada para cualquier sucesión de variables aleatorias independientes e idénticamente distribuidas.

3.4 Teorema Central del Límite

Teorema 3.4.1 (Teorema Central del Límite) *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con media μ y varianza finita σ^2. Entonces*

$$\lim_{n\to\infty} P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\ y^2/2} dy.$$

Demostración. *Sea $S_i = \frac{(X_i - \mu)}{\sigma}$ y $W_n = \sum_{i=1}^{n} \frac{(X_i-\mu)}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_i$. Con lo cual*

$$\begin{aligned} E\left[S_i\right] &= 0, \\ Var\left[S_i\right] &= 1, \end{aligned}$$

entonces

$$\begin{aligned} \Phi_{W_n}(t) &= E\left[e^{itw}\right] = E\left[e^{\frac{it}{\sqrt{n}} \sum_{i=1}^{n} S_i}\right] = \Pi_{i=1}^{n} \Phi_{S_i}\left(\frac{t}{\sqrt{n}}\right) \\ &= \left[\Phi_{S_i}\left(\frac{t}{\sqrt{n}}\right)\right]^n, \end{aligned}$$

y $\frac{t}{\sqrt{n}} \underset{n\to\infty}{\longrightarrow} 0$ *utilizando 3.3.3 se tiene*

$$\begin{aligned}
\Phi_{W_n}(t) &= \left[1 + \frac{it}{\sqrt{n}} E\left[S_1\right] + \frac{(it)^2}{2n} E\left[S_1^2\right] + R\left(\frac{t^2}{n}\right)\right]^n \\
&= \left[1 - \frac{t^2}{2n} + R\left(\frac{t^2}{n}\right)\right]^n \\
&= \left[1 - \frac{a_n(t)}{n}\right]^n,
\end{aligned}$$

donde $a_n(t) = \left(\frac{t^2}{2}\right)\left[1 - \frac{2R\left(\frac{t^2}{n}\right)}{t^2/n}\right]$. *Cuando* $n \to \infty$ *se tiene que* $\frac{R\left(\frac{t^2}{n}\right)}{t^2/n} \to 0$, *entonces*

$$\lim_{n\to\infty} \Phi_{W_n}(t) = e^{\frac{-t^2}{2}}.$$

Y si Z, *una v.a., se distribuye como una normal de parámetros* $\mu = 0$ *y* $\sigma^2 = 1$, *se tiene que*

$$\Phi_Z(t) = E\left[e^{itz}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itz - \frac{z^2}{2}} dz,$$

haciendo $v = z - it$ *se tiene que* $v^2 = z^2 - 2itz - t^2$. *Entonces*

$$\Phi_Z(t) = e^{\frac{-t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-v^2}{2}} dz = e^{\frac{-t^2}{2}}.$$

Por lo que $W_n \xrightarrow{D} Z$. ■

Ejemplo 3.4.2 *En un lanzamiento de tiro al blanco, un tirador obtiene 10, 9, 8, 7, o 6 puntos con probabilidades 0.05, 0.1, 0.15, 0.2 y 0.5 respectivamente. Lanza 100 tiros. ¿Cuál es la probabilidad de que su puntaje no exceda 730?*

Solución 3.4.3 *Definamos la variable*

$$X_i = \#\ de\ puntos\ del\ tiro\ en\ el\ i - \acute{e}simo\ lanzamiento,$$

donde $x_i = 6, 7, 8, 9, 10$. *Primero calculemos* $E[X_i]$ *y* $var[X_i]$

$$\begin{aligned} E[X_i] &= \sum_{i=6}^{10} xP(X_i = x) \\ &= 6(0.5) + 7(0.2) + 8(0.15) + 9(0.1) + 10(0.05) \\ &= 7, \end{aligned}$$

$$\begin{aligned} E^2[X_i] &= \sum_{i=6}^{10} x^2P(X_i = x) \\ &= 6^2(0.5) + 7^2(0.2) + 8^2(0.15) + 9^2(0.1) + 10^2(0.05) \\ &= 50.5, \end{aligned}$$

entonces

$$var[X_i] = 50.5 - 49 = 1.5$$

Por lo que

$$\begin{aligned} P\left[\sum_{i=1}^{100} X_i > 730\right] &= 1 - P\left[\sum_{i=1}^{100} X_i \leq 730\right] \\ &= 1 - P\left[\frac{\sum_{i=1}^{100} X_i - 7(100)}{\sqrt{1.5(100)}} \leq \frac{730 - 7(100)}{\sqrt{1.5(100)}}\right] \\ &\approx 1 - P[Z \leq 2.4494] = 1 - (.5 + .4929) = 0.0071 \end{aligned}$$

