38
2ej

# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

# SERIES ALEATORIAS

# T E S I S

QUE PARA OBTENER EL TITULO DE

MATEMATICO

PRESENTA:

REFUGIO TRUJILLO CORTEZ



MEXICO, D. F. 1994

VNIVERSDAD NACIONAL
AVFNºMA DE
MEXICO

**M. EN C. VIRGINIA ABRIN BATULE**
**Jefe de la División de Estudios Profesionales**
**Facultad de Ciencias**
**Presente**

Los abajo firmantes, comunicamos a Usted, que habiendo revisado el trabajo de Tesis que realiz(ó)ron el pasante(s) REFUGIO TRUJILLO CORTEZ

con número de cuenta 9052443-5 con el Título:

SERIES ALEATORIAS

Otorgamos nuestro **Voto Aprobatorio** y consideramos que a la brevedad deberá presentar su Examen Profesional para obtener el título de MATEMATICO

| GRADO | NOMBRE(S) | APELLIDOS COMPLETOS | FIRMA |
|---|---|---|---|
| DRA. | MARIA EMILIA | CABALLERO ACOSTA | |
| **Director de Tesis** | | | |
| DR. | GUILLERMO | GRABINSKY STEIDER | |
| MAT. | LUIS ALBERTO | BRISEÑO AGUIRRE | |
| M.C. | ARMANDO | GARCIA MARTINEZ | |
| **Suplente** | | | |
| M.C. | ALBERTO | CONTRERAS CRISTAN | |
| **Suplente** | | | |

A mis padres
a quienes nunca terminaré de agradecer
la oportunidad que me brindaron para llegar
al final de una de mis metas.

A mis hermanos
por su apoyo y comprensión durante todas
las etapas de mi vida.

A Rocío
por su amor y constante apoyo a ser mejor.

Cuco.

# Contenido

# Introducción

El clásico libro de Probabilidad "Limit Distributions for sums of Independent Random Variables" por B. V. Gnedenko y A. N. Kolmogorov fué publicado en 1949. Desde entonces la teoría de sumas de variables aleatorias independientes o dependientes se ha desarrollado rápidamente. Hoy, una recapitulación del estudio de esta área, y de sus resultados, requeriría de muchos libros. Más aún, el estudio de la convergencia casi segura de una serie infinita de variables aleatorias independientes es un tema muy amplio.

La presente tésis busca dar una presentación global de los principales resultados de convergencia casi segura de una serie aleatoria y algunas de las aplicaciones de ellos. Además, se vincula el tema anterior en el estudio de la convergencia (con una métrica adecuada) de esta serie aleatoria a otra variable aleatoria.

En el primer capítulo, se dan algunos conceptos de probabilidad, así como los resultados necesarios para el desarrollo del capítulo 2. Se define que es una familia "Uniformemente Integrable" de variables aleatorias en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y lo que es un Tiempo de paro. Veremos también lo que es una Martingala y su relación con los objetos anteriormente definidos.

Después, se estudia en el segundo capítulo la convergencia casi segura de una serie aleatoria. Se empezará estudiando la convergencia de esa serie en el caso de que sus sumandos tengan esperanza cero y continuaremos para el caso que sean acotados casi seguramente. En la última sección de éste segundo capítulo, mostramos el Teorema de Tres Series de Kolmogorov para la convergencia casi segura de $S_n$ y algunas consecuencias de éste.

En el tercer capítulo, se estudia el artículo de "Some elementary results on Poisson approximation in a sequence of Bernoulli trials". Este artículo presenta algunos resultados sobre la aproximación de ciertas variables aleatorias discretas a otra discreta y su "rapidez" de convergencia. Para ello, se introducirán dos métricas en el espacio de medidas de probabilidad sobre $(\mathbb{N}, \mathbb{B}(\mathbb{N}))$ y se demostrarán algunas propiedades de estas métricas.

Gracias a estas métricas podremos dar aproximaciones interesantes, que nos permitirán estudiar la distribución de $S_n$, cuando $S_n = \sum_{k=1}^{n} X_k$ donde $X_k$ es una variable aleatoria Bernoulli para cada $k \geq 1$, es decir, $X_k$ satisface:

$$\begin{aligned} P(X_k = 1) &= p_k \\ P(X_k = 0) &= 1 - p_k \quad \forall\, k. \end{aligned}$$

para alguna $p_k \in (0, 1)$.

Cuando estas variables aleatorias $X_k$ son independientes, entonces la sucesión $\{S_n\}_{n=1}^{\infty}$ es un caso especial, y de gran interés, de la teoría estudiada en el capítulo 2. Después, con esta herramienta se darán cotas sobre estas métricas para la aproximación Poisson a una suma de variables aleatorias independientes Bernoulli, y como consecuencia de esto se demostrará la clásica aproximación Poisson. Finalmente, estos resultados se generalizan para el caso de variables aleatorias dependientes.

En el último capítulo, vemos la relación que hay entre una de las métricas introducidas en el capítulo 3, con la métrica bien conocida de Prohorov. Esta última parte , tiene como finalidad principal mostrar la equivalencia entre estas dos métricas en el espacio de medidas de probabilidad sobre $(\mathbb{N}, \mathbb{B}(\mathbb{N}))$.

Es importante ver esta equivalencia, ya que las versiones de esta métrica para variables aleatorias discretas dadas en el capítulo 3 permiten estudiarla de manera fácil y directa, en cambio el estudio de la métrica de Prohorov en general resulta bastante más complicado, de modo que el poder mostrar la equivalencia facilita el estudio general de la métrica de Prohorov, y nos da un importante caso particular de ella, en el cual se extiende de manera muy directa que es lo que mide dicha métrica.

# Capítulo 1

# Preliminares

El propósito de este capítulo es dar algunos conceptos básicos de probabilidad, así como los resultados necesarios para el desarrollo del capítulo 2. Cabe mencionar, que todos los resultados se encuentran con su respectiva demostración, excepto el Teorema de Convergencia para Martingalas, pero se da la referencia para su localización.

## 1.1 Conceptos básicos y el Lema de Borel-Cantelli

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias definidas en este. En esta sección se estudiarán las diversas formas de convergencia de la sucesión $\{X_n\}_{n=1}^{\infty}$, además probaremos algunos resultados vinculados a estas.

**DEFINICIÓN 1.1.1**
*Una sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias converge casi seguramente a una variable aleatoria $X$ si existe un conjunto $A \in \mathcal{F}$ con $P(A) = 0$ tal que, para toda $\omega \in A^c$,*

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

*En este caso escribiremos $X_n \xrightarrow{\text{c.s.}} X$, o diremos que $\{X_n\}_{n=1}^{\infty}$ converge c.s. a $X$.*

### DEFINICIÓN 1.1.2

*Una sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias es de Cauchy casi seguramente si existe un conjunto $A \in \mathcal{F}$ con $P(A) = 0$ tal que, para toda $\omega \in A^c$, $\{X_n(\omega)\}_{n=1}^{\infty}$ es de Cauchy.*

*En general, se dice que una propiedad se cumple casi seguramente (abreviado c.s.) si existe un conjunto de medida cero, fuera del cual la propiedad se cumple.*

El siguiente teorema nos muestra una equivalencia entre las dos definiciones anteriores.

### TEOREMA 1.1.1

*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$. Entonces*

$$X_n \xrightarrow{c.s.} X \Longleftrightarrow \{X_n\}_{n=1}^{\infty} \quad \textit{es de Cauchy c.s.}$$

*Demostración*

Por la completez de los números Reales se sabe que

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \Longleftrightarrow \{X_n(\omega)\}_{n=1}^{\infty} \quad \textit{es de Cauchy},$$

y con esto se obtiene el resultado. □

Se puede mostrar que si $X_n \xrightarrow{c.s.} X$, entonces X es también una variable aleatoria y es única salvo por conjuntos de medida cero en $(\Omega, \mathcal{F}, P)$.

Para establecer el siguiente criterio de convergencia *c.s.* , recordemos que dada una sucesión $(A_n)_{n=1}^{\infty}$ de subconjuntos de $\Omega$, se define el $\limsup_{n\to\infty} A_n$ como

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$

Observemos que

$$\limsup_{n\to\infty} A_n = \{\omega \in \Omega \ / \ \omega \in A_n \ \textit{para una infinidad de enes}\}.$$

### TEOREMA 1.1.2

*a) Una sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias converge c.s. a X*

$$\Longleftrightarrow \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty} \{|X_m - X| \geq \epsilon\}) = 0 \quad \forall \epsilon > 0.$$

*b) Una sucesión $\{X_n\}_{n=1}^{\infty}$ es de Cauchy c.s.*

$$\Longleftrightarrow \lim_{n\to\infty} P(\bigcup_{m=1}^{\infty}\{|X_{m+n}-X_n|\geq\epsilon\})=0 \quad \forall\,\epsilon>0.$$

*Demostración*

a) Para cada $\epsilon>0$, tomemos $E_n(\epsilon)=\{\omega\in\Omega\ /\ |X_n(\omega)-X(\omega)|\geq\epsilon\}$ y sea $A=\{\omega\in\Omega\ /\ \lim_{n\to\infty}X_n(\omega)\neq X(\omega)\}$. Claramente

$$\begin{aligned} A &= \bigcup_{\epsilon>0}\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}E_m(\epsilon) \\ &= \bigcup_{k=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}E_m(\frac{1}{k}) \\ &= \bigcup_{\epsilon>0}\limsup_{n\to\infty}E_n(\epsilon), \end{aligned}$$

así

$$\begin{aligned} X_n \xrightarrow{c.s.} X &\Longleftrightarrow P(A)=0 \\ &\Longleftrightarrow P(\limsup_{n\to\infty}E_n(\epsilon))=0 \quad \forall\,\epsilon>0, \end{aligned}$$

además, como $\{\bigcup_{m=n}^{\infty}E_m(\epsilon)\}_{n=1}^{\infty}$ es una sucesión decreciente, entonces

$$\begin{aligned} \limsup_{n\to\infty}E_n(\epsilon) &= \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}E_m(\epsilon) \\ &= \lim_{n\to\infty}\bigcup_{m=n}^{\infty}E_m(\epsilon), \end{aligned}$$

y en consecuencia

$$\begin{aligned} P(\limsup_{n\to\infty}E_n(\epsilon)) &= P(\lim_{n\to\infty}\bigcup_{m=n}^{\infty}E_m(\epsilon)) \\ &= \lim_{n\to\infty}P(\bigcup_{m=n}^{\infty}E_m(\epsilon)) \\ &= 0. \end{aligned}$$

Por lo tanto

$$X_n \xrightarrow{c.s.} X \Longleftrightarrow \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty}\{\omega \in \Omega \ / \ |X_m(\omega) - X(\omega)| \geq \epsilon\}) = 0 \quad \forall \epsilon > 0.$$

b) Sea $\epsilon > 0$, Ahora tomemos $E_{n,m}(\epsilon) = \{\omega \in \Omega \ / \ |X_n(\omega) - X_m(\omega)| \geq \epsilon\}$ y $A = \{\omega \in \Omega \ / \ \{X_n(\omega)\}_{n=1}^{\infty}$ *no es de Cauchy*$\}$. De manera análoga a (a) se tiene que
$\{X_n\}_{n=1}^{\infty}$ es de Cauchy *c.s.*

$$\begin{aligned}
&\Longleftrightarrow \quad P(A) = 0 \\
&\Longleftrightarrow \quad P(\limsup_{n\to\infty} E_{n,m}(\epsilon)) = 0 \\
&\Longleftrightarrow \quad \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty}\{|X_m - X_n| \geq \epsilon\}) = 0 \quad \forall \epsilon > 0 \\
&\Longleftrightarrow \quad \lim_{n\to\infty} P(\bigcup_{m=1}^{\infty}\{|X_{m+n} - X_n| \geq \epsilon\}) = 0 \quad \forall \epsilon > 0.
\end{aligned}$$

□

## COROLARIO 1.1.1

*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias y sea $X$ una variable aleatoria si*

$$\sum_{n=1}^{\infty} P(\{|X_n - X| \geq \epsilon\}) < \infty \quad \forall \epsilon > 0.$$

*Entonces*

$$X_n \xrightarrow{c.s.} X.$$

*Demostración*

Es claro que

$$P(\bigcup_{m=n}^{\infty}\{|X_m - X| \geq \epsilon\}) \leq \sum_{m=n}^{\infty} P(\{|X_m - X| \geq \epsilon\}),$$

además

$$\begin{aligned}
\lim_{n\to\infty} P(\bigcup_{m=n}^{\infty}\{|X_m - X| \geq \epsilon\}) \quad &\leq \quad \lim_{n\to\infty} \sum_{m=n}^{\infty} P(\{|X_m - X| \geq \epsilon\}) \\
&= \quad 0.
\end{aligned}$$

Por lo tanto $X_n \xrightarrow{c.s.} X$. $\square$

El Lema de Borel-Cantelli es uno de los resultados importantes para el estudio de la convergencia *c.s.* de una serie de variables aleatorias independientes.

**TEOREMA 1.1.3** *( Lema de Borel-Cantelli )*
***Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sea $\{E_n\}_{n=1}^{\infty}$ una sucesión de conjuntos medibles.***
***a) Si $\sum_{n=1}^{\infty} P(E_n) < \infty$, entonces***

$$P(\limsup_{n\to\infty} E_n) = 0.$$

***b) Además, si los $E_n$ son independientes y $\sum_{n=1}^{\infty} P(E_n) = \infty$, entoces***

$$P(\limsup_{n\to\infty} E_n) = 1.$$

*Demostración*
a) Sea

$$\begin{aligned} E &= \limsup_{n\to\infty} E_n \\ &= \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m, \end{aligned}$$

si $F_n = \bigcup_{m=n}^{\infty} E_m$, entonces

$$E = \bigcap_{n=1}^{\infty} F_n \quad y \quad P(F_n) \leq \sum_{m=n}^{\infty} P(E_m),$$

y como $\{F_n\}_{n=1}^{\infty}$ es una sucesión que decrece a E podemos concluir que

$$\begin{aligned} P(E) &= \lim_{n\to\infty} P(F_n) \\ &\leq \lim_{n\to\infty} (\sum_{m=n}^{\infty} P(E_m)) \\ &= 0. \end{aligned}$$

b) Sea $E = \limsup_{n\to\infty} E_n$, entonces

$$E^c = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} E_m^c.$$

Sean $N, n \in \mathbb{N}$ tales que $N > n$. Entonces

$$\begin{aligned} P(\bigcap_{m=n}^{\infty} E_m^c) &\leq P(\bigcap_{m=n}^{N} E_m^c) \\ &= \prod_{m=n}^{N} [1 - P(E_m)] \\ &\leq \exp(-\sum_{m=n}^{N} P(E_m)) \quad \forall N > n, \end{aligned}$$

ya que los $E_m^c$ son independientes y $1 - x \leq \exp(-x)$. Por lo tanto

$$\begin{aligned} P(\bigcap_{m=n}^{\infty} E_m^c) &\leq \lim_{N\to\infty} \exp(-\sum_{m=n}^{N} P(E_m)) \\ &= \exp(-\sum_{m=n}^{\infty} P(E_m)), \end{aligned}$$

y si $\sum_{n=1}^{\infty} P(E_n) = \infty$, entonces

$$P(\bigcap_{m=n}^{\infty} E_m^c) = 0 \quad \forall n \geq 1,$$

esto implica que $P(E^c) = 0$. Por lo tanto

$$P(E) = 1.$$

Frecuentemente se utiliza la palabra evento en lugar de conjunto medible.

**COROLARIO 1.1.2**

*Sea $\{E_n\}_{n=1}^{\infty}$ una sucesión de eventos independientes. Entonces*

$$P(\limsup_{n\to\infty} E_n) = 0 \iff \sum_{n=1}^{\infty} P(E_n) < \infty.$$

**COROLARIO 1.1.3**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes. Entonces*

$$X_n \xrightarrow{c.s.} 0 \Longleftrightarrow \sum_{n=1}^{\infty} P(|X_n| \geq \epsilon) < \infty \quad \forall \epsilon > 0.$$

*Demostración*
Sea $E_n = \{|X_n| \geq \epsilon\}$ y $A = \{\omega \in \Omega \ / \ \lim_{n\to\infty} X_n(\omega) \neq 0\}$, entonces

$$\begin{aligned} X_n \xrightarrow{c.s.} 0 &\Longleftrightarrow P(A) = 0 \\ &\Longleftrightarrow P(\limsup_{n\to\infty} E_n(\epsilon)) = 0 \quad \forall \epsilon > 0 \\ &\Longleftrightarrow \sum_{n=1}^{\infty} P(E_n) < \infty \quad \forall \epsilon > 0. \end{aligned}$$

□

## 1.2 Truncación y una desigualdad de Kolmogorov

**DEFINICIÓN 1.2.1**
*Sea $X$ una variable aleatoria, decimos que $X$ está centrada en $c$ si*

$$E[(X - c)] = 0.$$

La elección de la constante c juega un papel importante en algunos problemas de la Teoría de Probabilidad.
**Ejemplo 1.2.1**
Si $X \in L^1(\Omega, \mathcal{F}, P)$ y elegimos a $c = E[X]$, decimos que X está centrada en su Esperanza.

**DEFINICIÓN 1.2.2**
*Sea $c > 0$ una constante. Sea*

$$X^c = \begin{cases} X & si\ |X| < c. \\ 0 & si\ |X| \geq c. \end{cases}$$

*La variable aleatoria así definida es llamada $X$ truncada en $c$.*

como $X^c$ es una variable aleatoria acotada, entonces

$$X^c \in L^p(\Omega, \mathcal{F}, P) \quad \forall p \geq 1$$

**DEFINICIÓN 1.2.3**

*Dos sucesiones de variables aleatorias $\{X_n\}_{n=1}^{\infty}$ y $\{X'_n\}_{n=1}^{\infty}$ son "Cola Equivalente" si difieren c.s. en un número finito de terminos. Esto es, para casi toda $\omega \in \Omega$, existe un número entero $N(\omega)$ tal que,*

$$X_n(\omega) = X'_n(\omega) \quad \forall n \geq N(\omega).$$

*En otras palabras*

$$P(\limsup_{n\to\infty}\{X_n \neq X'_n\}) = 0.$$

**DEFINICIÓN 1.2.4**

*Dos sucesiones de variables aleatorias $\{X_n\}_{n=1}^{\infty}$ y $\{X'_n\}_{n=1}^{\infty}$ son "Equivalentemente Convergentes" si convergen sobre el mismo conjunto excepto por un conjunto de probabilidad 0.*

**TEOREMA 1.2.1**

*Sean $\{X_n\}_{n=1}^{\infty}$ y $\{X'_n\}_{n=1}^{\infty}$ dos sucesiones de variables aleatorias tales que*

$$\sum_{n=1}^{\infty} P(X_n \neq X'_n) < \infty.$$

*Entonces las siguientes afirmaciones son verdaderas.*
*a) Las sucesiones $\{X_n\}_{n=1}^{\infty}$ y $\{X'_n\}_{n=1}^{\infty}$ son Cola Equivalente.*
*b) Las series $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} X'_n$ son Equivalentemente Convergentes.*

*Demostración*

a)

$$\begin{aligned}
P(\limsup_{n\to\infty}\{X_n \neq X'_n\}) &= \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty}\{X_m \neq X'_m\}) \\
&\leq \lim_{n\to\infty} \sum_{m=n}^{\infty} P(X_m \neq X'_m) \\
&= 0.
\end{aligned}$$

b) Sea

$$A = \{\omega \in \Omega \;/\; \sum_{n=1}^{\infty} X_n(\omega) \text{ converge}\} \text{ y } B = \{\omega \in \Omega \;/\; \sum_{n=1}^{\infty} X_n'(\omega) \text{ converge}\}$$

y sea $N = (A \cap B^c) \cup (B \cap A^c) = (A \cup B) \setminus (A \cap B)$. Entonces

$$N \subseteq \limsup_{n\to\infty} \{X_n \neq X_n'\},$$

por lo tanto $P(N) = 0$. □

**TEOREMA 1.2.2** *( Desigualdad de Kolmogorov )*
*Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes tales que $|X_k| \leq C$ c.s. para toda $1 \leq k \leq n$, y sea $S_k = \sum_{i=1}^{k} X_i \quad k = 1, 2, \cdots n$. Entonces*

$$P(\max_{1\leq k\leq n} |S_k - E[S_k]| > \epsilon) \geq 1 - \frac{(2C+\epsilon)^2}{Var(S_n)} \quad \forall \epsilon > 0.$$

*Demostración*
Sea

$$A = \{\omega \in \Omega \;/\; \max_{1\leq k\leq n} |S_k(\omega) - E[S_k]| > \epsilon\}.$$

Para cada $\epsilon > 0$, definimos

$$\begin{aligned} A_1 &= \{\omega \in \Omega \;/\; |S_1(\omega) - E[S_1]| > \epsilon\} \quad y\ \forall\, 1 < j \leq n \\ A_j &= \{\omega \in \Omega \;/\; |S_i(\omega) - E[S_i]| \leq \epsilon \text{ si } i < j, \text{ y } |S_j(\omega) - E[S_j] > \epsilon\}. \end{aligned}$$