Ejemplo 3.4.4 *Suponga que en una ciudad dada existen* N *votantes registrados en el padrón electoral. Suponga que el padrón está a su dispocición y que puede selecciónar, con muestreo aleatorio sin reemplazo,* n *votantes;* $1 \leq n \leq N$. *Si se desea estimar la proporción de votos declarados a favor del candidato* A *de tal manera que 0.001 sea la probabilidad de que la proporción muestral sea menor que*

0.50 cuando en realidad vale 0.52, ¿de que tamaño tiene que ser la muestra?

Primero definamos $X_i : \begin{cases} 1 & \text{si la i-ésima elección es a favor del candidato A,} \\ \\ 0 & \text{e.o.c,} \end{cases}$

tenemos que

$$P\left(X_i = 1\right) = \frac{VFA}{N} = 0.52,$$

donde VFA *es el número de votantes a favor del candidato A. Por lo que si definimos*

$$X : \#\ \text{de VFA en la muestra,}$$

X *es una variable de distribución hipergeométrica y además* $X = \sum_{i=1}^{n} X_i$*, así*

$$E\left[X\right] = 0.52n,$$

y

$$var\left(X\right) = \frac{\left(VFA\right)n\left(N - VFA\right)\left(N - n\right)}{N^2\left(N-1\right)},$$

que son la esperanza y varianza de una hipergeométrica. Reduciendo la varianza

$$\begin{aligned} var\left(X\right) &= \frac{0.52n\left(1-0.52\right)\left(N-n\right)}{\left(N-1\right)} \\ &= \frac{N-n}{N-1}0.52\left(0.48\right)n, \end{aligned}$$

(que es un resultado muy usado en muestreo simple)

$$\begin{aligned} P\left(\frac{\sum_{i=1}^{n} x_i}{n} < 0.5\right) &= P\left(Z < \frac{N-1}{N-n}\frac{0.5-0.52}{0.2496}\right) \\ &= P\left(Z < -0.080128205\frac{N-1}{N-n}\right) \\ &= 0.001, \end{aligned}$$

esto ocurre (por tablas) cuando

$$-0.080128205\frac{N-1}{N-n} = -3.08,$$

despejando n

$$n = N - (N-1)\,\frac{0.8012}{3.08},$$

N	n
$10,000$	$9,740$
$100,000$	$97,398$
$1,000,000$	$973,984$

como la proporción muestral está muy "cerca" a la real y el error que se pide es pequeño, entonces es natural que el tamaño de la muestra sea muy grande.

CONCLUSIONES

La importancia del Teorema Central del Límite es facilitar los cálculo numericos, sobre todo en la estadística, así como también es la base para muchas otras herramientas de ella. Claro que para los cálculos estadísticos pocas veces se cumplen todas las hipótesis (es común que para $n \geq 30$ se considere lo suficientemente grande para usar el T.C.L.) entonces dependerá del problema y de las hipótesis estadísticas planteadas para su resolución, para saber si es que se puede usar o no el T.C.L.

Suponga que se realizan n experimentos independientes obteniendo como resultados $X_1, X_2, ..., X_n$, los cuales son independientes e idénticamente distribuidos con media μ y varianza σ^2 $(0 < \sigma^2 < \infty)$. Esta es una situación clásica cuando queremos estimar μ y cuando los experimentos son sometidos a errores independientes, es decir, $X_n = \mu + \varepsilon_n$ donde los números $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ son variables aleatorias con media 0 y varianza σ^2.

Sabemos que el valor experimental de la media es $\overline{X}_n = (X_1 + X_2 + ... + X_n)/n$. Este valor tiene la propiedad de que $\overline{X}_n \to \mu$ (Ley Fuerte de los Grandes Números). Sin embargo $\overline{X}_n$ no es exactamente μ y se podría querer tener más información acerca de la distribución de $\overline{X}_n$. Se podría querer la estimación de $P\left[\left|\overline{X}_n - \mu\right| > a\right]$, es decir, la probabilidad de que el error sea mayor que a y si no se conoce exactamente la distribución de cada X_n alrededor de la media μ, entonces ese cálculo se vuelve imposible.

Sin embargo el Teorema Central del Límite nos puede ayudar a encontrar una estimación a esta probabilidad, esto si n es lo suficientemente grande. Podemos, entonces, observar que

$$P\left[\left|\frac{X_1 + X_2 + ... + X_n - n\mu}{\sigma\sqrt{n}}\right| \geq b\right] \approx 1 - \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} e^{-y^2/2} dy.$$

Es decir que en Muestreo Probabilístico, se supone que los estimadores $\overline{y}$ y $\overline{Y}$ se distribuyen en forma normal en torno a los parámetros que pretenden estimar. Esta suposición se basa en resultados análogos al Teorema Central del Límite, para poblaciones infinitas.

Por ejemplo, si se desea encontrar un intervalo de confianza al $100\,(1-\alpha)\,\%$ dado por

$$P\left[-Z_{(1-\alpha/2)} < Z < Z_{(1-\alpha/2)}\right] = 1 - \alpha,$$

donde Z es una variable aleatoria que se distribuye como una normal con media 0 y varianza 1, para la media poblacional $\overline{Y}$, como suponemos que $\overline{y} \sim N\left(\overline{Y}, \left(1 - \frac{n}{N}\right)\frac{S^2}{n}\right)$, (donde n es el tamaño de la muestra, N es el tamaño de la población y S^2 es la varianza de la población, que se puede reemplazar por la varianza de la muestra), tomando

$$Z = \frac{\overline{y} - \overline{Y}}{\sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}}}$$

tenemos que

$$P\left[-Z_{(1-\alpha/2)} < \frac{\overline{y} - \overline{Y}}{\sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}}} < Z_{(1-\alpha/2)}\right] = 1 - \alpha,$$

es decir,

$$P\left[\overline{y} - Z_{(1-\alpha/2)}\sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}} < \overline{Y} < \overline{y} + Z_{(1-\alpha/2)}\sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}}\right] = 1 - \alpha,$$

así los límites buscados serán

$$\overline{y} \pm Z_{(1-\alpha/2)}\sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}}.$$

Y este tipo de pruebas se realiza en toda la Estadística, constantemente, en muestreo, regresión y en muchas otras ramas lo que crea una necesidad de conocimiento del concepto y condiciones para su aplicación.

Además que los estudios sobre el Teorema Central del Límite, aún continúan, respecto a la rapidez en que convergen. Así como la variación del mismo, para aplicaciones en distintas areas.

Si desea hacer un estudio más profundo del tema es necesario por lo menos estudiar Teoría de la Medida, o Probabilidad Avanzada. Con lo cual se podrá consultar gran cantidad de bibliografía, en específico del desarrollo histórico del Teorema Central del Límite, que basa sus demostraciones en la integral de Lebesgue. Pero eso lo dejamos a consideración del lector, pues nuestro fin fue cubrir los Temas de Convergencia, de sucesiones de variables aleatorias, y Teoremas Límite, planeados para un curso de Probabilidad 2 en la Facultad de Ciencias, de la UNAM.