Es claro que

$$A = \bigcup_{k=1}^{n} A_k \quad y \quad A_i \cap A_j = \emptyset \quad \forall i \neq j.$$

Sea $Z_k = S_k - E[S_k]$, entonces

$$\begin{aligned} E[Z_n^2 I_A] &= \sum_{k=1}^{n} E[Z_n^2 I_{A_k}] \\ &= \sum_{k=1}^{n} E[(Z_k - (Z_k - Z_n))^2 I_{A_k}] \\ &= \sum_{k=1}^{n} E[Z_k^2 I_{A_k}] + \sum_{k=1}^{n} 2E[Z_k(Z_n - Z_k) I_{A_k}] + \sum_{k=1}^{n} E[(Z_n - Z_k)^2 I_{A_k}] \end{aligned}$$

pero

$$\begin{aligned} E[Z_k I_{A_k}(Z_n - Z_k)] &= E[Z_k I_{A_k}]E[Z_n - Z_k] \\ &= 0 \qquad \forall\, k = 1, \cdots, n. \end{aligned}$$

porque $Z_k I_{A_k}$, $Z_n - Z_k$ son independientes y $E[Z_n - Z_k] = 0$ para $k = 1, \cdots, n$. Entonces

$$E[Z_n^2 I_A] = \sum_{k=1}^{n} E[Z_k^2 I_{A_k}] + \sum_{k=1}^{n} E[(Z_n - Z_k)^2 I_{A_k}].$$

Ahora bien, sobre el conjunto $A_k$, se sabe que $|Z_{k-1}| \leq \epsilon$ y así sobre $A_k$ se tiene:

$$\begin{aligned} |Z_k| &= |Z_{k-1} + X_k - E[X_k]| \\ &\leq |Z_{k-1}| + |X_k| + |E[X_k]| \\ &\leq \epsilon + 2C. \end{aligned}$$

Además

$$\begin{aligned} E[(Z_n - Z_k)^2 I_{A_k}] &= E[(\sum_{j=k+1}^{n} (X_j - E[X_j]))^2 I_{A_k}] \\ &= E[I_{A_k}]E[(\sum_{j=k+1}^{n} (X_j - E[X_j]))^2] \\ &= P(A_k) \sum_{j=k+1}^{n} E[(X_j - E[X_j])^2] \\ &= P(A_k) \sum_{j=k+1}^{n} Var(X_j) \\ &\leq P(A_k) \sum_{k=1}^{n} Var(X_k), \end{aligned}$$

porque $I_{A_k}$ y $\sum_{j=k+1}^{n}(X_j - E[X_j])$ son variables aleatorias independientes (ya que son funciones que dependen de $X_1, \cdots, X_k$ y $X_{k+1}, \cdots, X_n$, respectivamente). Por lo tanto

$$
\begin{aligned}
E[Z_n^2 I_A] &= \sum_{k=1}^{n} E[Z_k^2 I_{A_k}] + \sum_{k=1}^{n} E[(Z_n - Z_k)^2 I_{A_k}] \\
&\leq (\epsilon + 2C)^2 \sum_{k=1}^{n} E[I_{A_k}] + \sum_{k=1}^{n} P(A_k) \sum_{k=1}^{n} Var(X_k) \\
&= [(\epsilon + 2C)^2 + \sum_{k=1}^{n} Var(X_k)] \sum_{k=1}^{n} P(A_k) \\
&= [(\epsilon + 2C)^2 + Var(S_n)] P(A).
\end{aligned}
$$

En consecuencia

$$E[Z_n^2 I_A] \leq [(\epsilon + 2C)^2 + Var(S_n)] P(A).$$

Por otro lado, como $Z_n^2 \leq \epsilon^2$ sobre el conjunto $\Omega \setminus A$,

$$
\begin{aligned}
E[Z_n^2 I_A] &= E[Z_n^2] - E[Z_n^2 I_{\Omega - A}] \\
&\geq Var(S_n) - \epsilon^2 P(\Omega - A) \\
&= Var(S_n) - \epsilon^2 + \epsilon^2 P(A)
\end{aligned}
$$

Así obtenemos

$$P(A)[(\epsilon + 2C)^2 + Var(S_n)] \geq Var(S_n) - \epsilon^2 + \epsilon^2 P(A),$$

y

$$
\begin{aligned}
P(A) &\geq \frac{Var(S_n) - \epsilon^2}{(\epsilon + 2C)^2 + Var(S_n) - \epsilon^2} \\
&= 1 - \frac{(\epsilon + 2C)^2}{(\epsilon + 2C)^2 + Var(S_n) - \epsilon^2} \\
&\geq 1 - \frac{(\epsilon + 2C)^2}{Var(S_n)}.
\end{aligned}
$$

□

## 1.3 Integrabilidad Uniforme

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Si $X \in L^1(\Omega, \mathcal{F}, P)$ entonces

$$\lim_{M\to\infty}\left(\int_{\{|X|\geq M\}} |X|\, dP\right) = 0$$

En efecto:
Sean $M_1 \leq M_2$, entonces $\{|X| \geq M_2\} \subseteq \{|X| \geq M_1\}$ y en consecuencia

$$\int_{\{|X|\geq M_2\}} |X|\, dP \leq \int_{\{|X|\geq M_1\}} |X|\, dP \quad (es\ decreciente)$$

Por lo tanto

$$\lim_{M\to\infty}\left(\int_{\{|X|\geq M\}} |X|\, dP\right) = \lim_{n\to\infty}\left(\int_{\{|X|\geq n\}} |X|\, dP\right)$$

Vamos a demostrar que el segundo limite vale cero.
Sea $X_n = I_{\{|X|\geq n\}}X$, claramente se tiene que $X_n \xrightarrow{c.s.} 0$ y $|X_n| \leq |X|$ para toda n, entonces por el Teorema de Convergencia Dominada

$$\lim_{n\to\infty}\left(\int_{\{|X|\geq n\}} |X|\, dP\right) = 0$$

□

Si ahora tenemos $\{X_\alpha\}_{\alpha\in T}$ una familia de funciones tales que para cada $\alpha \in T$ $X_\alpha \in L^1(\Omega, \mathcal{F}, P)$, para cada $\alpha \in T$ se cumple ( por lo anterior ) que:

$$\lim_{M\to\infty}\left(\int_{\{|X_\alpha|\geq M\}} |X_\alpha|\, dP\right) = 0.$$

Pero es claro que el límite no necesariamente es uniforme.

**Ejemplo 1.3.1**
Sean $X_n = n$ para toda $n \in \mathbb{N}$, $\Omega = [0,1]$, $\mu$ =medida de lebesgue en $[0,1]$. Entonces

$$\int_\Omega |X_n|\, d\mu = n \quad \forall\, n \in \mathbb{N}.$$

Por lo tanto $X_n \in L^1(\Omega, \mathcal{F}, P)$. Pero, claramente, dado $\epsilon > 0$ no existe una M que haga a

$$(\int_{\{|X_n| \geq M\}} |X_n| \, d\mu) < \epsilon \quad \forall n \in I\!N.$$

□

**DEFINICIÓN 1.3.1**
*Un subconjunto $\mathcal{H} \subset L^1(\Omega, \mathcal{F}, P)$ se le llama* *__Uniformemente Integrable__* *si*

$$\lim_{M \to \infty} (\sup_{X \in \mathcal{H}} \{\int_{\{|X| \geq M\}} |X| \, dP\}) = 0.$$

**Ejemplo 1.3.2**
Si $\mathcal{H}$ es una familia finita $\{X_1, X_2, \cdots, X_k\}$ y $\mathcal{H} \subset L^1(\Omega, \mathcal{F}, P)$, entonces $\mathcal{H}$ es Uniformemente integrable, ya que:
dada $\epsilon > 0$ y $1 \leq s \leq k$, existe $M_{(s,\epsilon)}$ tal que

$$(\int_{\{|X_s| \geq M\}} |X_s| \, dP) < \epsilon \quad \forall \, M \geq M_{(s,\epsilon)}$$

Además, si $M_\epsilon = \max_{1 \leq s \leq k}\{M_{(s,\epsilon)}\}$, entonces

$$\sup_{1 \leq s \leq k} \{\int_{\{|X_s| \geq M\}} |X_s| \, dP\} < \epsilon \quad \forall \, M \geq M_\epsilon.$$

□

**Ejemplo 1.3.3**
Si $\mathcal{H} \subset L^1(\Omega, \mathcal{F}, P)$ y si existe $Y \in L^1(\Omega, \mathcal{F}, P)$ tal que $|X| \leq |Y|$ para toda $X \in \mathcal{H}$. Entonces $\mathcal{H}$ es uniformemente integrable. En efecto:

$$\begin{aligned} \int_{\{|X| \geq M\}} |X| \, dP &\leq \int_{\{|Y| \geq M\}} |X| \, dP \quad \forall \, X \in \mathcal{H} \\ &\leq \int_{\{|Y| \geq M\}} |Y| \, dP \quad \forall \, X \in \mathcal{H}. \end{aligned}$$

Entonces

$$\sup_{X \in \mathcal{H}} \{\int_{\{|X| \geq M\}} |X| \, dP\} \leq \int_{\{|Y| \geq M\}} |Y| \, dP,$$

en consecuencia

$$\begin{aligned}\lim_{M\to\infty}(\sup_{X\in\mathcal{H}}\{\int_{\{|X|\geq M\}}|X|\,dP\}) &\leq \lim_{M\to\infty}\int_{\{|Y|\geq M\}}|Y|\,dP\\ &= 0.\end{aligned}$$

Por lo tanto $\mathcal{H}$ es uniformemente integrable. □

**TEOREMA 1.3.1**
*Sea $\mathcal{H}\subset L^1(\Omega,\mathcal{F},P)$. Entonces $\mathcal{H}$ es Uniformemente Integrable si y sólo si se cumplen las dos condiciones siguientes*
*a) dada $\epsilon>0$, existe $\delta_\epsilon>0$ tal que si $A\in\mathcal{F}$ y $P(A)<\delta_\epsilon$, entonces*

$$\sup_{X\in\mathcal{H}}\{\int_A|X|\,dP\}<\epsilon$$

*b)*

$$\sup_{X\in\mathcal{H}}\{\int_\Omega|X|\,dP\}<\infty$$

*Demostración*
$\Rightarrow$)
Sea $A\in\mathcal{F}$, entonces

$$\begin{aligned}\int_A|X|\,dP &= \int_{A\cap\{|X|\geq M\}}|X|\,dP+\int_{A\cap\{|X|<M\}}|X|\,dP &(1.1)\\ &\leq \int_{\{|X|\geq M\}}|X|\,dP+M\,P(A) &(1.2)\end{aligned}$$

Por hipótesis tenemos que, dado $\epsilon>0$, existe $M_0>0$ tal que si $M\geq M_0$, entonces

$$\sup_{X\in\mathcal{H}}\{\int_{\{|X|\geq M\}}|X|\,dP\}<\frac{\epsilon}{2}$$

para $M_0$ fija, sea $A\in\mathcal{F}$ tal que $P(A)<\frac{\epsilon}{2M_0}=\delta_\epsilon$. Entonces de la desigualdad 1.2 se cumple que

$$\sup_{X\in\mathcal{H}}\{\int_A|X|\,dP\}\leq\sup_{X\in\mathcal{H}}\{\int_{\{|X|\geq M_0\}}|X|\,dP\}+M_0P(A)<\epsilon$$

Por lo tanto

$$\sup_{X\in\mathcal{H}}\{\int_A |X|\,dP\} < \epsilon$$

Por otra parte,

$$\begin{aligned}\int_\Omega |X|\,dP &\leq \int_{\{|X|\geq M_0\}} |X|\,dP + \int_{\{|X|<M_0\}} |X|\,dP \\ &\leq \frac{\epsilon}{2} + M_0 \quad \forall X \in \mathcal{H}.\end{aligned}$$

Y en consecuencia

$$\sup_{X\in\mathcal{H}}\{\int_\Omega |X|\,dP\} < \frac{\epsilon}{2} + M_0.$$

$\Leftarrow$)
Utilizando (b) se tiene que

$$\begin{aligned}P(|X|\geq M) &= \int_{\{|X|\geq M\}} 1\,dP \\ &\leq \frac{1}{M}\int_{\{|X|\geq M\}} |X|\,dP \\ &\leq \frac{1}{M}(\sup_{X\in\mathcal{H}}\{\int_\Omega |X|\,dP\}) \\ &= \frac{K}{M} \quad (K<\infty)\end{aligned}$$

Además de (a) se sabe que
Dada $\epsilon > 0$, existe $\delta_\epsilon > 0$ tal que si $A \in \mathcal{F}$ y $P(A) < \delta_\epsilon$, entonces

$$\sup_{X\in\mathcal{H}}\{\int_A |X|\,dP\} < \epsilon.$$

Pero, dada $\delta_\epsilon > 0$, existe $M_0 > 0$ (suficientemente grande) tal que si $M \geq M_0$ se tiene que

$$P(|X|\geq M) \leq \frac{K}{M} \leq \frac{K}{M_0} < \delta_\epsilon.$$

En consecuencia

$$\sup_{X\in\mathcal{H}}\{\int_{\{|X|\geq M\}} |X|\,dP\} < \epsilon. \quad \forall\, M \geq M_0.$$

Y por lo tanto $\mathcal{H}$ es uniformemente integrable. □

## 1.4 Tiempos de paro

**DEFINICIÓN 1.4.1**
*Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad fijo y $T \subseteq \Re^+$. Si $(\mathcal{F}_\lambda)_{\lambda \in T}$ es una familia de sub-$\sigma$-álgebras de $\mathcal{F}$ tal que*

$$\lambda_1 \leq \lambda_2 \quad \Longrightarrow \quad \mathcal{F}_{\lambda_1} \subseteq \mathcal{F}_{\lambda_2}.$$

*En este caso diremos que $(\mathcal{F}_\lambda)_{\lambda \in T}$ es una <u>Filtración</u> en $(\Omega, \mathcal{F})$.*

En esta sección y durante todo este trabajo consideraremos a $T = I\!N$, de modo que la familia $(\mathcal{F}_\lambda)_{\lambda \in T}$ resulta ser una sucesión creciente $(\mathcal{F}_n)_{n \in I\!N}$ de sub-$\sigma$-álgebras de $\mathcal{F}$.

**DEFINICIÓN 1.4.2**
*Una función $\tau : \Omega \longrightarrow I\!N \cup \{\infty\}$ se llama Tiempo de Paro con respecto a $(\mathcal{F}_n)_{n \in I\!N}$ si*

$$\{\tau = n\} \in \mathcal{F}_n \quad \forall n \geq 1.$$

**Observación 1.4.1**
$\tau$ es tiempo de paro c.r.a. $(\mathcal{F}_n)_{n \in I\!N}$

$$\Longleftrightarrow \{\tau \leq n\} \in \mathcal{F}_n \quad \forall n \geq 1.$$

Mostremos esta equivalencia.
$\Rightarrow$)

$$\{\tau \leq n\} = \bigcup_{k=1}^{n} \{\tau = k\} \in \mathcal{F}_n$$

porque $\{\tau = k\} \in \mathcal{F}_k \subseteq \mathcal{F}_n$ para $k \leq n$.
$\Leftarrow$)

$$\{\tau = n\} = \{\tau \leq n\} \setminus \{\tau \leq n-1\} \in \mathcal{F}_n$$

ya que $\{\tau \leq n-1\} \in \mathcal{F}_{n-1} \subseteq \mathcal{F}_n$. $\square$

**Ejemplo 1.4.1**
Sean $B \in I\!B(\Re)$, $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias definidas en

$\Omega$ y $(\mathcal{F}_n)_{n\in\mathbb{N}}$ una sucesión creciente de $\sigma$-álgebras de $\Omega$, tales que para cada $n \in \mathbb{N}$ $X_n$ es $\mathcal{F}_n$-medible ( por ejemplo $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$ ). Definimos

$$\tau_B(\omega) = \begin{cases} \inf_n\{n \in \mathbb{N} \ / \ X_n(\omega) \in B\} & \text{si } \{n \in \mathbb{N} \ / \ X_n(\omega) \in B\} \neq \emptyset \\ \infty & \text{en otro caso} \end{cases}$$

Entonces $\tau_B$ es tiempo de paro con respecto a $(\mathcal{F}_n)_{n\in\mathbb{N}}$ y se llama el primer instante en que el proceso $\{X_n\}_{n=1}^{\infty}$ " le pega " a B ó primera entrada del proceso a B. Verifiquemos que es tiempo de paro.

$$\{\tau_B = n\} = \{X_n \in B\} \cap (\bigcap_{k=1}^{n-1}\{X_k \notin B\}) \in \mathcal{F}_n$$

porque $\{X_n \in B\} \in \mathcal{F}_n$ $y$ $\{X_k \notin B\} \in \mathcal{F}_k \subseteq \mathcal{F}_n$ para $k < n$. □

**DEFINICIÓN 1.4.3**

*Sea $\tau$ un tiempo de paro con respecto a $(\mathcal{F}_n)_{n\in\mathbb{N}}$. se define la "$\sigma$-álgebra parada en $\tau$" como*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \ / \ A \cap \{\tau = n\} \in \mathcal{F}_n \quad \forall n.\}$$

*y representa a los eventos que ocurren hasta el tiempo $\tau$.*

Veamos que efectivamente es $\sigma$-álgebra.

i) $\emptyset, \Omega \in \mathcal{F}_\tau$

ii) Sea $A \in \mathcal{F}_\tau$

$$\begin{aligned} &\Rightarrow \quad A \cap \{\tau = n\} \in \mathcal{F}_n \quad \forall n. \\ &\Rightarrow \quad (A \cap \{\tau = n\})^c \in \mathcal{F}_n \quad \forall n. \, (\textit{ya que } \mathcal{F}_n \textit{ es } \sigma - \acute{a}lgebra) \\ &\Rightarrow \quad A^c \cup \{\tau \neq n\} \in \mathcal{F}_n \quad \forall n. \end{aligned}$$

Por lo tanto

$$A^c \cap \{\tau = n\} = (A^c \cup \{\tau \neq n\}) \cap \{\tau = n\} \in \mathcal{F}_n \quad \forall n.$$

porque $A^c \cup \{\tau \neq n\} \in \mathcal{F}_n$ y $\{\tau = n\} \in \mathcal{F}_n$ por ser $\tau$ tiempo de paro.

iii) Sean $A_1, A_2, \cdots, A_k, \cdots \in \mathcal{F}_\tau$ y sea $A = \bigcup_{n=1}^{\infty} A_n$, tomemos $n \in I\!N$, entonces

$$\begin{aligned} A \cap \{\tau = n\} &= (\bigcup_{n=1}^{\infty} A_n) \cap \{\tau = n\} \\ &= \bigcup_{k=1}^{\infty} (A_k \cap \{\tau = n\}), \end{aligned}$$

pero $A_k \cap \{\tau = n\} \in \mathcal{F}_n$ y como $\mathcal{F}_n$ es $\sigma$-álgebra, entonces

$$A \cap \{\tau = n\} = \bigcup_{k=1}^{\infty} (A_k \cap \{\tau = n\}) \in \mathcal{F}_n \quad \forall n.$$

Por lo tanto $\mathcal{F}_\tau$ es $\sigma$-álgebra. □

**Observación 1.4.2**
Existe otra caracterización de $\sigma$-álgebra parada. Esto es:

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \;/\; A \cap \{\tau \leq n\} \in \mathcal{F}_n \; \forall n\},$$

es $\sigma$-álgebra parada.
*Demostración*

$$\begin{aligned} i)\; A \cap \{\tau \leq n\} &= A \cap (\bigcup_{k=1}^{n} \{\tau = k\}) = \bigcup_{k=1}^{n} (A \cap \{\tau = k\}) \in \mathcal{F}_n \quad \forall n. \\ ii)\; A \cap \{\tau = n\} &= A \cap (\{\tau \leq n\} \setminus \{\tau \leq n-1\}) \\ &= A \cap \{\tau \leq n\} \cap (\{\tau \leq n-1\})^c \in \mathcal{F}_n \quad \forall n. \end{aligned}$$

De (i) y (ii) se ve que las definiciones coinciden. □

**PROPIEDADES**
1) Sean $\tau_1$, $\tau_2$ dos tiempos de paro c.r.a $(\mathcal{F}_n)_{n \in I\!N}$ entonces

$$(\tau_1 + \tau_2),\; \min\{\tau_1, \tau_2\},\; \max\{\tau_1, \tau_2\}$$

son también tiempos de paro c.r.a. la misma filtración.
2) Si $\tau_1$, $\tau_2$ son tiempos de paro c.r.a. $(\mathcal{F}_n)_{n \in I\!N}$ y si $\tau_1 \leq \tau_2$, entonces

$$\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}.$$

3) Sean $\{X_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias, $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$ y $\tau$ un tiempo de paro c.r.a. $(\mathcal{F}_n)_{n\in\mathbb{N}}$, entonces la variable aleatoria

$$X_\tau(\omega) = \begin{cases} X_{\tau(\omega)}(\omega) & \text{si } \tau(\omega) < \infty \\ 0 & \text{si } \tau(\omega) = \infty \end{cases}$$

resulta ser $\mathcal{F}_\tau$-medible, ( se le llama "proceso parado en $\tau$" ).