APPENDICES

A. Construcción de los primeros cuatro momentos para la distribución binomial $(X \sim B(n,p))$

$$
\begin{aligned}
E[X] &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} (1-p)^{n-x} \\
&= np \sum_{x=0}^{n-1} \frac{(n-1)!}{(x)!\,(n-1-x)!} p^{x} (1-p)^{n-1-x} \\
&= np \sum_{x=0}^{n-1} x \binom{n}{x} p^x (1-p)^{n-1-x} \\
&= np
\end{aligned}
$$

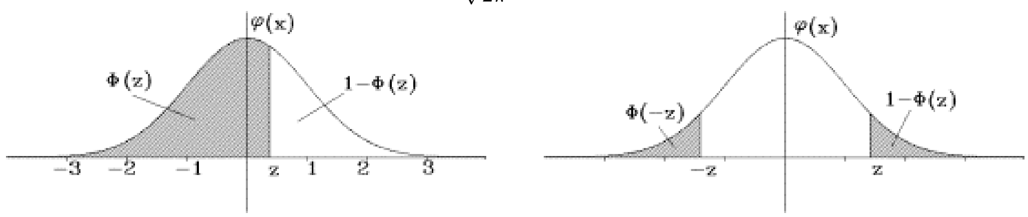
$$\begin{aligned}
E\left[X^2\right] &= E\left[X\left(X-1\right)+X\right]=E\left[X\right]+E\left[X\left(X-1\right)\right]\\
&= np+\sum_{x=0}^{n}x\left(x-1\right)\binom{n}{x}p^x\left(1-p\right)^{n-x}\\
&= np+\sum_{x=2}^{n}x\left(x-1\right)\frac{n!}{x!\left(n-x\right)!}p^x\left(1-p\right)^{n-x}\\
&= np+\sum_{x=2}^{n}\frac{n!}{\left(x-2\right)!\left(n-x\right)!}p^x\left(1-p\right)^{n-x}\\
&= np+n\left(n-1\right)p^2\sum_{x=2}^{n}\frac{\left(n-2\right)!}{\left(x-2\right)!\left(n-x\right)!}p^{x-2}\left(1-p\right)^{n-x}\\
&= np+n\left(n-1\right)p^2\sum_{x=0}^{n-2}\frac{\left(n-2\right)!}{\left(x\right)!\left(n-2-x\right)!}p^{x}\left(1-p\right)^{n-2-x}\\
&= np+n\left(n-1\right)p^2
\end{aligned}$$

$$\begin{aligned}
E\left[X^3\right] &= E\left[X\left(X-1\right)\left(X-2\right)+3X^2-2X\right]\\
&= -2E\left[X\right]+3E\left[X^2\right]+E\left[X\left(X-1\right)\left(X-2\right)\right]\\
&= np+3n\left(n-1\right)p^2+\sum_{x=0}^{n}x\left(x-1\right)\left(x-2\right)\binom{n}{x}p^x\left(1-p\right)^{n-x}\\
&= np+3n\left(n-1\right)p^2+\sum_{x=2}^{n}x\left(x-1\right)\left(x-2\right)\frac{n!}{x!\left(n-x\right)!}p^x\left(1-p\right)^{n-x}\\
&= np+3n\left(n-1\right)p^2\\
&\quad +n\left(n-1\right)\left(n-2\right)p^3\sum_{x=0}^{n-3}x\left(x-1\right)\left(x-2\right)\frac{\left(n-3\right)!}{x!\left(n-3-x\right)!}p^x\left(1-p\right)^{n-3-x}\\
&= np+3n\left(n-1\right)p^2+n\left(n-1\right)\left(n-2\right)p^3
\end{aligned}$$

$$\begin{aligned}
E\left[X^4\right] &= E\left[X\left(X-1\right)\left(X-2\right)\left(X-3\right)+6X^3-11X^2+6X\right] \\
&= np+7n\left(n-1\right)p^2+6n\left(n-1\right)\left(n-2\right)p^3 \\
&\quad +E\left[X\left(X-1\right)\left(X-2\right)\left(X-3\right)\right] \\
&= np+7n\left(n-1\right)p^2+6n\left(n-1\right)\left(n-2\right)p^3 \\
&\quad +\sum_{x=0}^{n} x\left(x-1\right)\left(x-2\right)\left(x-3\right)\binom{n}{x}p^x\left(1-p\right)^{n-x} \\
&= np+7n\left(n-1\right)p^2+6n\left(n-1\right)\left(n-2\right)p^3 \\
&\quad +n\left(n-1\right)\left(n-2\right)\left(n-3\right)p^4\sum_{x=0}^{n-4}\frac{\left(n-4\right)!}{x!\left(n-4-x\right)!}p^x\left(1-p\right)^{n-4-x} \\
&= np+7n\left(n-1\right)p^2+6n\left(n-1\right)\left(n-2\right)p^3+n\left(n-1\right)\left(n-2\right)\left(n-3\right)p^4
\end{aligned}$$