*Demostración*

1)

$$\{\tau_1 + \tau_2 = n\} = \bigcup_{k=1}^{n} \{\tau_1 = n - k\} \cap \{\tau_2 = k\}$$

Pero si $k = 1, \cdots, n$. se tiene que

$$\mathcal{F}_{n-k} \subseteq \mathcal{F}_n \; y \; \mathcal{F}_k \subseteq \mathcal{F}_n \quad \forall n.$$

En consecuencia $\tau_1 + \tau_2$ es tiempo de paro.

Para demostrar la otra parte de este inciso nótese que:

$$\begin{aligned} \{\max\{\tau_1, \tau_2\} \leq n\} &= \{\tau_1 \leq n\} \cap \{\tau_2 \leq n\} \in \mathcal{F}_n \quad \forall n. \\ \{\min\{\tau_1, \tau_2\} \leq n\} &= \{\tau_1 \leq n\} \cup \{\tau_2 \leq n\} \in \mathcal{F}_n \quad \forall n. \end{aligned}$$

Por lo tanto $\max\{\tau_1, \tau_2\}$ $y$ $\min\{\tau_1, \tau_2\}$ son tiempos de paro.

2) Sea $A \in \mathcal{F}_{\tau_1}$, entonces $A \cap \{\tau_1 \leq n\} \in \mathcal{F}_n$ para toda n, y como

$$\tau_1 \leq \tau_2 \Rightarrow \{\tau_2 \leq n\} \subseteq \{\tau_1 \leq n\} \quad \forall n,$$

entonces

$$\begin{aligned} A \cap \{\tau_2 \leq n\} &= A \cap \{\tau_2 \leq n\} \cap \{\tau_1 \leq n\} \\ &= (A \cap \{\tau_1 \leq n\}) \cap \{\tau_2 \leq n\} \in \mathcal{F}_n \quad \forall n. \end{aligned}$$

En consecuencia $A \in \mathcal{F}_{\tau_2}$. Por lo tanto $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

3) Sea $a \in \Re$, entonces

$$\{X_\tau \leq a\} \cap \{\tau = n\} = \{X_n \leq a\} \cap \{\tau = n\} \; \in \mathcal{F}_n \quad \forall n.$$

ya que $X_n$ es $\mathcal{F}_n$-medible y $\tau$ es tiempo de paro. Por lo tanto

$$\{X_\tau \leq a\} \in \mathcal{F}_\tau.$$

Por lo tanto $X_\tau$ es $\mathcal{F}_\tau$-medible. □

## 1.5 Martingalas

**DEFINICIÓN 1.5.1**
*Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad fijo y $(\mathcal{F}_n)_{n\in\mathbb{N}}$ una Filtración en $(\Omega, \mathcal{F}, P)$. Si $\{X_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias definidas en $(\Omega, \mathcal{F})$ y además si $X_n$ es $\mathcal{F}_n$-medible, entonces decimos que $\{X_n\}_{n=1}^{\infty}$ es MARTINGALA con respecto a $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si*

1. $X_n \in L^1(\Omega, \mathcal{F}, P) \quad \forall n \in \mathbb{N}$.
2. $E[X_{n+1} / \mathcal{F}_n] = X_n \quad c.s. \quad \forall n \in \mathbb{N}$.

Esta definición es un caso particular de una definición más general para una familia no necesariamente numerable de variables aleatorias, pero para nuestro objetivo es suficiente este caso.

**Observación 1.5.1**
Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias y sea $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$, entonces $X_n$ es $\mathcal{F}_n$-medible y en este caso diremos solamente que $(X_n)_{n\in\mathbb{N}}$ es martingala, sin especificar cual es la filtración, pues se trata de la filtración " natural ".

**PROPIEDADES**

1) $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ es Martingala

$$\iff \int_A X_n \, dP = \int_A X_{n+1} \, dP \quad \forall n \ y \ \forall A \in \mathcal{F}_n.$$

2) Si $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ es una Martingala

$$\Longrightarrow E[X_n] = C \quad \forall n.$$

donde C es una constante.

3) Si $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ y $(Y_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ son Martingalas

$$\Longrightarrow ((aX_n + bY_n), \mathcal{F}_n)_{n\in\mathbb{N}} \quad es\ Martingala.$$

*Demostración*
1)

$$\begin{aligned} & \int_A X_n \, dP = \int_A X_{n+1} \, dP \quad \forall A \in \mathcal{F}_n \; y \; \forall n. \\ \Longleftrightarrow \quad & \int_A (X_n - X_{n+1}) \, dP = 0 \quad \forall A \in \mathcal{F}_n \; y \; \forall n. \\ \Longleftrightarrow \quad & E[X_n - X_{n+1} \; / \; \mathcal{F}_n] = 0 \quad c.s. \quad \forall n. \\ \Longleftrightarrow \quad & E[X_{n+1} \; / \; \mathcal{F}_n] = X_n \quad c.s. \quad \forall n. \end{aligned}$$

(2) es consecuencia del inciso anterior y (3) de la linealidad de la esperanza condicional. □

**Ejemplo 1.5.1**
Sean $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes e integrables tales que $E[X_k] = 0$ para toda $k \geq 1$. Sea $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$. Entonces

$$(S_n, \mathcal{F}_n)_{n \in I\!N} \quad es \; Martingala.$$

Aquí $S_n = \sum_{k=1}^{n} X_k$.
*Demostración*
Para cada $n \in I\!N$ $S_n \in L^1(\Omega, \mathcal{F}, P)$, además

$$\begin{aligned} E[S_{n+1} \; / \; \mathcal{F}_n] &= E[X_{n+1} + S_n \; / \; \mathcal{F}_n] \\ &= S_n + E[X_{n+1} \; / \; \mathcal{F}_n] \\ &= S_n \quad c.s. \quad \forall n. \end{aligned}$$

Por lo tanto $(S_n, \mathcal{F}_n)_{n \in I\!N}$ es Martingala. □

**TEOREMA 1.5.1** *( Teorema de Paro o de muestreo opcional )*
***Sea $(X_n, \mathcal{F}_n)_{n \in I\!N}$ una martingala y $\tau_1, \tau_2$ dos tiempos de paro con respecto a $(\mathcal{F}_n)_{n \in I\!N}$ tales que $\tau_1 \leq \tau_2$. Si $\tau_2$ es acotado, entonces***

$$E[X_{\tau_2} \; / \; \mathcal{F}_{\tau_1}] = X_{\tau_1}.$$

*Demostración*
Por la propiedad (3) de tiempos de paro tenemos que $X_{\tau_i}$ es $\mathcal{F}_{\tau_i}$-medible i=1,2. y como $\tau_2$ es acotado, existe $n \in I\!N$ tal que $\tau_2 \leq n$, entonces

$$|X_{\tau_i}| \leq \max_{1 \leq k \leq n} \{|X_k|\} \quad i = 1, 2.$$

Por lo tanto $X_{\tau_1}, X_{\tau_2} \in L^1(\Omega, \mathcal{F}, P)$.
Es suficiente demostrar que

$$\int_A X_{\tau_1}\, dP = \int_A X_{\tau_2}\, dP \quad \forall\, A \in \mathcal{F}_{\tau_1}.$$

Primer caso, si $\tau_2 - \tau_1 \leq 1$:
Sea $A \in \mathcal{F}_{\tau_1}$

$$\begin{aligned}
\int_A (X_{\tau_1} - X_{\tau_2})\, dP &= \int_{A\cap\{\tau_2>\tau_1\}} (X_{\tau_1} - X_{\tau_2})\, dP \\
&= \sum_{k=1}^{n} \int_{\{\tau_1=k\}\cap A\cap\{\tau_2>\tau_1\}} (X_{\tau_1} - X_{\tau_2})\, dP \\
&= \sum_{k=1}^{n-1} \int_{\{\tau_1=k\}\cap A\cap\{\tau_2=k+1\}} (X_k - X_{k+1})\, dP,
\end{aligned}$$

porque $\tau_2 - \tau_1 \leq 1$ y la integral sobre $\{\tau_2 = k\}$ es cero.
Ahora bien, como

$$\{\tau_1 = k\} \cap A \in \mathcal{F}_k \quad y \quad \{\tau_2 > k\} = \{\tau_2 \leq k\}^c \in \mathcal{F}_k$$

y $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ es Martingala, entonces

$$\int_{\{\tau_1=k\}\cap A\cap\{\tau_2=k+1\}} (X_k - X_{k+1})\, dP = 0. \quad \forall\, k = 1, \cdots, (n-1).$$

Por lo tanto

$$\int_A (X_{\tau_1} - X_{\tau_2})\, dP = 0 \quad \forall\, A \in \mathcal{F}_{\tau_1}.$$

Segundo caso, caso general:
Sea

$$\tau_1 = \tau^{(0)} \leq \tau^{(1)} \leq \cdots \leq \tau^{(m)} = \tau_2$$

tiempos de paro tales que $\tau^{(k+1)} - \tau^{(k)} \leq 1$, por ejemplo

$$\tau^{(k)} = \max\{\tau_1, \min\{\tau_2, k\}\} \quad \forall k = 0, 1, \cdots, m.$$

los cuales se demostró que son tiempos de paro ( *propiedad 1, sección 1.4* ).
Y se aplica el primer caso m veces. □

**COROLARIO 1.5.1**
*Sea $(X_n)_{n\in\mathbb{N}}$ una martingala y sea $\{\tau_n\}_{n=1}^{\infty}$ una sucesión creciente de tiempos de paro tales que para cada $n \in \mathbb{N}$ $\tau_n \leq n$, entonces*

$$(X_{\tau_n})_{n\in\mathbb{N}} \quad \textit{es martingala.}$$

Observemos que aunque no se mencione la filtración, se trata de la filtración " natural ".

El siguiente resultado es de gran importancia para el estudio de la convergencia *c.s.* de la sucesión $\{S_n\}_{n=1}^{\infty}$. Este resultado no se demuestra para no hacer más extenso este trabajo, pero su prueba se puede encontrar en [21].

**TEOREMA 1.5.2** *( Teorema de convergencia para Martingalas )*
*Sea $(X_n)_{n\in\mathbb{N}}$ una martingala tal que*

$$\sup_n\{E[|X_n|]\} < \infty.$$

*Entonces la sucesión $\{X_n\}_{n=1}^{\infty}$ converge casi seguramente a una variable áleatoria $X \in L^1(\Omega, \mathcal{F}, P)$.*

# Capítulo 2

# Convergencia c.s. de una Serie Aleatoria

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias definidas en este. En este capítulo estudiaremos la convergencia de la sucesión $\{S_n\}_{n=1}^{\infty}$ cuando las $X_n$'s son independientes. Cabe señalar que durante todo este capítulo $S_n$ representará la suma parcial y $S$ la serie total de variables aleatorias, es decir,

$$S_n = \sum_{k=1}^{n} X_k \quad y \quad S = \sum_{n=1}^{\infty} X_n$$

## 2.1 Convergencia de $S_n$ si $E[X_k] = 0$

En esta sección se trabajará con sucesiones de variables aleatorias independientes las cuales tienen esperanza cero. Este es un caso muy interesante porque bajo estas condiciones la sucesión $(S_n)_{n \in \mathbb{N}}$ resulta ser una martingala, ( *ejemplo 1.5.1* ).

**TEOREMA 2.1.1**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que $E[X_n] = 0$ para toda $n$. Si*

$$\sum_{n=1}^{\infty} Var(X_n) < \infty,$$

*entonces $\{S_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias que converge c.s. a una variable aleatoria $S \in L^1(\Omega, \mathcal{F}, P)$.*

*Demostración*

Ya que $E[X_n] = 0$, entonces $(S_n)_{n\in\mathbb{N}}$ es martinagala ( *ejemplo 1.5.1* ). Además como

$$\begin{aligned} E[|S_n|] &\leq (E[S_n^2])^{\frac{1}{2}} \quad (D.\ de\ Schwarz) \\ &= (E[(\sum_{k=1}^{n} X_k)^2])^{\frac{1}{2}} \\ &= (\sum_{k=1}^{n} E[X_k^2])^{\frac{1}{2}} \quad (Independencia\ y\ E[X_k] = 0) \\ &\leq (\sum_{n=1}^{\infty} E[X_n^2])^{\frac{1}{2}} \\ &= (\sum_{n=1}^{\infty} Var(X_n))^{\frac{1}{2}} < \infty. \end{aligned}$$

Entonces

$$\sup_n \{E[|S_n|]\} < \infty$$

De esta forma, del *teorema 1.5.2* se tiene que

$$S_n \xrightarrow{c.s.} S = \sum_{n=1}^{\infty} X_n \quad y \quad S \in L^1(\Omega, \mathcal{F}, P).$$

□

**COROLARIO 2.1.1**

*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes. Si*

$$\sum_{n=1}^{\infty} Var(X_n) < \infty$$

*entonces*

$$\sum_{n=1}^{\infty} (X_n - E[X_n]) \quad converge\ c.s.$$

*Demostración*
Si tomamos $\{X'_n = X_n - E[X_n]\}_{n=1}^{\infty}$, esta sucesión cumple con las hipótesis del teorema anterior. $\square$

**Observación 2.1.1**
El teorema anterior es un caso particular de un resultado más general el cual nos dice que:
Dada $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que para cada $n \in \mathbb{N}$ $E[X_n] = 0$, y

$$\sum_{n=1}^{\infty} E[|X_n|^p] < \infty \quad \textit{para alguna } 0 < p \leq 2,$$

entonces

$$S_n \xrightarrow{c.s.} S.$$

Este resultado se probará en la última sección de este capítulo.

El siguiente ejemplo es una aplicación interesante del teorema anterior.

**Ejemplo 2.1.1**
Recordemos que en un curso de Cálculo se demuestra que la serie $\sum_{n=1}^{\infty} \frac{1}{n}$ diverge, pero que la serie $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ converge condicionalmente. Ahora analicemos el siguiente caso:
Considere el experimento de lanzar una moneda una infinidad de veces y sean $\{X_n\}_{n=1}^{\infty}$ la sucesión de variables aleatorias definidas como:

$$X_n = \begin{cases} 1 & \text{si en la n-ésima tirada sale aguila} \\ -1 & \text{en otro caso} \end{cases}$$

y la suma aleatoria

$$S_n = \sum_{k=1}^{n} Y_k \quad \textit{donde} \quad Y_k = \frac{X_k}{k}$$

además supongamos que $p = P(X_n = 1)$, $P(X_n = -1) = q = 1 - p$ con $p \in (0,1)$ y que las $X_1, X_2, \cdots, X_n, \cdots$ son independientes. Nuestra pregunta es la siguiente:

¿ Bajo que condiciones de $p$ la sucesión $\{S_n\}_{n=1}^{\infty}$ converge *c.s.* ?.
Para el caso $p = \frac{1}{2}$, basados en nuestra " intuición ", afirmaríamos que $S_n$ converge *c.s.* , lo cual es cierto. Demostremos esta afirmación.
Para $p = \frac{1}{2}$ se tiene que

$$E[Y_n] = \frac{1}{n}[1P(X_n = 1) + (-1)P(X_n = -1)] = 0 \quad \forall n.$$

y también

$$Var(Y_n) = \frac{1}{n^2}[1P(X_n = 1) + 1P(X_n = -1)] = \frac{1}{n^2} \quad \forall n.$$

entonces

$$\sum_{n=1}^{\infty} Var(Y_n) = \sum_{n=1}^{\infty}(\frac{1}{n^2}) < \infty.$$

Por lo tanto, del *teorema 2.1.1* se tiene que

$$\sum_{n=1}^{\infty} Y_n = \sum_{n=1}^{\infty} \frac{X_n}{n} \quad \textit{converge c.s.}$$

□

Ahora bien, ¿ Qué pasa para el caso $p \in (0,1)$ $y$ $p \neq \frac{1}{2}$ ?. En este caso, se tiene que $\{S_n\}_{n=1}^{\infty}$ diverge *c.s.* , pero esto se demostrará en la siguiente sección.

Otra aplicación directa del teorema de convergencia para martingalas y de la teoría de tiempos de paro es el siguiente resultado.

**TEOREMA 2.1.2**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que*

$$E[X_n] = 0 \quad \forall n \quad y \quad E[\sup_n X_n] < \infty$$

*entonces $S_n$ converge sobre el conjunto $A = \{\omega \in \Omega \ / \ \sup_n S_n(\omega) < \infty\}$ casi seguramente.*

*Demostración*
Bajo estas condiciones se sabe que $(S_n)_{n \in \mathbb{N}}$ es martingala.

Sea $M \in I\!N$ y sea $t_M(\omega) = \inf\{n \in I\!N \ / \ S_n(\omega) > M\}$, observemos que si $\{EI \ I\!N \ / \ S_n(\omega) > M\} = \emptyset$ entonces $t_M(\omega) = \infty$. Es claro que por el *ejemplo 1.4.1* , $t_M$ es tiempo de paro. Sea

$$\tau_n = \min\{t_M, n\}$$

como $\tau_n$ también es tiempo de paro para toda n, y

$$\tau_n \leq \tau_{n+1} \leq n+1 \quad \forall n.$$

entonces, por el *corolario 1.5.1*

$$(S_{\tau_n}, \mathcal{F}_{\tau_n})_{n \in N} \quad \textit{es martingala.}$$

Ahora bien, como $t_M$ es el ínfimo y $S_{t_M} > M$, entonces
si $t_M > n$ entonces $S_{\tau_n} = S_n \leq M$ y por lo tanto

$$(S_{\tau_n})^+ = \max\{0, S_{\tau_n}\} \leq M \quad \textit{si } t_M > n,$$

si $t_M \leq n$ entonces $S_{\tau_n} = S_{t_M} = S_{t_M - 1} + X_{t_M} \leq M + \sup_n X_n$, por lo tanto

$$\begin{aligned}(S_{\tau_n})^+ &= \max\{0, S_{\tau_n}\} \\ &\leq \max\{0, M + \sup_n X_n\} \quad \textit{si } t_M \leq n,\end{aligned}$$

luego entonces

$$\begin{aligned}E[(S_{\tau_n})^+] &= E[(S_{\tau_n})^+ I_{\{t_M > n\}}] + E[(S_{\tau_n})^+ I_{\{t_M \leq n\}}] \\ &\leq M + E[\max\{0, M + \sup_n X_n\}] < \infty.\end{aligned}$$

Por lo tanto

$$\sup_n\{E[(S_{\tau_n})^+]\} < \infty.$$

Por otro lado, se tiene que

$$\begin{aligned}E[|S_{\tau_n}|] &= E[(S_{\tau_n})^+] + E[(S_{\tau_n})^-] \\ &\leq E[(S_{\tau_n})^+] + E[(S_{\tau_n})^+ - S_{\tau_n}] \\ &= 2E[(S_{\tau_n})^+] - C.\end{aligned}$$

Por lo tanto

$$\begin{aligned}\sup_n\{E[|S_{\tau_n}|]\} &\leq 2\sup_n\{E[(S_{\tau_n})^+]\} - C \\ &< \infty.\end{aligned}$$

Así, por el *teorema 1.5.2* ,

$$S_{\tau_n} \xrightarrow{c.s.} S$$

pero $S_{\tau_n} = S_n$ sobre el conjunto $\{t_M = \infty\}$. Entonces, sobre cada conjunto $A_M = \{t_M = \infty\}$ el $\lim_{n\to\infty} S_n$ existe *c.s.* , es decir

$$\exists\, N_M \in \mathcal{F} \text{ tal que } \forall\, \omega \in (A_M - N_M) \text{ el } \lim_{n\to\infty} S_n(\omega) \text{ existe y } P(N_M) = 0$$

además, como $A = \bigcup_{M=1}^{\infty} A_M$ y si tomamos $N = \bigcup_{M=1}^{\infty} N_M$, entonces

$$\forall\, \omega \in (A \setminus N) \text{ el } \lim_{n\to\infty} S_n(\omega) \text{ existe y } P(N) = 0.$$

Por lo tanto $S_n$ converge sobre A *c.s.* □

El siguiente teorema nos relaciona la convergencia de una serie de números reales positivos con la convergencia de $S_n$.

**TEOREMA 2.1.3**
*Sean $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes y $\{b_n\}_{n=1}^{\infty}$ una sucesión de números reales positivos tales que $E[X_n] = 0$ para toda $n$, y $\sum_{n=1}^{\infty} b_n < \infty$. Si*

$$\sum_{n=1}^{\infty}(b_n)^{1-\frac{p}{2}} E[|X_n|^p] < \infty \quad \textit{para alguna } p \geq 2,$$

*entonces*

$$S_n \xrightarrow{c.s.} S$$

*Demostración*
Supongamos que $p > 2$, ya que para el caso $p = 2$ se tiene por el *teorema 2.1.1*. Vamos a demostrar que bajo las hipótesis del teorema,

$$\sum_{n=1}^{\infty}(b_n)^{1-\frac{p}{2}} E[|X_n|^p] < \infty \Longrightarrow \sum_{n=1}^{\infty} E[X_n^2] < \infty \quad \forall\, p > 2$$

y por el *teorema 2.1.1* tendremos el resultado.
Sea $C_n = (E[|X_n|^p])^{\frac{2}{p}}$. Entonces

$$\begin{aligned}
b_n < C_n &\iff (b_n)^{\frac{p}{2}} < (C_n)^{\frac{p}{2}} \\
&\iff (C_n)^{-\frac{p}{2}} < (b_n)^{-\frac{p}{2}} \\
&\iff ((C_n)^{-\frac{p}{2}})^{1-\frac{2}{p}} < ((b_n)^{-\frac{p}{2}})^{1-\frac{2}{p}} \quad (\textit{ ya que } 1 - \frac{2}{p} > 0\ ) \\
&\iff (C_n)^{1-\frac{p}{2}} < (b_n)^{1-\frac{p}{2}}.
\end{aligned}$$

Y en consecuencia

$$\begin{aligned}
(E[|X_n|^p])^{\frac{2}{p}} &= C_n[I_{\{C_n \le b_n\}} + I_{\{C_n > b_n\}}] \\
&\le b_n + C_n I_{\{C_n > b_n\}} \\
&= b_n + (C_n)^{1-\frac{p}{2}}(C_n)^{\frac{p}{2}} I_{\{C_n > b_n\}} \\
&\le b_n + (b_n)^{1-\frac{p}{2}}(C_n)^{\frac{p}{2}} \\
&= b_n + (b_n)^{1-\frac{p}{2}} E[|X_n|^p]
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
\sum_{n=1}^{\infty}(E[|X_n|^p])^{\frac{2}{p}} &\le \sum_{n=1}^{\infty} b_n + \sum_{n=1}^{\infty}(b_n)^{1-\frac{2}{p}} E[|X_n|^p] \\
&< \infty \qquad (\textit{ por hipótesis }\ )
\end{aligned}$$

Ahora bien, por la desigualdad de Jensen tenemos que

$$\begin{aligned}
\sum_{n=1}^{\infty} E[|X_n|^2] &= \sum_{n=1}^{\infty}((E[|X_n|^2])^{\frac{p}{2}})^{\frac{2}{p}} \\
&\le \sum_{n=1}^{\infty}(E[|X_n|^p])^{\frac{2}{p}} \quad (\textit{ ya que } \frac{p}{2} > 1\ ) \\
&< \infty
\end{aligned}$$

Por lo tanto

$$\sum_{n=1}^{\infty} E[|X_n|^2] < \infty.$$

□

**Ejemplo 2.1.2**
Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes todas con distribución $X \sim Normal(\mu, \sigma^2)$ donde $\mu = 0$ y $0 < \sigma^2 < \infty$, es decir,

$$\begin{aligned} P(X_n \leq x) &= P(X \leq x) \\ &= \int_{-\infty}^{x} (2\pi\sigma^2)^{-\frac{1}{2}} e^{(-x^2/2\sigma^2)}\, dx \quad \forall\, x \in \Re. \end{aligned}$$

Entonces $\sum_{n=1}^{\infty} \frac{X_n}{n}$ converge *c.s.*
*Demostración*
Se puede demostrar que si $X$ es una variable aleatoria Normal con parámetros $\mu$ y $0 < \sigma^2 < \infty$, entonces (ver [1] para detalles)

$$E[|X_n|^p] = E[|X|^p] < \infty \quad \forall 0 < p < \infty.$$

y en particular se cumple si $p \geq 2$. Además, $E[\frac{X_n}{n}] = \frac{1}{n}E[X_n] = 0$ para toda n, y si tomamos $b_n = \frac{1}{n^2}$, se tiene que

$$\begin{aligned} \sum_{n=1}^{\infty} b_n &< \infty \quad y \\ \sum_{n=1}^{\infty} (b_n)^{1-\frac{p}{2}} E[|\frac{X_n}{n}|^p] &= \sum_{n=1}^{\infty} (\frac{1}{n})^{2-p} (\frac{1}{n})^p E[|X_n|^p] \\ &= E[|X|^p] \sum_{n=1}^{\infty} (\frac{1}{n})^2 \quad \forall p \geq 2 \\ &< \infty. \end{aligned}$$

Por lo tanto, del teorema anterior, se tiene que

$$\sum_{n=1}^{\infty} \frac{X_n}{n} \quad \textit{converge c.s.}$$

$\square$

Los pasos que conducen a nuestro siguiente teorema y a su demostración concentran nuestra atención en la definición de integrabilidad uniforme para la sucesión $\{S_n\}_{n=1}^{\infty}$, ya que, si $\{S_n\}_{n=1}^{\infty}$ es uniformemente integrable, entonces por el *teorema 1.3.1* se tiene que

$$\sup_n\{E[|S_n|]\} < \infty,$$

y por lo tanto

$$S = \sum_{n=1}^{\infty} X_n \text{ } \textit{converge c.s} \quad (\textit{ teorema } 1.5.2 \text{ }).$$

En otras palabras, todas las condiciones sobre $S_n$, las cuales den como resultado que $\{S_n\}_{n=1}^{\infty}$ es uniformemente integrable, implican que $S_n \xrightarrow{c.s.} S$. Existen una gran variedad de dichas condiciones, pero solamente mencionaremos dos de gran importancia.

**TEOREMA 2.1.4**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que $E[X_n] = 0$ para cada $n \in I\!N$. Cada una de las condiciones siguientes implican que $S_n \xrightarrow{c.s.} S$.*
*a) Si existe $Y \in L^1(\Omega, \mathcal{F}, P)$ tal que $|S_n| \leq Y$ para toda $n$.*
*b) Si existe $G : \Re^+ \longrightarrow \Re^+$ tal que*

$$\lim_{t\to\infty} \frac{G(t)}{t} = \infty \quad y \quad \sup_n\{E[G(|S_n|)]\} < \infty$$

*Demostración*
Vamos a demostrar que cada una de estas condiciones implican que $\{S_n\}_{n=1}^{\infty}$ es uniformemente integrable y por lo mencionado anteriormente

$$S_n \xrightarrow{c.s.} S$$

a) Sea $M > 0$, entonces

$$\begin{aligned} \int_{\{|S_n|\geq M\}} |S_n| \, dP &\leq \int_{\{|S_n|\geq M\}} Y \, dP \\ &\leq \int_{\{Y\geq M\}} Y \, dP \end{aligned}$$

en consecuencia

$$\sup_n\{\int_{\{|S_n|\geq M\}} |S_n|\, dP\} \leq \int_{\{Y\geq M\}} Y\, dP,$$

y de esta forma se obtiene que

$$\lim_{M\to\infty}(\sup_n\{\int_{\{|S_n|\geq M\}} |S_n|\, dP\}) \leq \lim_{M\to\infty}(\int_{\{Y\geq M\}} Y\, dP) = 0.$$

Por lo tanto $\{S_n\}_{n=1}^{\infty}$ es U.I.

b) Sea $\epsilon > 0$, existe $N_\epsilon \in I\!N$ tal que $\frac{1}{N_\epsilon} < \epsilon$
Para $N_\epsilon$ existe $M_0 \in \Re^+$ tal que

$$N_\epsilon \leq \frac{G(t)}{t} \quad (\, t \leq \frac{G(t)}{N_\epsilon}\,) \quad \forall t \geq M_0$$

entonces

$$\begin{aligned}
\int_{\{|S_n|\geq M_0\}} |S_n|\, dP &\leq \int_{\{|S_n|\geq M_0\}} \frac{G(|S_n|)}{N_\epsilon}\, dP \\
&\leq \frac{1}{N_\epsilon}\int_\Omega G(|S_n|)\, dP \\
&\leq \frac{1}{N_\epsilon}\sup_n\{E[G(|S_n|)]\} \\
&\leq \epsilon K \quad (\, K < \infty\,)
\end{aligned}$$

y además sabemos que

$$\int_{\{|S_n|\geq M\}} |S_n|\, dP \leq \int_{\{|S_n|\geq M_0\}} |S_n|\, dP \quad \forall\, M > M_0 \;y\; \forall\, n.$$

Así, dada $\epsilon' = \frac{\epsilon}{K} > 0$, existe $M_0 \in \Re^+$ tal que

$$\begin{aligned}
\sup_n\{\int_{\{|S_n|\geq M\}} |S_n|\, dP\} &\leq \sup_n\{\int_{\{|S_n|\geq M_0\}} |S_n|\, dP\} \\
&\leq \epsilon' K \\
&= \epsilon \quad \forall\, M \geq M_0
\end{aligned}$$

Por lo tanto $\{S_n\}_{n=1}^{\infty}$ es U.I. □

**COROLARIO 2.1.2**
*Sea* $\{X_n\}_{n=1}^{\infty}$ *una sucesión de variables aleatorias independientes tales que* $E[X_n] = 0$ *para cada* $n \in \mathbb{N}$. *Si*

$$\sup_n\{E[|S_n|^p]\} < \infty \quad \textit{para alguna } p \geq 1.$$

*Entonces*

$$S_n \xrightarrow{c.s.} S \quad y \quad S \in L^1(\Omega, \mathcal{F}, P)$$

*Demostración*
Para el caso $p = 1$ se tiene por el *teorema 2.1.1* . Para el caso $p > 1$, sea $G : \Re^+ \longrightarrow \Re^+$ definida como $G(x) = x^p$, en tal caso $G(x)$ cumple con las hipótesis del teorema anterior ( inciso (b) ). □

## 2.2 Convergencia de $S_n$ si $|X_k| \leq C$ *c.s.*

En la presente sección se trabajará con sucesiones de variables aleatorias independientes y acotadas casi seguramente.

Nuestro primer resultado nos da el " recíproco " del corolario.2.1.1, pero bajo la condición de que la sucesión sea acotada por una constante positiva.

**TEOREMA 2.2.1**
*Sea* $\{X_n\}_{n=1}^{\infty}$ *una sucesión de variables aleatorias independientes tales que* $|X_n| \leq C$ *c.s. para* $n \geq 1$ *y para alguna* $C > 0$. *Si*

$$\sum_{n=1}^{\infty}(X_n - E[X_n]) \textit{ converge c.s.}$$

*Entonces*

$$\sum_{n=1}^{\infty} Var(X_n) < \infty$$

*Demostración*
Sean $n \in \mathbb{N}$, $\epsilon > 0$ y $Y_1 = X_{n+1}, Y_2 = X_{n+2}, \cdots, Y_k = X_{n+k}, \cdots$ tomemos $S_k = \sum_{i=1}^{k} Y_i$ para $k = 1, 2, \cdots$, entonces

$$
\begin{aligned}
P(\max_{1\le k\le m}|S_k - E[S_k]| > \epsilon) &\geq 1 - \frac{(2C+\epsilon)^2}{Var(S_m)} \quad \forall m. \quad (\textit{teorema } 1.2.2)\\
&= 1 - \frac{(2C+\epsilon)^2}{\sum_{i=1}^{m} Var(Y_i)}\\
&= 1 - \frac{(2C+\epsilon)^2}{\sum_{i=n+1}^{n+m} Var(X_i)} \quad \forall m.
\end{aligned}
$$

Pero

$$
\begin{aligned}
P(\max_{1\le k\le m}|S_k - E[S_k]| > \epsilon) &\leq P(\sup_{k\ge 1}|S_k - E[S_k]| > \epsilon)\\
&= P(\sup_{k\ge 1}|\sum_{i=n+1}^{n+k}(X_i - E[X_i])| > \epsilon)\\
&= P(\sup_{k\ge 1}|T_{n+k} - T_n| > \epsilon)\\
&= P(\bigcup_{k=1}^{\infty}\{|T_{n+k} - T_n| > \epsilon\})
\end{aligned}
$$

donde $T_k = \sum_{i=1}^{k}(X_i - E[X_i])$ para $k = 1, 2, \cdots$. Entonces

$$
P(\bigcup_{k=1}^{\infty}\{T_{n+k} - T_n| > \epsilon\}) \geq 1 - \frac{(2C+\epsilon)^2}{\sum_{i=n+1}^{n+m} Var(X_i)} \quad \forall m.
$$

Por lo tanto

$$
P(\bigcup_{k=1}^{\infty}\{|T_{n+k} - T_n| > \epsilon\}) \geq 1 - \frac{(2C+\epsilon)^2}{\sum_{i=n+1}^{\infty} Var(X_i)}
$$

Ahora bien, si suponemos que $\sum_{n=1}^{\infty} Var(X_n) = \infty$, entonces

$$
P(\bigcup_{k=1}^{\infty}\{|T_{n+k} - T_n| > \epsilon\}) = 1 \quad \forall \epsilon > 0, \quad \forall n.
$$

Por lo tanto $\{T_n\}_{n=1}^{\infty}$ no es de cauchy *c.s.* ( *teorema 1.1.2* ), y en consecuencia

$$
T_n = \sum_{i=1}^{n}(X_i - E[X_i]) \quad \textit{diverge c.s} \;!!
$$

De esta forma se obtiene el resultado. □

**Observación 2.2.1**
Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que $|X_n| \leq C$ *c.s.* para $n \geq 1$ y para alguna constante $C > 0$. Entonces, del *corolario 2.1.1* y del teorema anterior tenemos que

$$\sum_{n=1}^{\infty}(X_n - E[X_n]) \text{ converge c.s.} \Longleftrightarrow \sum_{n=1}^{\infty} Var(X_n) \text{ converge}$$

**TEOREMA 2.2.2**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que $|X_n| \leq C$ c.s. para $n \geq 1$ y $C > 0$. Entonces*

$$\sum_{n=1}^{\infty} X_n \text{ converge c.s.}$$

*si y sólo si las dos siguientes condiciones se cumplen:*
*a)* $|\sum_{n=1}^{\infty} E[X_n]| < \infty$
*b)* $\sum_{n=1}^{\infty} Var(X_n) < \infty$

*Demostración*
Primero supongamos que $\sum_{n=1}^{\infty} X_n$ converge *c.s.*
Para cada $n \in \mathbb{N}$ sea $(\Omega', \mathcal{F}, P_n)$ donde $\Omega' = \Re$, $\mathcal{F} = \mathbb{B}(\Re)$ y
$P_n = \lambda_{F_{X_n}}$ ( Medida inducida por $F_{X_n}$ )
definimos $X_n' : \Omega' \longrightarrow \Re$ como $X_n'(\omega) = \omega$. Entonces

$$\begin{aligned} F_{X_n'}(x) &= P_n(X_n' \leq x) \\ &= P_n((-\infty, x]) \\ &= \lambda_{F_{X_n}}((-\infty, x]) \\ &= F_{X_n}(x) - F_{X_n}(-\infty) \\ &= F_{X_n}(x). \end{aligned}$$

Por lo tanto

$$F_{X_n} \equiv F_{X_n'}$$

Así, hemos construido una sucesión $\{X_n'\}_{n=1}^{\infty}$ de variables aleatorias independientes con las siguientes propiedades:

a) $X_n$ y $X'_n$ tienen la misma distribución para cada $n \in \mathbb{N}$.
b) La sucesión combinada $\{X_n, X'_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias independientes, es decir, para todo $n \neq m$ y para todo $A, B \in \mathbb{B}(\Re)$, se tiene que

$$P((X_n, X'_n) \in A, (X_m, X'_m) \in B) = P((X_n, X'_n) \in A)P((X_m, X'_m) \in B).$$

Tomemos $Y_n = X_n - X'_n$ para $n = 1, 2, \cdots$ Es claro que $\{Y_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias independientes tales que,

$$\begin{aligned} |Y_n| &\leq |X_n| + |X'_n| \leq 2C \quad c.s. \\ E[Y_n] &= E[X_n] - E[X'_n] = 0 \quad y \\ Var(Y_n) &= 2Var(X_n) \quad \forall\, n. \end{aligned}$$

Además, como $\sum_{n=1}^{\infty} X_n$ converge *c.s.* entonces $\sum_{n=1}^{\infty} X'_n$ converge *c.s.* y en consecuencia

$$\sum_{n=1}^{\infty} Y_n \textit{ converge c.s.}$$

De la *observación 2.2.1* se tiene que

$$\sum_{n=1}^{\infty} Var(Y_n) < \infty$$

Por lo tanto

$$\sum_{n=1}^{\infty} Var(X_n) = \frac{1}{2} \sum_{n=1}^{\infty} Var(Y_n) < \infty.$$

Y del *corolario 2.1.1* podemos concluir que

$$\sum_{n=1}^{\infty} (X_n - E[X_n]) \textit{ converge c.s.}$$

Y por lo tanto

$$\sum_{n=1}^{\infty} E[X_n] = \sum_{n=1}^{\infty} X_n - \sum_{n=1}^{\infty} (X_n - E[X_n]) \textit{ converge c.s.}$$

Recíprocamente, supongase que pasa (a) y (b). Entonces, por el *corolario 2.1.1*,

$$\sum_{n=1}^{\infty}(X_n - E[X_n]) \textit{ converge c.s.}$$

Por lo tanto

$$\sum_{n=1}^{\infty} X_n = \sum_{n=1}^{\infty}(X_n - E[X_n]) + \sum_{n=1}^{\infty} E[X_n] \textit{ converge c.s.}$$

□

**Observación 2.2.2**
Con esta teoría estamos en condiciones de demostrar la otra parte del ejemplo 2.1.1. Recordemos que en la sección anterior se demostró que

$$\sum_{n=1}^{\infty} Y_n = \sum_{n=1}^{\infty} \frac{X_n}{n} \quad \textit{converge c.s. para } p = \frac{1}{2}.$$

Ahora bien, para el caso $p \in (0,1)$ *y* $p \neq \frac{1}{2}$ vamos a probar que $\sum_{n=1}^{\infty} Y_n$ diverge c.s.
*Demostración*
Sea $p \in (0,1)$ *y* $p \neq \frac{1}{2}$, entonces

$$E[Y_n] = \frac{1}{n}[1P(X_n = 1) + (-1)P(X_n = -1)] = \frac{1}{n}[p + (-1)(1-p)] = \frac{1}{n}[2p-1]$$

y también

$$Var(Y_n) = \frac{1}{n^2}[E[(X_n)^2] - (E[X_n])^2] = \frac{1}{n^2}[1 - (2p-1)^2] = \frac{4}{n^2}[p - p^2].$$

Por lo tanto

$$\sum_{n=1}^{\infty} Var(Y_n) = \sum_{n=1}^{\infty} \frac{4}{n^2}[p - p^2] < \infty,$$

Pero

$$\sum_{n=1}^{\infty} E[Y_n] = \sum_{n=1}^{\infty} \frac{1}{n}[2p-1] = \infty.$$

Y del *teorema 2.2.2* , se tiene que

$$\sum_{n=1}^{\infty} Y_n = \sum_{n=1}^{\infty} \frac{X_n}{n} \quad \textit{diverge c.s.}$$

□

## 2.3 Teorema de Kolmogorov

En esta última sección probaremos el Teorema de tres series de Kolmogorov para la convergencia de la sucesión $\{S_n\}_{n=1}^{\infty}$. Este teorema es uno de los resultados más importantes en el estudio de la convergencia de la sucesión $\{S_n\}_{n=1}^{\infty}$. Dicho resultado nos da condiciones necesarias y suficientes para que $S_n$ converja *c.s.* , además nos permite establecer otros resultados vinculados con dicha convergencia.