B. Tabla de la Normal Estándar

$$\Phi\left(x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{1}{2}y^{2}}dx$$



Segunda cifra decimal del valor de z										
z	**0.00**	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.4878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.8	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.9	1									

C. Criterios de convergencia en $\mathbb{R}$

1. **Teorema C..1**

Si $\sum a_k$ *y* $\sum b_k$ *son series con términos no negativos, si* $\sum b_k$ *converge y si* $a_k \leq b_k$ *para todo* k *suficientemente grande, entonces* $\sum a_k$ *converge.*

Si $\sum a_k$ *converge, entonces* $\sum |a_k|$ *converge.*

Criterio de la Razón *(a) Si* $a_k \neq 0$ *y* $\overline{\lim} \left| \frac{a_{k+1}}{a_k} \right| < 1$ *entonces* $\sum a_k$ *es absolutamente convergente.*

(b) Si $a_k \neq 0$ *y* $\underline{\lim} \left| \frac{a_{k+1}}{a_k} \right| > 1$ *entonces* $\sum a_k$ *es divergente.*

Criterio de la Raiz *(a) Si* $\overline{\lim} \sqrt[k]{|a_k|} < 1$ *entonces* $\sum a_k$ *es absolutamente convergente.*

(b) Si $\underline{\lim} \sqrt[k]{|a_k|} > 1$ *entonces* $\sum a_k$ *es divergente.*

Criterio de la integral *Si* $\sum a_k$ *es una serie de términos no negativos y* f *es una función continua no creciente sobre el intervalo* $[1, \infty\rangle$ *tal que* $f(k) = a_k$, *entonces* $\sum a_k$ *y* $\int_1^\infty f$ *o ambas convergen o ambas divergen.*

Criterio de las series alternantes *Si* $\{a_k\}$ *es una sucesión no creciente de términos positivos y* $\lim a_k = 0$ *entonces* $\sum_{k=1}^{\infty} (-1)^{k+1} a_k$ *converge.*

Las demostraciones de estos teoremas se pueden encontrar en cualquier libro de Cálculo Diferencial e Integral, (por ejemplo en el Hasser 2)

BIBLIOGRAFÍA

[1] Chung Kai L. A *Course in Probability Theory.* Academic Press, Estados Unidos, 1974.

[2] Dudley R. M., *Real Analysis and Probability*, Belmont: Wadsworth & Brooks /Cole Advanced Boocks & Software, Estados Unidos, 1989.

[3] Feller, William. *An Introduction to Probability Theory and Applications*, Segunda edición, Estados Unidos, 1965.

[4] García A. Miguel. *Introducción a la teoría de la Probabilidad I. Primer Curso*, Fondo de Cultura Económica, México, 2005

[5] García A. Miguel. *Introducción a la teoría de la Probabilidad II. Segundo Curso*, Fondo de Cultura Económica, México, 2005.

[6] Gnedenko, B.V. *The theory of Probability and the elements of Statistics*, Chelsea Publishing Company, Estados Unidos, 1989.

[7] Hasser Norman B. *Análisis Matemático, Vol. 2. Curso intermedio.* Trillas, México, 1995.

[8] Heinz Bauer, *Probability theory and elements of measure theory*, E.D. Academic Press, Segunda edición en Inglés, 1981.

[9] Hernández A. Fabian , *Cálculo de probabilidades*, Sociedad Matemática Mexicana, México, 2003.

[10] *Notas del curso de Probabilidad II* de Carrasco Licea Guadalupe

[11] *Notas del curso de Probabilidad II* de Rincón Solis Luis Antonio