**TEOREMA 2.3.1** *( Tres series de Kolmogorov )*
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Entonces la serie $\sum_{n=1}^{\infty} X_n$ converge c.s. si sólo si para alguna $c > 0$, las siguientes 3 series convergen*
*a)* $\sum_{n=1}^{\infty} P(|X_n| \geq c)$
*b)* $\sum_{n=1}^{\infty} E[X_n^c]$
*c)* $\sum_{n=1}^{\infty} Var(X_n^c)$
*Además, si (a), (b), y (c) convergen para alguna $c > 0$, entonces convergen para toda $c > 0$.*

*Demostración*
Supongamos que $\sum_{n=1}^{\infty} X_n$ converge *c.s.* Entonces $X_n \xrightarrow{c.s.} 0$, y del *corolario 1.1.3* se tiene que

$$\sum_{n=1}^{\infty} P(|X_n| \geq c) < \infty \quad \forall c > 0.$$

Además, como $\sum_{n=1}^{\infty} P(\{X_n \neq X_n^c\}) = \sum_{n=1}^{\infty} P(|X_n| \geq c) < \infty$ y en vista del *teorema 1.2.1*

$$\sum_{n=1}^{\infty} X_n \quad y \quad \sum_{n=1}^{\infty} X_n^c \quad \textit{son equivalentemente convergentes}$$

Por lo tanto

$$\sum_{n=1}^{\infty} X_n^c \textit{ converge c.s.} \quad \forall c > 0$$

En consecuencia (a) y (b) son inmediatas del *teorema 2.2.2* .
Inversamente, si (b) y (c) convergen para alguna $c > 0$, del *teorema 2.2.2* se tiene que $\sum_{n=1}^{\infty} X_n^c$ converge y como (a) converge, entonces

$$\sum_{n=1}^{\infty} X_n \quad y \quad \sum_{n=1}^{\infty} X_n^c \quad \textit{son equivalentemente convergentes}$$

Por lo tanto

$$\sum_{n=1}^{\infty} X_n \textit{ converge c.s.}$$

□

**COROLARIO 2.3.1**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que $\sum_{n=1}^{\infty}(X_n)^2 < \infty$ c.s. Entonces*

$$S_n \xrightarrow{c.s.} S \Longleftrightarrow \sum_{n=1}^{\infty} E[X_n I_{\{|X_n|<1\}}] < \infty$$

*Demostración*
Supongamos que $\sum_{n=1}^{\infty}(X_n)^2$ converge *c.s.* , entonces por el teorema anterior se tiene que

$$\sum_{n=1}^{\infty} E[(X_n)^2 I_{\{|X_n|<1\}}] = \sum_{n=1}^{\infty} E[(X_n)^2 I_{\{|X_n|^2<1\}}] < \infty$$

y por lo tanto

$$\sum_{n=1}^{\infty} Var(X_n I_{\{|X_n|<1\}}) < \infty.$$

También por el mismo teorema se tiene

$$\sum_{n=1}^{\infty} P(|X_n| \geq 1) = \sum_{n=1}^{\infty} P(|X_n|^2 \geq 1) < \infty.$$

Por lo tanto, aplicando otra vez el teorema anterior se tiene que

$$S_n \xrightarrow{c.s.} S \iff \sum_{n=1}^{\infty} E[X_n I_{\{|X_n|<1\}}] < \infty$$

□

Ahora, usaremos este teorema para demostrar la Observación.2.1.1. Este resultado, como ya se dijo, es una generalización del teorema.2.1.1.

**TEOREMA 2.3.2**
*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes tales que $E[X_n] = 0$ para toda n. Si*

$$\sum_{n=1}^{\infty} E[|X_n|^p] < \infty \quad \textit{para alguna } 0 < p \leq 2$$

*Entonces*

$$S_n \xrightarrow{c.s.} S$$

*Demostración*
Vamos a demostrar que estas hipótesis implican (a), (b) y (c) del teorema anterior.
Sea $c \geq 0$
a)Para $p > 0$

$$\begin{aligned} P(|X_n| \geq c) &= E[I_{\{|X_n| \geq c\}}] \\ &\leq \frac{E[|X_n|^p I_{\{|X_n| \geq c\}}]}{c^p} \quad (\textit{ por que } 1 \leq \frac{|X_n|}{c}) \\ &\leq \frac{E[|X_n|^p]}{c^p}. \end{aligned}$$

Por lo tanto

$$\sum_{n=1}^{\infty} P(|X_n| \geq c) \leq \sum_{n=1}^{\infty} \frac{E[|X_n|^p]}{c^p} < \infty$$

b) La demostración de este inciso la vamos a hacer por casos.
Primer caso : si $p > 1$,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{|E[X_n^c]|}{c} &= \sum_{n=1}^{\infty} \frac{|E[X_n I_{\{|X_n|<c\}}]|}{c} \\
&= \sum_{n=1}^{\infty} \frac{|E[X_n I_{\{|X_n|\geq c\}}]|}{c} \quad (\textit{ ya que } E[X_n]=0 \text{ }) \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|}{c} I_{\{|X_n|\geq c\}}] \quad (\textit{D. de Jensen}) \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p} I_{\{|X_n|\geq c\}}] \quad (\textit{ por que } 1 \leq \frac{|X_n|}{c} \text{ }) \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p}] \\
&< \infty.
\end{aligned}$$

Segundo caso : si $0 < p \leq 1$,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{|E[X_n^c]|}{c} &= \sum_{n=1}^{\infty} \frac{|E[X_n I_{\{|X_n|<c\}}]|}{c} \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|}{c} I_{\{|X_n|<c\}}] \quad (\textit{D. de Jensen}) \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p} I_{\{|X_n|<c\}}] \quad (\textit{ porque } \frac{|X_n|}{c} < 1 \text{ }) \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p}] \\
&< \infty.
\end{aligned}$$

Por lo tanto si $p > 0$, se tiene que

$$\begin{aligned}
|\sum_{n=1}^{\infty} \frac{E[X_n^c]}{c}| &\leq \sum_{n=1}^{\infty} \frac{|E[X_n^c]|}{c} \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p}] < \infty.
\end{aligned}$$

c) como $p \leq 2$, entonces

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{E[|X_n^c|^2]}{c} &= \sum_{n=1}^{\infty} E[\frac{|X_n|^2}{c} I_{\{|X_n|<c\}}] \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p} I_{\{|X_n|<c\}}] \quad ( \textit{porque } \frac{|X_n|}{c} < 1 ) \\
&\leq \sum_{n=1}^{\infty} E[\frac{|X_n|^p}{c^p}] \\
&< \infty,
\end{aligned}$$

y además

$$\sum_{n=1}^{\infty} (E[|X_n^c|])^2 \leq \sum_{n=1}^{\infty} E[|X_n^c|^2] < \infty,$$

Por lo tanto

$$\sum_{n=1}^{\infty} Var(X_n^c) = \sum_{n=1}^{\infty} E[|X_n^c|^2] - \sum_{n=1}^{\infty} (E[|X_n^c|])^2 < \infty.$$

Y por lo tanto $S_n \xrightarrow{c.s.} S$. □

**COROLARIO 2.3.2**

*Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes con esperanza finita. Si*

$$\sum_{n=1}^{\infty} E[|X_n|^p] < \infty \quad \textit{para alguna } 0 < p \leq 1.$$

*Entonces*

$$S_n \xrightarrow{c.s.} S$$

*Demostración*

Vamos a demostrar que se cumplen las condiciones (a), (b) y (c) del *teorema 2.3.1* .

Los incisos (a) y (c) estan demostrados en el inciso (a) y (c) del teorema anterior ( observese que no usamos la hipótesis de $E[X_n] = 0$ ), el inciso

(b) está demostrado en el segundo caso del inciso (b) del mismo teorema $(0 < p \leq 1)$. □

**Ejemplo 2.3.1**
Sean $X_1, X_2, \cdots, X_n, \cdots$ variables aleatorias independientes exponencial con parámetro $\lambda_n$, es decir,

$$P(X_n \leq x) = \int_0^x \lambda_n^{-1} e^{-\lambda_n x}\, dx \quad \forall\, x \geq 0.$$

Vamos a demostrar que si $\sum_{n=1}^{\infty} \lambda_n^{-1} < \infty$, entonces $\sum_{n=1}^{\infty} X_n$ converge *c.s.*
*Demostración*
Como

$$\begin{aligned} E[|X_n|] &= E[X_n] \\ &= \int_{\Re} \lambda_n(\lambda_n^{-1} e^{-\lambda_n x})\, dx \\ &= \int_0^{\infty} e^{-\lambda_n x}\, dx \\ &= \lambda_n^{-1}, \end{aligned}$$

entonces

$$\sum_{n=1}^{\infty} E[|X_n|] = \sum_{n=1}^{\infty} \lambda_n^{-1} < \infty.$$

por lo tanto, por el corolario anterior

$$\sum_{n=1}^{\infty} X_n \quad \textit{converge c.s.}$$

□

# Capítulo 3

# Algunos resultados elementales sobre la aproximación Poisson en una sucesión de ensayos Bernoulli

En este capítulo se estudiará el artículo: "Some elementary results on Poisson approximation in a sequence of Bernoulli trials". Este artículo presenta algunos resultados sobre la aproximación de ciertas variables aleatorias discretas a otra discreta y su "rápidez" de convergencia. Se utilizarán series aleatorias para un caso muy particular: $S_n = \sum_{k=1}^{n} X_k$ y para cada $k = 1, 2, \cdots$, $X_k$ es una variable aleatoria de Bernoulli, es decir, $X_k$ satisface:

$$\begin{aligned} P(X_k = 1) &= p_k \\ P(X_k = 0) &= 1 - p_k \quad \forall\, k. \end{aligned}$$

para alguna $p_k \in (0, 1)$.

Gracias a estas métricas podremos además dar aproximaciones interesantes, que nos permiten estudiar el comportamiento (en el límite cuando $n \to \infty$) de la distribución $S_n$ cuando $n \to \infty$. Es claro que estas series aleatorias no convergen *c.s.* y esto es cierto para cualquier valor de $p_k \in (0, 1)$ que se elija. En cambio podemos esperar que se de la convergencia en un sentido menos fuerte.

En la primera sección de este capítulo definiremos dos métricas en el espacio de medidas de probabilidad sobre $(\mathbb{N}, \mathbb{B}(\mathbb{N}))$. En la segunda sección demostraremos varias propiedades de estas métricas.

Después, en la tercera y cuarta sección se darán cotas sobre estas métricas para la aproximación Poisson a una suma de variables aleatorias independientes Bernoulli, y como consecuencia de esto se demostrará la aproximación clásica de Poisson.

Finalmente, en la última sección, estos resultados se generalizan para el caso de variables aleatorias dependientes.

## 3.1 Estructura básica de dos funciones distancia

Supongase que todas las variables aleatorias en esta sección, y en todo el capítulo, están definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Sea $X_1, X_2, \cdots, X_n \cdots$, una sucesión de variables aleatorias Bernoulli, es decir, $X_k$ toma los valores 0 ó 1 para toda $k \geq 1$ y sea $S_n = \sum_{k=1}^{n} X_k$ la variable aleatoria que representa el número de éxitos (o el número de $X_k$'s que toman el valor 1). Es un resultado clásico que la distribución de probabilidad de $S_n$ se aproxima a una distribución Poisson si $n$ es suficientemente grande y la probabilidad de éxito es suficientemente pequeña, (Ver [1]). Aquí veremos variaciones de este resultado.

Dado un $\lambda > 0$, $Y(\lambda)$ denotará en esta sección, y en todo este capítulo, a una variable aleatoria Poisson con parámetro o media igual a $\lambda$. Esto es,

$$P(Y(\lambda) \leq m) = \sum_{k=0}^{m} \frac{\lambda^k e^{-\lambda}}{k!} \quad m = 0, 1, 2, \cdots$$

además denotemos a $\mathcal{Z}$ como el espacio de todas las variables aleatorias sobre los enteros no negativos, es decir,

$$\mathcal{Z} = \{X : \Omega \longrightarrow N_+ \ / \ X \ es\ \mathcal{F} - medible\}.$$

donde $N_+$ denota al conjunto de todos los enteros no negativos y sea $\mathcal{P}(N_+)$ el conjunto de todas las medidas de probabilidad sobre $N_+$.

Para caracterizar la eficacia de esa aproximación, consideraremos varias métricas en $\mathcal{P}(N_+)$ para medir la " semejanza " entre 2 distribuciones de probabilidad sobre $N_+$.

**DEFINICIÓN 3.1.1**
*Sean $Q_1, Q_2 \in \mathcal{P}(N_+)$. Definimos*

$$\begin{aligned} d(Q_1,Q_2) &= \sup_{A\subset N_+} \{|Q_1(A) - Q_2(A)|\} \\ d_0(Q_1,Q_2) &= \sup_{m\geq 0}\{|Q_1(-\infty,m] - Q_2(-\infty,m]|\}. \end{aligned}$$

*En particular si $X$ y $Y$ son variables aleatorias que pertenecen a $\mathcal{Z}$, definimos*

$$\begin{aligned} d(P_X,P_Y) = d(X,Y) &= \sup_{A\subset N_+} \{|P(X\in A) - P(Y\in A)|\} \\ d_0(P_X,P_Y) = d_0(X,Y) &= \sup_{m\geq 0}\{|P(X\leq m) - P(Y\leq m)|\}. \end{aligned}$$

Cabe mencionar, que usaremos $d(X,Y)$ y $d_0(X,Y)$ para simplificar la notación, pero realmente se trata de una distancia en $\mathcal{P}(N_+)$, ya que, a cada variable aleatoria $X \in \mathcal{Z}$ se le asocia una única distribución de probabilidad $P_X$ la cual pertenece a el espacio $\mathcal{P}(N_+)$.

Es claro que $d_0(X,Y) \leq d(X,Y)$ para todo $X, Y \in \mathcal{Z}$ y además $d$ y $d_0$ dependen solamente de la distribución de $X$ y $Y$. $d$ y $d_0$ así definidos son métricas en $\mathcal{P}(N_+)$ ó en $\mathcal{Z}$. En efecto:
Dadas $X, Y, Z \in \mathcal{Z}$ y $P_X, P_Y$ y $P_Z$ sus distribuciones de probabilidad tenemos que:
i)

$$d(X,Y) \geq 0 \quad y \quad d(X,Y) = d(Y,X)$$

ii)

$$\begin{aligned} d(X,Y) = 0 &\iff \sup_{A\subset N_+} \{|P(X\in A) - P(Y\in A)|\} = 0 \\ &\iff |P(X\in A) - P(Y\in A)| = 0 \quad \forall A \subset N_+ \\ &\iff P(X\in A) = P(Y\in A) \quad \forall A \subset N_+ \\ &\iff P_X = P_Y. \end{aligned}$$

iii) Como

$|P(X \in A) - P(Y \in A)| \leq |P(X \in A) - P(Z \in A)| + |P(Z \in A) - P(Y \in A)|$,

entonces

$$\begin{aligned}
&\sup_{A \subset N_+} \{|P(X \in A) - P(Y \in A)|\} \\
&\leq \sup_{A \subset N_+} \{|P(X \in A) - P(Z \in A)| + |P(Z \in A) - P(Y \in A)|\} \\
&\leq \sup_{A \subset N_+} \{|P(X \in A) - P(Z \in A)|\} + \sup_{A \subset N_+} \{|P(Z \in A) - P(Y \in A)|\}.
\end{aligned}$$

Por lo tanto $d(X,Y) \leq d(X,Z) + d(Z,Y)$. □

Daremos dos expresiones alternativas para la métrica $d$ las que serán de gran ayuda.

**TEOREMA 3.1.1**

*Sean $X \in \mathcal{Z}$ y $Y \in \mathcal{Z}$, entonces*

*i) $d(X,Y) = d_1(X,Y) = \frac{1}{2}\sum_{n=0}^{\infty} |P(X = n) - P(Y = n)|$ y*

*ii) $d(X,Y) = d_2(X,Y) = \frac{1}{2}\sup_{|h|\leq 1} \{|E[h(X)] - E[h(Y)]|\}$*

*donde el supremo se toma sobre todas las funciones $h : \Re \longrightarrow \Re$ medibles tales que $|h| \leq 1$.*

*Demostración*

i)

$d_1(X,Y) \leq d(X,Y)$:

Sean $X \in \mathcal{Z}$ y $Y \in \mathcal{Z}$. Tomemos $p_n = P(X = n)$, $q_n = P(Y = n)$ y $J = \{n : p_n \geq q_n\}$, entonces

$$\begin{aligned}
2d_1(X,Y) &= \sum_{n=0}^{\infty} |p_n - q_n| \\
&= \sum_{n \in J} p_n - q_n + \sum_{n \in J^c} q_n - p_n \\
&\leq |\sum_{n \in J} p_n - q_n| + |\sum_{n \in J^c} p_n - q_n| \\
&= |P(X \in J) - P(Y \in J)| + |P(X \in J^c) - P(Y \in J^c)| \\
&\leq 2 \sup_{A \subset N_+} \{|P(X \in A) - P(Y \in A)|\} \\
&= 2d(X,Y).
\end{aligned}$$

Por lo tanto

$$d_1(X,Y) \leq d(X,Y).$$

$d(X,Y) \leq d_1(X,Y)$:
Sea $A \subset N_+$, entonces
$2|P(X \in A) - P(Y \in A)|$

$$\begin{aligned}
&= |P(X \in A) - P(Y \in A)| + |P(X \in A) - P(Y \in A)| \\
&= |P(X \in A) - P(Y \in A)| + |P(X \in A^c) - P(Y \in A^c)| \\
&= |\sum_{n \in A} p_n - q_n| + |\sum_{n \in A^c} p_n - q_n| \\
&\leq \sum_{n \in A} |p_n - q_n| + \sum_{n \in A^c} |p_n - q_n| \\
&= \sum_{n=0}^{\infty} |p_n - q_n| \\
&= 2d_1(X,Y),
\end{aligned}$$

y en consecuencia

$$2d(X,Y) = 2 \sup_{A \subset N_+} \{|P(X \in A) - P(Y \in A)|\} \leq 2d_1(X,Y).$$

Por lo tanto

$$d(X,Y) \leq d_1(X,Y).$$

De esta forma se obtiene que

$$d(X,Y) = d_1(X,Y).$$

ii) Vamos a demostrar que $d_2(X,Y) = d_1(X,Y)$.
$d_1(X,Y) \leq d_2(X,Y)$:
Sean $X \in \mathcal{Z}$ y $Y \in \mathcal{Z}$. Tomemos

$$p_n = P(X = n), \quad q_n = P(Y = n)$$

y

$$h(n) = \begin{cases} 1 & \text{si } p_n \geq q_n \\ -1 & \text{si } p_n < q_n \end{cases}$$

Entonces

$$
\begin{aligned}
2d_1(X,Y) &= \sum_{n=0}^{\infty} |p_n - q_n| \\
&= \sum_{n=0}^{\infty} (h(n)p_n - h(n)q_n) \\
&= \sum_{n=0}^{\infty} h(n)p_n - \sum_{n=0}^{\infty} h(n)q_n \\
&= E[h(X)] - E[h(Y)] \\
&\leq |E[h(X)] - E[h(Y)]| \\
&\leq \sup_{|h|\leq 1} \{|E[h(X)] - E[h(Y)]|\} \\
&= 2d_2(X,Y).
\end{aligned}
$$

Por lo tanto $d_1(X,Y) \leq d_2(X,Y)$.

$d_2(X,Y) \leq d_1(X,Y)$:
Sea $h : \Re \longrightarrow \Re$ tal que $|h| \leq 1$, entonces

$$
\begin{aligned}
2|E[h(X)] - E[h(Y)]| &= |\sum_{n=0}^{\infty} h(n)p_n - \sum_{n=0}^{\infty} h(n)q_n| \\
&= |\sum_{n=0}^{\infty} h(n)(p_n - q_n)| \\
&\leq \sum_{n=0}^{\infty} |h(n)||p_n - q_n| \\
&\leq \sum_{n=0}^{\infty} |p_n - q_n| \\
&= 2d_1(X,Y),
\end{aligned}
$$

y en consecuencia

$$2d_2(X,Y) = 2 \sup_{|h|\leq 1} \{|E[h(X)] - E[h(Y)]|\} \leq 2d_1(X,Y).$$

Así, $d_2(X,Y) \leq d_1(X,Y)$. Por lo tanto

$$d_2(X,Y) = d_1(X,Y).$$

□

Es importante señalar que estas expresiones alternativas para la métrica $d$ se han estudiado en otras áreas. Por ejemplo, $d(X,Y)$ conocida como la "norma de variación" y $2d(X,Y)$ como un operador normal sobre el espacio de medidas con signo, (Ver [22]).

## 3.2 Propiedades de $d$ y $d_0$

En esta sección probaremos algunas propiedades de las métricas $d$ y $d_o$. La primera propiedad nos da cotas superiores para estas métricas.

**TEOREMA 3.2.1**
*Sean $X$ y $Y$ variables aleatorias con valores en $N_+$. Entonces*
*a)* $d(X,Y) \leq P(X \neq Y)$.
*b)* $d_0(X,Y) \leq \max\{P(X < Y), P(X > Y)\}$.

*Demostración*
a) Sea $A \subset N_+$, entonces

$$\begin{aligned} P(X \in A) &= P(\{X \in A\} \cap \{X = Y\}) + P(\{X \in A\} \cap \{X \neq Y\}) \\ &\leq P(Y \in A) + P(X \neq Y). \end{aligned}$$

En consecuencia

$$P(X \in A) - P(Y \in A) \leq P(X \neq Y).$$

De manera análoga se obtiene que $P(Y \in A) - P(X \in A) \leq P(X \neq Y)$, y por lo tanto

$$d(X,Y) = \sup_{A \subset N_+} |P(X \in A) - P(Y \in A)| \leq P(X \neq Y)$$

b)Sea $m \in N_+$, entonces

$$\begin{aligned} P(X \leq m) &= P(\{X \leq m\} \cap \{Y \leq X\}) + P(\{X \leq m\} \cap \{Y > X\}) \\ &\leq P(Y \leq m) + P(Y > X). \end{aligned}$$

En consecuencia

$$P(X \leq m) - P(Y \leq m) \leq P(Y > X).$$

Análogamente se tiene que $P(Y \leq m) - P(X \leq m) \leq P(X > Y)$, y por lo tanto

$$d_0(X,Y) = \sup_{m \in N_+} |P(X \leq m) - P(Y \leq m)| \leq \max\{P(X > Y), P(Y > X)\}.$$

□

El siguiente corolario fué propuesto por Freedman [15].

**COROLARIO 3.2.1**
*Sean $Y(\mu_1)$, $Y(\mu_2)$ dos variables aleatorias independientes Poisson con parámetros $0 < \mu_1 < 1$, $0 < \mu_2 < 1$ respectivamente. Entonces*

$$d(Y(\mu_1), Y(\mu_2)) \leq |\mu_1 - .\mu_2|.$$

*Demostración*
Supongase que $0 < \mu_1 < \mu_2$ y sean

$$X = Y(\mu_1) \quad y \quad Z = Y(\mu_2 - \mu_1)$$

variables aleatorias con parámetros $\mu_1$ y $(\mu_2 - \mu_1)$ respectivamente, entonces por el teorema anterior se tiene que

$$\begin{aligned} d(Y(\mu_1), Y(\mu_2)) &= d(X, X+Z) \\ &\leq P(X \neq X+Z) \\ &= P(Z \neq 0) \\ &= 1 - \exp(-(\mu_2 - \mu_1)) \\ &\leq \mu_2 - \mu_1. \end{aligned}$$

Por lo tanto

$$d(Y(\mu_1), Y(\mu_2)) \leq \mu_2 - \mu_1.$$

Observe que se usó el resultado siguiente: la suma de variables aleatorias independientes Poisson resulta ser una variable aleatoria Poisson con parámetro igual a la suma de los parámetros de cada una de las variables. Además, se

usó la conocida desigualdad $(1+t) \leq \exp(t)$, si $t \in \Re$.
De manera análoga se obtiene que

$$d(Y(\mu_1), Y(\mu_2)) \leq \mu_1 - \mu_2 \quad si\ 0 < \mu_2 < \mu_1.$$

Por lo tanto $d(Y(\mu_1), Y(\mu_2)) \leq |\mu_1 - \mu_2|$. □

La siguiente propiedad nos dice que las distribuciones de X y de Y pueden ser cambiadas por otras " más cercanas " si una variable aleatoria Z, independiente de X y Y, es sumada a ambas variables.

**TEOREMA 3.2.2**
*Sean $X$, $Y$ y $Z$ variables aleatorias en $\mathcal{Z}$ tales que $Z$ es independiente de $X$ y de $Y$, entonces*

$$d^*(X+Z, Y+Z) \leq d^*(X,Y).$$

*donde $d^* = d$ ó $d_0$.*

*Demostración*
Para $d^* = d$:
Sea $A \subset N_+$, como $P(X \in A) \leq P(Y \in A) + d(X,Y)$, entonces

$$\begin{aligned}
P(X+Z \in A) &= \sum_{n=0}^{\infty} P(X \in A \setminus \{n\}, Z = n) \\
&= \sum_{n=0}^{\infty} P(X \in A \setminus \{n\}) P(Z = n) \quad (\ \textit{independencia}\ ) \\
&\leq \sum_{n=0}^{\infty} P(Z = n)[P(Y \in A \setminus \{n\}) + d(X,Y)] \\
&= P(Y+Z \in A) + d(X,Y).
\end{aligned}$$

En consecuencia

$$P(X+Z \in A) - P(Y+Z \in A) \leq d(X,Y).$$

Analogamente se obtiene que

$$P(Y+Z \in A) - P(X+Z \in A) \leq d(X,Y).$$

Por lo tanto

$$d(X+Z,Y+Z) = \sup_{A\subset N_+} \{|P(X+Z\in A) - P(Y+Z\in A)|\} \leq d(X,Y).$$

Un argumento similar da el resultado para el caso cuando $d^* = d_0$. □

El siguiente corolario del teorema anterior nos da una propiedad de subaditividad para $d$ y $d_0$. Esta propiedad tiene diferentes pruebas, y se pueden encontar en [10] y [11].

**COROLARIO 3.2.2**
*Sean $X_1, X_2, \cdots, X_n$ y $Y_1, Y_2, \cdots, Y_n$ variables aleatorias independientes que pertenecen a $\mathcal{Z}$, entonces*

$$d^*(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) \leq \sum_{k=1}^{n} d^*(X_k, Y_k).$$

*donde $d^* = d$ ó $d_0$.*

*Demostración*
Para $n = 2$,
$d^*(X_1 + X_2, Y_1 + Y_2)$

$$\begin{aligned} &\leq d^*(X_1+X_2, Y_1+X_2) + d^*(Y_1+X_2, Y_1+Y_2) \quad (\textit{ D. Triangular }) \\ &\leq d^*(X_1,Y_1) + d^*(X_2,Y_2) \quad (\textit{ Teorema anterior }). \end{aligned}$$

Por lo tanto

$$d^*(X_1+X_2, Y_1+Y_2) \leq d^*(X_1,Y_1) + d^*(X_2,Y_2).$$

Y el resultado se sigue por inducción sobre n. □

Finalmente presentaremos una forma más débil del corolario anterior que nos permite generalizar al caso en que las $X_n$'s y $Y_n$'s sean dependientes.

**TEOREMA 3.2.3**
*Sean $X_1, X_2, \cdots, X_n$ y $Y_1, Y_2, \cdots, Y_n$ variables aleatorias que pertenecen a $\mathcal{Z}$, entonces*
*a)* $d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) \leq \sum_{k=1}^{n} P(X_k \neq Y_k)$.
*b)* $d_0(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) \leq \sum_{k=1}^{n} \max\{P(X_k < Y_k), P(X_k > Y_k)\}$.

*Demostración*
a) Por el *teorema 3.2.1 ( inciso (a) )* tenemos que

$$d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) \leq P(\sum_{k=1}^{n} X_k \neq \sum_{k=1}^{n} Y_k),$$

pero

$$\begin{aligned} P(\sum_{k=1}^{n} X_k \neq \sum_{k=1}^{n} Y_k) &\leq P(\bigcup_{k=1}^{n} \{X_k \neq Y_k\}) \\ &\leq \sum_{k=1}^{n} P(X_k \neq Y_k). \end{aligned}$$

Por lo tanto

$$d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) \leq \sum_{k=1}^{n} P(X_k \neq Y_k).$$

b) Como

$$\begin{aligned} P(\sum_{k=1}^{n} X_k < \sum_{k=1}^{n} Y_k) &\leq P(\bigcup_{k=1}^{n} \{X_k < Y_k\}) \\ &\leq \sum_{k=1}^{n} P(X_k < Y_k) \quad y \\ P(\sum_{k=1}^{n} X_k > \sum_{k=1}^{n} Y_k) &\leq P(\bigcup_{k=1}^{n} \{X_k > Y_k\}) \\ &\leq \sum_{k=1}^{n} P(X_k > Y_k), \end{aligned}$$

entonces

$$\begin{aligned} &\max\{P(\textstyle\sum_{k=1}^{n} X_k < \sum_{k=1}^{n} Y_k), P(\sum_{k=1}^{n} X_k > \sum_{k=1}^{n} Y_k)\} \\ &\qquad \leq \max\{\sum_{k=1}^{n} P(X_k < Y_k), \sum_{k=1}^{n} P(X_k > Y_k)\} \\ &\qquad \leq \sum_{k=1}^{n} \max\{P(X_k < Y_k), P(X_k > Y_k)\}, \end{aligned}$$

además, por el *teorema 3.2.1 ( inciso (b) )* se sabe que

$$d_0(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) \leq \max\{P(\sum_{k=1}^{n} X_k < \sum_{k=1}^{n} Y_k), P(\sum_{k=1}^{n} X_k > \sum_{k=1}^{n} Y_k)\}.$$

Por lo tanto

$$\begin{aligned} d_0(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k) &\leq \max\{P(\sum_{k=1}^{n} X_k < \sum_{k=1}^{n} Y_k), P(\sum_{k=1}^{n} X_k > \sum_{k=1}^{n} Y_k)\} \\ &\leq \sum_{k=1}^{n} \max\{P(X_k < Y_k), P(X_k > Y_k)\}. \end{aligned}$$

□

## 3.3 Cotas de la métrica $d$ para la aproximación Poisson a una suma de variables aleatorias independientes Bernoulli

En esta sección estudiaremos la aproximación Poisson a una suma de variables aleatorias independientes Bernoulli. Dicha aproximación la haremos utilizando la métrica $d$ y la siguiente observación.

**Observación 3.3.1**
Sea $f(\lambda) = \frac{1}{2}\lambda^2 + \lambda e^{-\lambda} + e^{-\lambda} - 1$ para todo $\lambda \in \Re^+ \cup \{0\}$. Afirmamos que

$$f(\lambda) \geq 0 \quad \forall \lambda \geq 0.$$

Y en consecuencia $1 - e^{-\lambda} - \lambda e^{-\lambda} \leq \frac{1}{2}\lambda^2$.
*Demostración*
$f(0) = 0$ y

$$\begin{aligned} f'(\lambda) &= \lambda + e^{-\lambda} - \lambda e^{-\lambda} - e^{-\lambda} \quad \forall \lambda \geq 0 \\ &= \lambda - \lambda e^{-\lambda} \quad \forall \lambda \geq 0 \\ &= \lambda(1 - e^{-\lambda}) \quad \forall \lambda \geq 0 \\ &\geq 0. \end{aligned}$$

Entonces $f(\lambda)$ es creciente y $f(0) = 0$. Por lo tanto $f(\lambda) \geq 0$. $\square$

Primeramente consideraremos el caso de una sola variable aleatoria X Bernoulli. Si queremos que se cumpla que $E[X] = E[Y]$, entonces tomamos $Y(p)$ variable aleatoria Poisson con parámetro $p$. Por otro lado, si queremos que $P(X = 0) = P(Y = 0)$, entonces tomamos $Y(\lambda)$ variable aleatoria Poisson con parámetro $\lambda = -\log(1-p)$. Para estos casos tenemos el siguiente resultado.

**TEOREMA 3.3.1**

*Sea X una variable aleatoria Bernoulli con parámetro p,* $0 < p < 1$. *Sea* $\lambda = -\log(1-p)$, *entonces*

*i)* $d(X, Y(p)) \leq p(1 - e^{-p}) \leq p^2$ *y*

*ii)* $d(X, Y(\lambda)) \leq 1 - e^{-\lambda} - \lambda e^{-\lambda} \leq \frac{1}{2}\lambda^2$

*Demostración*

Utilizaremos la primera identidad dada en el *teorema 3.1.1* .

i)

$$\begin{aligned} |P(X=0) - P(Y=0)| &= |1 - p - e^{-p}| \\ &= e^{-p} + p - 1 \quad (\text{ ya que } e^{-p} \geq 1 - p \text{ }) \\ |P(X=1) - P(Y=1)| &= |p - pe^{-p}| \\ &= p(1 - e^{-p}), \end{aligned}$$

además

$$|P(X = n) - P(Y = n)| = P(Y = n) \quad \forall n \geq 2.$$

Por lo tanto

$$\begin{aligned} d(X,Y) &= \frac{1}{2}[(e^{-p} + p - 1) + (p - pe^{-p}) + P(Y \geq 2)] \\ &= \frac{1}{2}[2p + e^{-p} - pe^{-p} - 1 + (1 - e^{-p} - pe^{-p})] \\ &= \frac{1}{2}[2p - 2pe^{-p}] \\ &= p - pe^{-p} \\ &= p(1 - e^{-p}) \\ &\leq p^2 \quad (\text{ porque } 1 - e^{-p} \leq p \text{ }) \end{aligned}$$

ii) Como $e^{-\lambda} = 1 - p$ y

$$1 = P(Y \geq 0) \geq P(Y = 0) + P(Y = 1) = e^{-\lambda} + \lambda e^{-\lambda} \tag{3.1}$$

entonces

$$\begin{aligned} |P(X = 0) - P(Y = 0)| &= |(1 - p) - e^{-\lambda}| \\ &= |(1 - p) - (1 - p)| \\ &= 0 \\ |P(X = 1) - P(Y = 1)| &= |p - \lambda e^{-\lambda}| \\ &= |(1 - e^{-\lambda}) - \lambda e^{-\lambda}| \\ &= 1 - e^{-\lambda} - \lambda e^{-\lambda} \quad (\ por\ 3.1\ ), \end{aligned}$$

además

$$|P(X = n) - P(Y = n)| = P(Y = n) \quad \forall n \geq 2.$$

Por lo tanto

$$\begin{aligned} d(X, Y(\lambda)) &= \frac{1}{2}[1 - e^{-\lambda} - \lambda e^{-\lambda} + P(Y \geq 2)] \\ &= \frac{1}{2}[1 - e^{-\lambda} - \lambda e^{-\lambda} + (1 - e^{-\lambda} - \lambda e^{-\lambda}] \\ &= 1 - e^{-\lambda} - \lambda e^{-\lambda} \\ &\leq \frac{1}{2}\lambda^2 \quad (\ \textit{Observación}\ 3.3.1\ ) \end{aligned}$$

□

Haciendo uso de la subaditividad para $d$ ( *corolario 3.2.2* ) y en conjunción del teorema anterior se tiene el siguiente:

**COROLARIO 3.3.1**

*Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes Bernoulli con probabilidad de éxito $p_1, p_2, \cdots p_n$ respectivamente y $0 < p_k < 1$.*
*Si $\lambda_k = -\log(l - p_k)$ para $k = 1, \cdots n$. Entonces*

$$\begin{aligned} d\left(\sum_{k=1}^{n} X_k, Y\left(\sum_{k=1}^{n} p_k\right)\right) &\leq \sum_{k=1}^{n} p_k^2 \\ d\left(\sum_{k=1}^{n} X_k, Y\left(\sum_{k=1}^{n} \lambda_k\right)\right) &\leq \frac{1}{2}\sum_{k=1}^{n} \lambda_k^2 \end{aligned}$$

*Demostración*

$$
\begin{aligned}
d(\sum_{k=1}^{n} X_k, Y(\sum_{k=1}^{n} p_k)) &= d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y(p_k)) \\
&\leq \sum_{k=1}^{n} d(X_k, Y(p_k)) \\
&\leq \sum_{k=1}^{n} p_k^2 \\
d(\sum_{k=1}^{n} X_k, Y(\sum_{k=1}^{n} \lambda_k)) &= d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y(\lambda_k)) \\
&\leq \sum_{k=1}^{n} d(X_k, Y(\lambda_k)) \\
&\leq \frac{1}{2}\sum_{k=1}^{n} \lambda_k^2
\end{aligned}
$$

□

Una prueba alternativa de la primera desigualdad de este teorema, usando variables aleatorias indicadoras, es dada en [26]. Además, esta desigualdad es un caso particular de un resultado general de Le Cam [22].

Como una aplicación directa de estos resultados tenemos la aproximación Poisson clásica.

**Ejemplo 3.3.1**

Sea $S_n$ una variable aleatoria con distribución Binomial con parámetros $(n, p_n)$ y supongase que $\lim_{n\to\infty} p_n = 0$ y $\lim_{n\to\infty} np_n = \lambda$, $0 < \lambda < \infty$, entonces

$$\lim_{n\to\infty} d(S_n, Y(\lambda)) = 0.$$

Observese que $Y(\lambda)$ representa una variable aleatoria Poisson.

*Demostración*

Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes Bernoulli todas con parámetro $p_n$, entonces por el corolario anterior y por el *corolario 3.2.1*

$$\begin{aligned} d(S_n, Y(\lambda)) &\leq d(S_n, Y(np_n)) + d(Y(np_n), Y(\lambda)) \quad (\ D.\ triangular\ ) \\ &= d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y(p_n)) + d(Y(np_n), Y(\lambda)) \\ &\leq np_n^2 + |np_n - \lambda|. \end{aligned}$$

Por lo tanto

$$\lim_{n\to\infty} d(S_n, Y(\lambda) \leq \lim_{n\to\infty} (np_n^2 + |np_n - \lambda|) = 0.$$

□

## 3.4 Cotas de la métrica $d_0$ para la aproximación Poisson a una suma de variables aleatorias independientes Bernoulli

De igual forma, primeramente consideraremos el caso de una sola variable aleatoria X Bernoulli.

**Observación 3.4.1**
i) Sea $f(p) = 2e^{-p} + pe^{-p} + p - 2 \quad \forall p \in \Re^+ \cup \{0\}$. Afirmamos que

$$f(p) \geq 0 \quad \forall p \geq 0.$$

Y en consecuencia $1 - e^{-p} - pe^{-p} \leq e^{-p} + p - 1$.
ii) Sea $g(p) = \frac{1}{2}p^2 - e^{-p} - p + 1 \quad \forall p \in \Re^+ \cup \{0\}$. Afirmamos que

$$g(p) \geq 0 \quad \forall p \geq 0.$$

Y en consecuencia $e^{-p} + p - 1 \leq \frac{1}{2}p^2$.
*Demostración*
i) $f(0) = 0$ y

$$\begin{aligned} f'(p) &= -2e^{-p} + e^{-p} - pe^{-p} + 1 \quad \forall p \geq 0 \\ &= -e^{-p} - pe^{-p} + 1 \quad \forall p \geq 0 \\ &\geq 0, \quad (\ por\ 3.1\ ) \end{aligned}$$

Entonces $f(p)$ es creciente y $f(0) = 0$. Por lo tanto

$$f(p) \geq 0 \quad \forall p \geq 0.$$

ii) $g(0) = 0$ y

$$\begin{aligned} g'(p) &= p + e^{-p} - 1 \quad \forall p \geq 0 \\ &\geq 0, \quad (\textit{ porque } e^{-p} \geq 1 - p \;) \end{aligned}$$

Entonces $g(p)$ es creciente y $g(0) = 0$. Por lo tanto $g(p) \geq 0 \; \forall p \geq 0$. □

El siguiente resultado fué demostrado por Daley [11].

**TEOREMA 3.4.1**

*Sea $X$ una variable aleatoria Bernoulli con parámetro $p$, $0 < p < 1$, entonces*

$$d_0(X, Y(p)) \leq \frac{1}{2}p^2.$$

*Demostración*

$$\begin{aligned} |P(X \leq 0) - P(Y \leq 0)| &= |1 - p - e^{-p}| \\ &= e^{-p} + p - 1 \quad (\textit{ porque } e^{-p} \geq 1 - p \;) \\ |P(X \leq 1) - P(Y \leq 1)| &= |1 - e^{-p} - pe^{-p}| \\ &= 1 - e^{-p} - pe^{-p} \quad (\textit{ por } 3.1 \;), \end{aligned}$$

y además como $P(Y \leq 1) \leq P(Y \leq n)$ para toda $n \geq 2$, entonces

$$\begin{aligned} |P(X \leq n) - P(Y \leq m)| &= |1 - P(Y \leq n)| \quad \forall n \geq 2 \\ &= 1 - P(Y \leq n) \quad \forall n \geq 2 \\ &\leq 1 - P(Y \leq 1) \quad \forall n \geq 2 \\ &= 1 - e^{-p} - pe^{-p} \quad \forall n \geq 2. \end{aligned}$$

Y usando la ***observación 3.4.1*** se tiene que

$$\begin{aligned} d_0(X, Y(p)) &= \sup_{n \geq 0} |P(X \leq n) - P(Y \leq n)| \\ &= \max\{e^{-p} + p - 1, 1 - e^{-p} - pe^{-p}\} \\ &= e^{-p} + p - 1 \\ &\leq \frac{1}{2}p^2 \end{aligned}$$

Por lo tanto $d_0(X, Y(p)) \leq \frac{1}{2}p^2$. □

**Observación 3.4.2**
Si $X$ es Bernoulli con parámetro $p$ y $\lambda = -\log(1-p)$, $0 < p < 1$, entonces
i) $P(X \geq t) \leq P(Y(\lambda) \geq t) \quad \forall t$.
ii) $P(X \neq Y(\lambda)) \leq \frac{1}{2}\lambda^2$.
*Demostración*
i) Dada la variable aleatoria $Y(\lambda)$ definimos a

$$X' = I_{\{Y(\lambda) \geq 1\}},$$

la cual tiene la misma distribucón que $X$ porque

$$\begin{aligned} P(X'=1) &= P(Y(\lambda) \geq 1) \\ &= 1 - P(Y(\lambda) = 0) \\ &= 1 - e^{-\lambda} \\ &= p \\ &= P(X = 1), \end{aligned}$$

además, se cumple que $X' \leq Y(\lambda)$ y por lo tanto

$$\begin{aligned} P(X \geq t) &= P(X' \geq t) \\ &\leq P(Y(\lambda) \geq t). \end{aligned}$$

ii) Tomamos la misma $X'$ que antes y tenemos que

$$\begin{aligned} P(X \neq Y(\lambda)) &= P(X' \neq Y(\lambda)) \\ &= P(X'=1, X' \neq Y(\lambda)) + P(X'=0, X' \neq Y(\lambda)) \\ &= P(Y(\lambda) \geq 1, Y(\lambda) \neq 1) + P(Y(\lambda) < 1, Y(\lambda) \neq 0) \\ &= P(Y(\lambda) \geq 2) \\ &= 1 - P(Y(\lambda) \leq 1) \\ &= 1 - e^{-\lambda} - \lambda e^{-\lambda} \\ &\leq \frac{1}{2}\lambda^2 \quad (\textit{ observación } 3.3.1\ ) \end{aligned}$$

Por lo tanto $P(X \neq Y(\lambda)) \leq \frac{1}{2}\lambda^2$. □

Como corolario del teorema anterior tenemos el siguiente resultado que fué probado por Daley [11].

**COROLARIO 3.4.1**
*Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes Bernoulli con probabilidad de éxito $p_1, p_2, \cdots p_n$ respectivamente y $0 < p_k < 1$.*
*Si $\lambda_k = -\log(l-p)$ para todo $k = 1, \cdots n$. Entonces*

$$d_0\left(\sum_{k=1}^{n} X_k, Y\left(\sum_{k=1}^{n} p_k\right)\right) \leq \frac{1}{2}\sum_{k=1}^{n} p_k^2$$

*y además, para cada $m \in N_+$*

$$\begin{aligned} P\left(Y\left(\sum_{k=1}^{n} \lambda_k\right) \leq m\right) &\leq P\left(\sum_{k=1}^{n} X_k \leq m\right) \\ &\leq P\left(Y\left(\sum_{k=1}^{n} \lambda_k\right) \leq m\right) + \frac{1}{2}\sum_{k=1}^{n} \lambda_k^2 \end{aligned}$$

*Demostración*
Usando el *corolario 3.2.2* y el *teorema 3.4.1*

$$\begin{aligned} d_0\left(\sum_{k=1}^{n} X_k, Y\left(\sum_{k=1}^{n} p_k\right)\right) &\leq \sum_{k=1}^{n} d_0(X_k, Y(p_k)) \\ &\leq \frac{1}{2}\sum_{k=1}^{n} p_k^2. \end{aligned}$$

Para la otra parte, sean $X'_k = I_{\{Y_k \geq 1\}}$, por la *observación 3.4.2* se tiene que

$$\begin{aligned} \sum_{k=1}^{n} X'_k &\leq \sum_{k=1}^{n} Y_k \\ &= Y\left(\sum_{k=1}^{n} \lambda_k\right), \end{aligned}$$

por lo tanto

$$\begin{aligned} P\left(\sum_{k=1}^{n} Y_k \leq m\right) &\leq P\left(\sum_{k=1}^{n} X'_k \leq m\right) \\ &= P\left(\sum_{k=1}^{n} X_k \leq m\right) \quad y \end{aligned}$$

$P(\sum_{k=1}^{n} X_k \leq m)$

$$\begin{aligned}
&= P(\sum_{k=1}^{n} X_k \leq m, \sum_{k=1}^{n} X_k = \sum_{k=1}^{n} Y_k) + P(\sum_{k=1}^{n} X_k \leq m, \sum_{k=1}^{n} X_k \neq \sum_{k=1}^{n} Y_k) \\
&\leq P(\sum_{k=1}^{n} Y_k \leq m) + P(\sum_{k=1}^{n} X_k \neq \sum_{k=1}^{n} Y_k) \\
&\leq P(\sum_{k=1}^{n} Y_k \leq m) + \sum_{k=1}^{n} P(X_k \neq Y_k) \\
&\leq P(\sum_{k=1}^{n} Y_k \leq m) + \frac{1}{2}\sum_{k=1}^{n} \lambda_k^2 \quad (\textit{ observación } 3.4.2 ).
\end{aligned}$$

□

## 3.5 Extensión para el caso de variables aleatorias dependientes

En esta sección acotaremos el error de la aproximación Poisson con la metrica $d$ cuando las variables aleatorias $X_n$'s son dependientes Bernoulli. Además, demostraremos un resultado más general para el caso de una suma de variables aleatorias que toman valores en $N_+$. Nuestro primer resultado fué provado por Serfling [26] y ha sido sujeto a diferentes pruebas (Ver [16]).

**TEOREMA 3.5.1**
*Sean $X_1, X_2, \cdots, X_n$ variables aleatorias Bernoulli. Definimos*

$$\begin{aligned}
\theta_1 &= P(X_1 = 1) \quad y \\
\theta_i &= P(X_i = 1 \ / \ \mathcal{F}_{i-1}) \quad \forall\, 2 \leq i \leq n.
\end{aligned}$$

*donde $\mathcal{F}_{i-1} = \sigma\{X_1, \cdots, X_{i-1}\}$. Sean $X_1^*, X_2^*, \cdots, X_n^*$ variables aleatorias independientes Bernoulli con parámetros $p_1, p_2, \cdots, p_n$ respectivamente. Entonces*

$$d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X_k^*) \leq \sum_{k=1}^{n} E[|\theta_k - p_k|]$$

*Demostración*

Sean $U_1, U_2, \cdots, U_n$ una sucesión de variables aleatorias independientes uniformemente distribuidas sobre el intervalo $[0,1]$. Definimos

$$\begin{aligned} X_i^1 &= I_{\{U_i \le \theta_i\}} \quad \forall \le i \le n. \\ X_i^2 &= I_{\{U_i \le p_i\}} \quad \forall 1 \le i \le n. \end{aligned}$$

Estas $X_n^1$'s y $X_n^2$'s estan definidas en un espacio común de probabilidad, las cuales cumplen las siguientes propiedas:

i) $X_n^1$'s son variables aleatorias ( posiblemente dependientes ) Bernoulli con la misma distribución que las $X_n$'s.

ii) $X_n^2$'s son variables aleatorias independientes Bernoulli con parámetros $0 < p_n < 1$, respectivamente. Entonces

$$\begin{aligned} P(X_k \neq X_k^*) &= P(X_k^1 \neq X_k^2) \\ &= E[E[I_{\{X_k^1 \neq X_k^2\}} \ / \ \theta_k]] \\ &= E[E[I_{\{X_k^1=1\}\cap\{X_k^2=0\}} + I_{\{X_k^1=0\}\cap\{X_k^2=1\}} \ / \ \theta_k]] \\ &= E[E[I_{\{X_k^1=1\}\cap\{X_k^2=0\}} \ / \ \theta_k]] + E[E[I_{\{X_k^1=0\}\cap\{X_k^2=1\}} \ / \ \theta_k]]. \end{aligned}$$

Pero

$$\begin{aligned} E[I_{\{X_k^1=1\}\cap\{X_k^2=0\}} \ / \ \theta_k] &= E[I_{\{U_k \le \theta_k\}\cap\{U_k > p_k\}} \ / \ \theta_k] \\ &= E[I_{\{p_k < U_k \le \theta_k\}} I_{\{p_k < \theta_k\}} \ / \ \theta_k] \\ &= \sum_{y \in J_k} E[I_{\{p_k < U_k \le \theta_k\}} I_{\{p_k < \theta_k\}} \ / \ \theta_k = y] I_{\{\theta_k = y\}} \\ &= \sum_{y \in J_k} E[I_{\{p_k < U_k \le y\}} I_{\{p_k < y\}}] I_{\{\theta_k = y\}} \\ &= \sum_{y \in J_k} P(p_k < U_k \le y) I_{\{p_k < y\}} I_{\{\theta_k = y\}} \\ &= \sum_{y \in J_k} (y - p_k) I_{\{p_k < y\}} I_{\{\theta_k = y\}} \\ &= (\theta_k - p_k) I_{\{p_k < \theta_k\}}, \end{aligned}$$

donde $J_k = \{$todos los posibles valores que toma $\theta_k\}$. En consecuencia

$$E[I_{\{X_k^1=1\}\cap\{X_k^2=0\}} \ / \ \theta_k] = (\theta_k - p_k) I_{\{p_k < \theta_k\}},$$

de manera análoga se obtiene que

$$E[I_{\{X_k^1=0\}\cap\{X_k^2=1\}} \ / \ \theta_k] = (p_k - \theta_k) I_{\{\theta_k < p_k\}}.$$

Entonces

$$\begin{aligned} P(X_k \neq X_k^*) &= E[E[I_{\{X_k^1=1\}\cap\{X_k^2=0\}} \ / \ \theta_k]] + E[E[I_{\{X_k^1=0\}\cap\{X_k^2=1\}} \ / \ \theta_k]] \\ &= E[(\theta_k - p_k) I_{\{p_k < \theta_k\}}] + E[(p_k - \theta_k) I_{\{\theta_k < p_k\}}] \\ &= E[|\theta_k - p_k|]. \end{aligned}$$

Por lo tanto

$$\begin{aligned} d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X_k^*) &\leq \sum_{k=1}^{n} P(X_k \neq X_k^*) \quad (\ \textit{teorema}\ 3.2.3\ ) \\ &= \sum_{k=1}^{n} E[\theta_k - p_k|]. \end{aligned}$$

□

**Observación 3.5.1**
i) Es claro que la cota para $d$ es también una cota para $d_0$.
ii) $E[|\theta_k - p_k|]$ denota la cantidad $E[|\theta_k(X_1, X_2, \cdots, X_{k-1}) - p_k|]$. Para $k = 1$ esto es simplemente $E[|\theta_1 - p_1|]$
iii) Un hecho importante del teorema es que los valores de $p_k$ pueden ser seleccionados arbitrariamente. Por ejemplo, una elección natural y conveniente es $p_k = E[\theta_k]$.

Antes de enunciar una forma más general del teorema anterior, necesitamos establecer la relación entre una sucesión $X_1, X_2, \cdots, X_n \in \mathcal{Z}$ (no necesariamente Bernoulli) y las varaibles aleatorias Bernoulli definidas como

$$X_k = I_{\{X_k=1\}} \quad \forall\, 1 \leq k \leq n.$$

**TEOREMA 3.5.2**
*Sean $X_1, X_2, \cdots, X_n$ variables aleatorias que toman valores en $N_+$. Sean*

$$X_k' = I_{\{X_k=1\}} \quad \forall\, 1 \leq k \leq n.$$

*Entonces*

$$d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X'_k) \leq \sum_{k=1}^{n} P(X_k \geq 2) \quad y$$

$$\begin{aligned} P(\sum_{k=1}^{n} X'_k \leq m) - \sum_{k=1}^{n} P(X_k \geq 2) &\leq P(\sum_{k=1}^{n} X_k \leq m) \\ &\leq P(\sum_{k=1}^{n} X'_k \leq m). \end{aligned}$$

*Demostración*
Usando el *teorema 3.2.3* y que

$$\begin{aligned} P(X'_k \neq X_k) &= P(X'_k \neq X_k, X'_k = 1) + P(X'_k \neq X_k, X'_k = 0) \\ &= P(X_k \neq 1, X_k = 1) + P(X_k \neq 0, X_k \neq 1) \\ &= P(X_k \geq 2) \quad \forall 1 \leq k \leq n \end{aligned}$$

podemos concluir que

$$\begin{aligned} d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X'_k) &\leq \sum_{k=1}^{n} P(X'_k \neq X_k) \\ &= \sum_{k=1}^{n} P(X_k \geq 2) \end{aligned}$$

Por otro lado, sea $m \in N_+$, entonces

$$P(\sum_{k=1}^{n} X'_k \leq m) - P(\sum_{k=1}^{n} X_k \leq m) \leq d_0(\sum_{k=1}^{n} X'_k, \sum_{k=1}^{n} X_k)$$

pero

$$\begin{aligned} d_0(\sum_{k=1}^{n} X'_k, \sum_{k=1}^{n} X_k) &\leq d(\sum_{k=1}^{n} X'_k, \sum_{k=1}^{n} X_k) \\ &\leq \sum_{k=1}^{n} P(X_k \geq 2), \end{aligned}$$

y en consecuencia

$$P(\sum_{k=1}^{n} X_k' \leq m) - \sum_{k=1}^{n} P(X_k \geq 2) \leq P(\sum_{k=1}^{n} X_k \leq m).$$

Por último, como $X_k' \leq X_k$, entonces

$$P(\sum_{k=1}^{n} X_k \leq m) \leq P(\sum_{k=1}^{n} X_k' \leq m).$$

□

Para terminar este capítulo, vamos a demostrar un teorema para el caso de una suma de variables aleatorias que toman valores en $N_+$.
Sean $X_1, X_2, \cdots, X_n \in \mathcal{Z}$ y

$$\theta_1 = P(X_1 = 1) \quad y \tag{3.2}$$

$$\theta_i = P(X_i = 1 \ / \ \mathcal{F}_{i-1}) \quad \forall\, 2 \leq i \leq n, \tag{3.3}$$

donde $\mathcal{F}_{i-1} = \sigma\{X_1, X_2, \cdots, X_{i-1}\}$.
Además, sean $X_1^*, X_2^*, \cdots, X_n^*$ variables aleatorias independientes Bernoulli con parámetros $p_1, p_2, \cdots, p_n$ respectivamente, $0 < p_i < 1$ y

$$\lambda_i = -\log(1 - p_i),\ \gamma_i = E[|\theta_i - p_i|]\ y\ \delta_i = P(X_i \geq 2) \quad 1 \leq i \leq n.$$

Definimos

$$\alpha = \sum_{k=1}^{n} p_k^2 \quad \beta = \sum_{k=1}^{n} \lambda_k \quad \gamma = \sum_{k=1}^{n} \gamma_k \quad \delta = \sum_{k=1}^{n} \delta_k$$

**TEOREMA 3.5.3**
*Sean $X_1, X_2, \cdots, X_n$, $p_1, p_2, \cdots, p_n$, $\lambda_1, \lambda_2, \cdots, \lambda_n$ y $\alpha, \beta, \gamma, \delta$ definidos anteriormente. Entonces*

$$d(\sum_{k=1}^{n} X_k, Y(\sum_{k=1}^{n} p_k)) \leq \alpha + \gamma + \delta \tag{3.4}$$

$$d(\sum_{k=1}^{n} X_k, Y(\sum_{k=1}^{n} \lambda_k)) \leq \frac{1}{2}\beta + \gamma + \delta \tag{3.5}$$

$$d_0(\sum_{k=1}^{n} X_k, Y(\sum_{k=1}^{n} p_k)) \leq \frac{1}{2}\alpha + \gamma + \delta \tag{3.6}$$

*y para cada $m \in N_+$*

$$P(Y(\sum_{k=1}^{n}\lambda_k)) \leq m) - \gamma - \delta \ \leq \ P(\sum_{k=1}^{n} X_k \leq m)$$
$$\leq \ P(Y(\sum_{k=1}^{n}\lambda_k) \leq m) + \frac{1}{2}\beta + \gamma$$

*Demostración*
Vamos a demostrar la primera desigualdad y de manera análoga se obtienen las otras desigualdades.
Sea $X'_k = I_{\{X_k=1\}}$ para todo $1 \leq k \leq n$, entonces por el teorema anterior tenemos que

$$\begin{aligned} d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X'_k) &\leq \sum_{k=1}^{n} P(X_k \geq 2) \\ &= \delta. \end{aligned}$$

Ahora bien, sean $X_1^*, X_2^*, \cdots, X_n^*$ variables aleatorias independientes Bernoulli con parámetros $p_1, p_2, \cdots, p_n$ respectivamente, entonces por el *teorema 3.5.2*

$$d(\sum_{k=1}^{n} X'_k, \sum_{k=1}^{n} X_k^*) \leq \sum_{k=1}^{n} E[|\theta'_k - p_k|]$$

donde $\theta'_1, \theta'_2, \cdots, \theta'_n$ estan definidas relativamente para las $X'_n$'s por las ecuaciones (3.2) *y* (3.3). Entonces, si $\mathcal{F}'_{i-1} = \sigma\{X'_1, X'_2, \cdots, X'_{i-1}\}$.

$$\theta'_1 = P(X'_1 = 1) = P(X_1 = 1) = \theta_1 \quad y$$

$$\begin{aligned} \theta'_i &= P(X'_i = 1 \ / \ \mathcal{F}'_{i-1}) \\ &= E[I_{\{X'_i=1\}} \ / \ \mathcal{F}'_{i-1}] \\ &= E[E[I_{\{X_i=1\}} \ / \ \mathcal{F}'_{i-1}] \ / \ \mathcal{F}_{i-1}] \\ &= E[E[I_{\{X_i=1\}} \ / \ \mathcal{F}_{i-1}] \ / \ \mathcal{F}'_{i-1}] \quad (\textit{porque } \mathcal{F}'_{i-1} \subseteq \mathcal{F}_{i-1}) \\ &= E[P(X_i = 1 \ / \ \mathcal{F}_{i-1}) \ / \ \mathcal{F}'_{i-1}] \\ &= E[\theta_i \ / \ \mathcal{F}'_{i-1}] \end{aligned}$$

y en consecuencia

$$\begin{aligned}
E[|\theta_i' - p_i|] &= E[|E[\theta_i \,/\, \mathcal{F}_{i-1}'] - p_i|] \\
&= E[|E[\theta_i - p_i \,/\, \mathcal{F}_{i-1}']|] \\
&\leq E[E[|\theta_i - p_i| \,/\, \mathcal{F}_{i-1}']] \quad (\ D.\ de\ Jensen\ ) \\
&= E[|\theta_i - p_i|]
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
d(\sum_{k=1}^{n} X_k', \sum_{k=1}^{n} X_k^*) &\leq \sum_{k=1}^{n} E[|\theta_k' - p_k|] \\
&\leq \sum_{k=1}^{n} E[|\theta_k - p_k|] \\
&= \gamma
\end{aligned}$$

Por otro lado, por el *corolario 3.3.1*

$$\begin{aligned}
d(\sum_{k=1}^{n} X_k^*, Y(\sum_{k=1}^{n} p_k)) &\leq \sum_{k=1}^{n} p_k^2 \\
&= \alpha
\end{aligned}$$

Por lo tanto, usando la desigualdad del triángulo, y las desigualdades anteriores obtenemos

$$\begin{aligned}
&d(\textstyle\sum_{k=1}^{n} X_k, Y(\sum_{k=1}^{n} p_k)) \\
&\leq d(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X_k') + d(\sum_{k=1}^{n} X_k', \sum_{k=1}^{n} X_k^*) + d(\sum_{k=1}^{n} X_k^*, Y(\sum_{k=1}^{n} p_k)) \\
&\leq \alpha + \gamma + \delta
\end{aligned}$$

□

**Ejemplo 3.5.1**

Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias dependientes Bernoulli con probabilidad de éxito $p = P(X_k = 1)$ para todo $k \geq 1$ y probabilidad de transición

$$\begin{aligned}
p^1 &= P(X_k = 1 \,/\, X_{k-1} = 1) \quad y \\
p^0 &= P(X_k = 1 \,/\, X_{k-1} = 0) \quad \forall\, k \geq 2,
\end{aligned}$$

es decir, la probabilidad de que $X_k$, para $k \geq 2$, tome el valor 1 ó 0 solamente depende del valor que tome $X_{k-1}$. Esta sucesión de variables aleatorias es conocida comunmente como un "proceso Markov-dependiente". Bajo estas hipótesis se tiene que

$$\theta_1 = P(X_1 = 1) = p$$

y para todo $k \geq 2$

$$\theta_k = P(X_k = 1 \,/\, X_1, \cdots, X_{k-1}) = \begin{cases} p^1 & \text{si } X_{k-1} = 1, \\ p^0 & \text{si } X_{k-1} = 0. \end{cases}$$

Y además,

$$\begin{aligned} p &= P(X_k = 1) \\ &= P(X_k = 1, X_{k-1} = 1) + P(X_k = 1, X_{k-1} = 0), \end{aligned}$$

pero

$$\begin{aligned} &P(X_k = 1, X_{k-1} = 1) + P(X_k = 1, X_{k-1} = 0) \\ &= P(X_k = 1 \,/\, X_{k-1} = 1)P(X_{k-1} = 1) + P(X_k = 1 \,/\, X_{k-1} = 0)P(X_{k-1} = 0) \\ &= p^1 p + p^0(1-p) \end{aligned}$$

en consecuencia

$$p = p^1 p + p^0(1-p), \tag{3.7}$$

y despejando a $p$ de esta ecuación se obtiene

$$p = \frac{p^0}{(1-p^1) + p^0}.$$

Por otro lado,

$$\begin{aligned} E[\theta_k] &= p^1 P(\theta_k = p^1) + p^0 P(\theta_k = p^0) \\ &= p^1 P(X_{k-1} = 1) + p^0 P(X_{k-1} = 0) \\ &= p^1 p + p^0(1-p) \\ &= p, \quad (por\ 3.7\ ) \end{aligned}$$

entonces $E[\theta_k] = p$.

Ahora, tomemos una sucesión $\{X_n^*\}_{n=1}^{\infty}$ de variables aleatorias independientes Bernoulli con parámetros $p_n = p$ para toda n. Entonces, usando (3.1)

$$\begin{aligned}
\gamma_k &= E[|\theta_k - p|] \\
&= |p^1 - p|P(\theta_k = p^1) + |p^0 - p|P(\theta_k = p^0) \\
&= |p^1 - p|P(X_{k-1} = 1) + |p^0 - p|P(X_{k-1} = 0) \\
&= |p^1 - p|p + |p^0 - p|(1-p) \\
&= |p^1 - [p^1 p + p^0(1-p)]|p + |p^0 - [p^1 p + p^0(1-p)]|(1-p) \\
&= |p^1(1-p) - p^0(1-p)|p + |p^0 - p^1 p - p^0 + p^0 p|(1-p) \\
&= p(1-p)|p^1 - p^0| + p(1-p)|p^1 - p^0| \\
&= 2p(1-p)|p^1 - p^0|,
\end{aligned}$$

en consecuencia $\gamma_k = 2p(1-p)|p^1 - p^0|$ para todo $k \geq 2$. Entonces, usando el *teorema 3.5.3 (ecuación (3.4))*

$$\begin{aligned}
d(\sum_{k=1}^{n} X_k, Y(np)) &\leq \alpha + \gamma + \delta \\
&= \sum_{k=1}^{n} p_k^2 + \sum_{k=1}^{n} \gamma_k + \sum_{k=1}^{n} \delta_k \\
&= \sum_{k=1}^{n} (p)^2 + \sum_{k=2}^{n} 2(1-p)|p^1 - p^0| + \sum_{k=1}^{n} P(X_k \geq 2) \\
&= n(p)^2 + 2(n-1)(1-p)|p^1 - p^0|
\end{aligned}$$

Por lo tanto

$$d(\sum_{k=1}^{n} X_k, Y(np)) \leq n(p)^2 + 2(n-1)(1-p)|p^1 - p^0| \tag{3.8}$$

Utilizando esto, tenemos que si

$$\lim_{n\to\infty} p^1 = 0, \quad \lim_{n\to\infty} p^0 = 0 \quad y \quad \lim_{n\to\infty} np = \lambda < \infty$$

entonces

$$\lim_{n\to\infty} d(S_n, Y(\lambda)) = 0.$$

En efecto:
Usando el *corolario 3.2.1* y la ecuación (3.8) se tiene que

$$\begin{aligned} d(S_n, Y(\lambda)) &\leq d(S_n, Y(np)) + d(Y(np), Y(\lambda)) \\ &\leq n(p)^2 + 2(n-1)(1-p)|p^1 - p^0| + |np - \lambda| \end{aligned}$$

Por lo tanto $\lim_{n\to\infty} d(S_n, Y(\lambda)) = 0$. □

Este ejemplo nos da un resultado que generaliza la clásica aproximación Poisson para variables aleatorias Bernoulli dependientes que satisfagan las condiciones mencionadas en este ejemplo.

# Capítulo 4

# La Métrica $d$ y la Convergencia débil

En este capítulo estudiaremos la relación que existe entre la métrica $d$, definida en el capítulo anterior, y la convergencia débil de una sucesión de medidas de probabilidad de Borel en el espacio de los números enteros no negativos.

Para establecer esa relación, utilizaremos la métrica de Prohorov la cual es una métrica para el espacio de medidas de Probabilidad sobre $(E, \mathbb{B}(E))$, donde $(E, L)$ es un espacio métrico con su respectiva métrica $L$.

En la primera sección de este capítulo, introduciremos la métrica de Prohorov y enunciaremos, sin demostración, el Teorema de Prohorov.

La segunda sección, tiene como finalidad principal mostrar la equivalencia entre la métrica $d$ y la métrica de Prohorov en el espacio de medidas de probabilidad sobre $(\mathbb{N}, \mathbb{B}(\mathbb{N}))$. Y de esta forma, tendremos una expresión más simple para la métrica de Prohorov en el espacio de medidas de probabilidad sobre $(\mathbb{N}, \mathbb{B}(\mathbb{N}))$.

## 4.1 El Teorema de Prohorov

Sean $(E, L)$ un espacio métrico ($L$ denota la métrica), $\mathbb{B}(E)$ la $\sigma$-álgebra de Borel de subconjuntos de $E$, y $\mathcal{P}(E)$ la familia de medidas de Borel de probabilidad sobre $E$.

En el espacio de $\mathcal{P}(E)$, definimos la métrica de Prohorov como:

$$\rho(P,Q) = \inf\{\epsilon > 0 \ / \ P(F) \leq Q(F^\epsilon) + \epsilon \ \forall F \in \mathcal{C}\},$$

donde $\mathcal{C}$ es la familia de todos los conjuntos cerrados de $E$ y

$$F^\epsilon = \{x \in E \ / \ \inf_{y \in F}\{L(x,y)\} < \epsilon\}.$$

Vamos a demostrar que $\rho$ es una métrica para el espacio $\mathcal{P}(E)$, para ver esto necesitamos el siguiente:

**TEOREMA 4.1.1**
*Sean $P,Q \in \mathcal{P}(E)$ y $\epsilon > 0$. si*

$$P(F) \leq Q(F^\epsilon) + \epsilon \quad \forall F \in \mathcal{C}, \tag{4.1}$$

*entonces*

$$Q(F) \leq P(F^\epsilon) + \epsilon \quad \forall F \in \mathcal{C},$$

*Demostración*
Sea $F \in \mathcal{C}$, tomemos $F_1 = E \setminus F^\epsilon$, observese que $F_1 \in \mathcal{C}$ y $F \subset E \setminus F_1^\epsilon$, aplicando (4.1) a $F_1$ se tiene que

$$\begin{aligned} P(F^\epsilon) &= 1 - P(F_1) \\ &\geq 1 - Q(F_1^\epsilon) - \epsilon \\ &\geq Q(F) - \epsilon \end{aligned}$$

Por lo tanto

$$Q(F) \leq P(F^\epsilon) + \epsilon \quad \forall F \in \mathcal{C}.$$

□

Se sigue del teorema anterior que $\rho(P,Q) = \rho(Q,P)$ para todo $P,Q \in \mathcal{P}(E)$. También, si $\rho(P,Q) = 0$, entonces $P(F) = Q(F)$ para todo $F \in \mathcal{C}$, y en consecuencia $P(F) = Q(F)$ para todo $F \in \mathbb{B}(E)$. Por lo tanto

$$\rho(P,Q) = 0 \Longleftrightarrow P = Q.$$

Finalmente, para ver la desigualdad del triángulo, sean $P, Q, R \in \mathcal{P}(E)$ tales que $\rho(P,Q) \leq \epsilon$, y $\rho(Q,R) \leq \delta$. Como $(\overline{F^\epsilon})^\delta \subseteq F^{\epsilon+\delta}$ y $F^\epsilon \subseteq \overline{F^\epsilon}$, entonces

$$\begin{aligned} P(F) &\leq Q(F^\epsilon) + \epsilon \\ &\leq Q(\overline{F^\epsilon}) + \epsilon \\ &\leq R((\overline{F^\epsilon})^\delta) + \epsilon + \delta \\ &\leq R(F^{\epsilon+\delta}) + \epsilon + \delta \quad \forall F \in \mathcal{C}. \end{aligned}$$

Por lo tanto $\rho(P,R) \leq \epsilon + \delta$.

Sea $\overline{C}(E)$ el espacio de todas las funciones de $(E, L)$ en $\Re$, continuas y acotadas con la norma

$$\|f\| = \sup_{x \in E}\{|f(x)|\}.$$

**DEFINICIÓN 4.1.1**

*Sea $(E, L)$ un espacio métrico, se dice que una sucesión de probabilidades $\{P_n\}_{n=1}^\infty \subset \mathcal{P}(E)$ converge débilmente a $P \in \mathcal{P}(E)$ si*

$$\lim_{n\to\infty} \int f\, dP_n = \int f\, dP \quad \forall f \in \overline{C}(E).$$

*En este caso escribiremos $P_n \overset{D}{\Longrightarrow} P$.*

Antes de enunciar el teorema de Prohorov, el cual establece que la convergencia débil es equivalente a la convergencia con la métrica de Prohorov, recordemos que dado $A \subset E$ y $P \in \mathcal{P}(E)$, se dice que $A$ es un conjunto de $P$-continuidad si

$$A \in \mathbb{B}(E) \quad y \quad P(\partial A) = 0$$

donde $\partial A$ es la frontera de $A$.

La demostracion de este teorema puede ser encontrada en [14].

**TEOREMA 4.1.2** *( Teorema de Prohorov )*

*Sean $(E, L)$ un espacio métrico, $\{P_n\}_{n=1}^\infty \subset \mathcal{P}(E)$ y $P \in \mathcal{P}(E)$. De las siguientes condiciones, de la (b) a la (f) son equivalentes y son consecuencia de la condición (a), además, si $(E, L)$ es un espacio separable entonces las seis son equivalentes.*

*a)* $\lim_{n\to\infty} \rho(P_n, P) = 0$.

*b)* $P_n \overset{D}{\Longrightarrow} P$.
*c)* $\lim_{n\to\infty} \int f\, dP_n = \int f dP \quad \forall f \in \overline{\mathcal{I}}(E)$.
*d)* $\limsup_{n\to\infty} P_n(F) \leq P(F) \quad \forall F \in \mathcal{C}$.
*e)* $\liminf_{n\to\infty} P_n(F) \geq P(F) \quad \forall F \in \mathcal{C}$.
*)*$\lim_{n\to\infty} P_n(A) = P(A)$ ***para todo conjunto A de P-continuidad.***

## 4.2 Convergencia débil cuando $E = N_+$

En esta sección estudiaremos la convergencia débil cuando el espacio $E$ es el conjunto de enteros no negativos, como en el capítulo anterior lo denotaremos por $N_+$, y la métrica $L$ es la del valor absoluto, es decir,

$$L(x,y) = |x-y| \quad \forall\, x,y \in N_+.$$

El objetivo principal de esta sección, y de este capítulo, es demostrar que dada una sucesión de $\{P_n\}_{n=1}^{\infty} \subset \mathcal{P}(N_+)$ y $P \in \mathcal{P}(N_+)$, se tiene que

$$\lim_{n\to\infty} \rho(P_n, P) = 0 \Longleftrightarrow \lim_{n\to\infty} d(P_n, P) = 0$$

donde $d(P_n, P) = \sup_{A \subset N_+}\{|P_n(A) - P(A)|\}$.

Para llevar acabo este objetivo, primeramente haremos algunas observaciones de gran utilidad.

**Observación 4.2.1**
Dado $F \subset N_+$ y $0 < r \leq 1$ se tiene que

$$F = F^r = \{x \in N_+ \ / \ \inf_{y\in F}\{L(x,y)\} < r\},$$

esto es cierto ya que

$$L(x,y) \geq 1 \quad \forall x \in S \setminus F \ y \ \ \forall y \in F.$$

**Observación 4.2.2**
Dado $F \subset N_+$ y $\mathcal{C}$ la familia de conjuntos cerrados de $N_+$, es fácil ver que $F \in \mathcal{C}$ ya que el conjunto de puntos de acumulación de $F$ es vacio y en consecuencia

$$\overline{F} = F \cup F' = F$$

donde $\overline{F}$ =la cerradura de $F$ y $F'$ =el conjunto de puntos de acumulación de $F$. Por lo tanto $N_+ = \mathcal{C}$.

**TEOREMA 4.2.1**
*Sea una sucesión de $\{P_n\}_{n=1}^{\infty} \subset \mathcal{P}(N_+)$ y sea $P \in \mathcal{P}(N_+)$. Entonces*

$$\lim_{n\to\infty} \rho(P_n, P) = 0 \iff \lim_{n\to\infty} d(P_n, P) = 0.$$

*Demostración*
$\Rightarrow$)
Supongamos que $\lim_{n\to\infty} \rho(P_n, P) = 0$ y sea $\epsilon > 0$.
Tomamos a $\epsilon' = \min\{1, \epsilon\}$, por hipótesis tenemos que existe $M \in \mathbb{N}$ ( que depende de $\epsilon'$ ) tal que

$$\rho(P_n, P) < \epsilon' \quad \forall n \geq M,$$

entonces existe $r \in \{r > 0 \ / \ P_n(F) - P(F^r) \leq r \quad \forall F \in \mathcal{C}\}$ tal que

$$\begin{aligned} P_n(F) - P(F^r) &\leq r \\ &< \epsilon' \quad \forall F \in \mathcal{C} \ y \ \forall n \geq M, \end{aligned}$$

y en consecuencia, por el *teorema 4.1.1*

$$\begin{aligned} P(F) - P_n(F^r) &\leq r \\ &< \epsilon' \quad \forall F \in \mathcal{C} \ y \ \forall n \geq M. \end{aligned}$$

Por lo tanto

$$\begin{aligned} |P_n(F) - P(F^r)| &\leq r \\ &< \epsilon' \quad \forall F \in \mathcal{C} \ y \ \forall n \geq M. \end{aligned}$$

Pero $F = F^r$ porque $0 < r \leq \epsilon' \leq 1$ (*obsevación 4.2.1* ), entonces

$$\begin{aligned} |P_n(F) - P(F)| &< \epsilon' \\ &\leq \epsilon \quad \forall F \in \mathcal{C} \ y \ \forall n \geq M. \end{aligned}$$

además, sabemos que $N_+ = \mathcal{C}$ (*observación 4.2.2* ), entonces

$$\begin{aligned} d(P_n P) &= \sup_{F \subset N_+} \{|P_n(F) - P(F)|\} \\ &< \epsilon \quad \forall n \geq M. \end{aligned}$$

Por lo tanto

$$\lim_{n\to\infty} d(P_n, P) = 0.$$

Observese que $\epsilon'$ depende de $\epsilon$ y en consecuencia $M$ depende de $\epsilon$.

$\Leftarrow$)

Supongamos que $\lim_{n\to\infty} d(P_n, P) = 0$ y sea $\epsilon > 0$. Tomemos a $\epsilon' = \min\{1, \epsilon\}$, por hipótesis existe $M \in \mathbb{N}$ tal que

$$d(P_n, P) < \epsilon' \quad \forall n \geq M,$$

entonces

$$|P_n(A) - P(A)| < \epsilon' \quad \forall n \geq M \text{ y } \forall A \subset N_+,$$

y en particular

$$|P_n(F) - P(F)| < \epsilon' \quad \forall n \geq M \text{ y } \forall F \in \mathcal{C}.$$

Pero $F = F^{\epsilon'} \quad \forall F \in \mathcal{C}$ porque $\epsilon' \leq 1$. Entonces

$$\begin{aligned} |P_n(F) - P(F^{\epsilon'})| &< \epsilon' \\ &\leq \epsilon \quad \forall n \geq M \text{ y } \forall F \in \mathcal{C}. \end{aligned}$$

En consecuencia

$$\rho(P_n, P) < \epsilon \quad \forall n \geq M,$$

y por lo tanto

$$\lim_{n\to\infty} \rho(P_n, P) = 0.$$

$\square$.

## COROLARIO 4.2.1

*Sea $(N_+, L)$ el espacio métrico de los enteros no negativos con la métrica del valor absoluto. Sea $\{P_n\}_{n=1}^{\infty} \subset \mathcal{P}(N_+)$ y $P \in \mathcal{P}(N_+)$. Las siguientes condiciones son equivalentes.*

*a)* $\lim_{n\to\infty} d(P_n, P) = 0$.

*b)* $\lim_{n\to\infty} \rho(P_n, P) = 0$.

*c)* $P_n \xrightarrow{D} P$.

*d)* $\lim_{n\to\infty} \int f\, dP_n = \int f dP \quad \forall f \in \overline{\mathfrak{f}}(N_+)$.

*e)* $\limsup_{n\to\infty} P_n(F) \leq P(F) \quad \forall F \in \mathcal{C}$.

)$\liminf_{n\to\infty} P_n(F) \geq P(F) \quad \forall F \in \mathcal{C}$.

*g)* $\lim_{n\to\infty} P_n(A) = P(A)$ *para todo conjunto $A$ de $P$-continuidad.*

*Demostración*
Por el teorema anterior sabemos que

$$\lim_{n\to\infty} \rho(P_n, P) = 0 \iff \lim_{n\to\infty} d(P_n, P) = 0,$$

además, las condiciones, de la (b) a la (g), son equivalentes si el espacio $(N_+, L)$ es separable, pero en este caso lo es porque $N_+$ es númerable. □

Gracias a este resultado, podemos afirmar que la sucesión de variables aleatorias Binomial dada en el *ejemplo 3.1.1* converge débilmente a una variable aleatoria Poisson con parámetro $\lambda$.

En general, si nosotros queremos demostrar la convergencia débil de una sucesión de variables aleatorias con valores en $N_+$ a otra variable aleatoria con valores en $N_+$, es suficiente mostrar la convergencia con la métrica $d$.

# Bibliografía


[1] Ash, Robert B. (1972).
"Real Analysis and Probability".
Academic Press Inc. Orlando Florida.

[2] Billingsley, P. (1968).
"Convergence of Probability Measures".
Wiley, New York.

[3] Chen L. H. Y. (1974).
"On the Converge of Poisson binomial to Poisson distribution".
Ann. Probability. **2**, 178-180.

[4] Chen L. H. Y. (1975).
"Poisson approximation for dependent trials".
Ibid. **3**, 534-545.

[5] Chow, Y. S. (1966).
"Some Convergence Theorems for Independence Ramdon Variables".
Ann. Math. Statist. **37**, 1482-1493.

[6] Chow, Y. S. and Robbins, H. (1961).
"On Sums of Independent Ramdon Rariables with Infinite Moments and Fair Games .
Proc. Nat. Acad. Sci. U.S.A. **47**, 330-335.

[7] Chow, Y. S. and Lai. T. L. (1973).
"Limiting Behavior of Weighted Sums of Independent Ramdon Variables.
Ann. of Prob. **1**, 810-824.

[8] Chung, K-L. (1968).
"A Course in Probability Theory".
Harcourt, New York.

[9] Chung, K-L., and Erdõ s, P. (1947).
"On the Lower Limit of Sums of Independent Ramdon Variables". Ann. of Math. **48**, 1003-1012.

[10] Cinlar E. (1972).
"Superposition of point processes, Stochastic Point Processes".
P. A. Lewis ed. pp. 549-606.

[11] Daley D. J. (1975)
Private Comunication.

[12] Dharmadhikari, S. W. and Jogdeo, K. (1969).
"Bounds on Moments of Sums of Independent Variables".
Ann. Math. Statist. **40**, 1506-1509.

[13] Durret, R. (1985).
"Probability: Theory and Examples".
Pacific Grove, California.

[14] Ethier, Stewart N. and Kurtz, Thomas G. (1968).
"Markov Processes, Caracterization and Convergence".
Wiley, New York.

[15] Freedman D. (1974).
"The Poisson approximation for dependent events". Ann. Probability. **2**, 256-269.

[16] Galambos J. (1973).
"A General Poisson Limit theorem of probability theory".
Duke Math. J. **40**, 486-581.

[17] Gastwirth J. L. (1977).
"A Probability model of a pyramid scheme".
Amer. Statist. **31**, 79-82.

[18] Gnedenko, B. V. and A. N. Kolmogorov. (1954).
"Limit Distributions for Sums of Independent Ramdon Variables".
Addison-Wesley, Reading, Mass.

[19] Heyde, C. C. (1968a).
"On Almost Sure Converge for Sums of Independent Ramdon Variables".
Sankhya Ser. A. **30**, 353-358.

[20] Kesten, H. (1972).
"Sums of Independent Ramdon Variables Without Moment Conditions".
Ann. Math. Statist. **43**, 701-732.

[21] Laha, Rohatgi. (1979).
"Probability Theory".
Wiley, New York.

[22] Le Cam L. (1960).
"An approximation theorem for the Poisson binomial distribution".
Pacific. J. Math. **10**, 1181-1197.

[23] Miller, H. D. (1967).
"A Note on Sums of Independent Ramdon Variables With Infinite First Moments".
Ann. Math. Statist. **38**, 751-758.

[24] Petrov, V. V. (1975).
"Sums of Independent Ramdon Variables".
Springer-Verlag, New York.

[25] Pruitt, W. E. (1966).
"Summability of independent Ramdon Variables".
J. Math. Mech. **15**, 769-776.

[26] Serfling R. J. (1975).
"A General approximation theorem".
Ann. Probability. **3**, 726-731.

[27] Stout, W. F. (1967).
"Some Results on the Complete and Almost Sure Converge of Linear Combinations of Independent Ramdon Variables and Martingales Differences".
Ann. Math. Statist. **39**, 1549-1562.

[28] Stout, W. F. (1974).
"Almost Sure Convergence".
Academic Press, New York.

[29] Strassen, V. (1965).
"Almost Sure Behavior of Sums of Independent Ramdon Variables And Martingales".
Proc. Symp. Math. Statist and Probability, 5th, Berkeley, 1965, **2**, 315-344.