Independencia condicionada. Propiedades de invarianza y aplicaciones

Andrés Felipe Muñoz Tello

Universidad del Valle
Facultad de Ciencias Naturales y Exactas
Departamento de Matemáticas
Santiago de Cali
2013

Independencia condicionada. Propiedades de invarianza y aplicaciones

Andrés Felipe Muñoz Tello

Trabajo de grado presentado como requisito
parcial para optar al título de Magister en Ciencias Matemáticas

Dirigido por:

Dr. Miguel A. Marmolejo L.

Universidad del Valle
Facultad de Ciencias Naturales y Exactas
Departamento de Matemáticas
Santiago de Cali
2013

Agradecimientos

Agradezco a mi madre, a mi hermana, a el Dr. Miguel Marmolejo y a quienes compartieron junto conmigo la experiencia de la maestría.

Andrés Felipe Muñoz Tello

Resumen

Este trabajo versa sobre la relación de la independencia condicionada, que se introduce a continuación. Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos y $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$; si para cada subconjunto finito $\{i_1, \cdots, i_k\}$ de I, $k \geq 2$, y cada escogencia $E_{i_j} \in \mathcal{E}_{i_j}$, $j = 1, 2, \cdots, k$, se cumple

$$P(\cap_{j=1}^{k} E_{i_j}|\mathcal{G}) = \prod_{j=1}^{k} P(E_{i_j}|\mathcal{G}) \quad (c.t.p.),$$

entonces se dice que la familia $\{\mathcal{E}_i\}_{i\in I}$ es condicionalmente independiente dada $\mathcal{G}$ (o simplemente $\mathcal{G}$ independiente). Esta relación juega un papel importante en la probabilidad y la estadística; en particular, la $\mathcal{G}$-independencia de dos sub-σ-álgebras de $\mathcal{F}$ interviene en la definición y estudio de los procesos markovianos y de los procesos recíprocos.

El principal aporte de este trabajo consiste en establecer algunas propiedades de invarianza de la relación de independencia condicionada; esto es, dar condiciones necesarias o suficientes para que se preserve la relación de independencia condicionada cuando se hacen cambios en uno de los tres objetos que intervienen en su definición: hacer más grande ó pequeña cada clase $\mathcal{E}_i$, hacer más grande ó pequeña la sub-σ-álgebra $\mathcal{G}$ o cambiar la función de probabilidad $P(\cdot)$. Estos resultados se pueden ver como generalizaciones de los establecidos por Van Putten y Van Schuppen [17] (caso $I = \{1, 2\}$) y de resultados conocidos sobre familias de eventos independientes (caso $\mathcal{G} = \{\emptyset, \Omega\}$). Como una aplicación de estas propiedades de invarianza se establecen versiones condicionadas de algunos resultados clásicos de la teoría de probabilidad. Los aportes fundamentales de este trabajo aparecen en el artículo *Versión condicionada de una generalización del lema de Borel-Cantelli* de Marmolejo et al. [10] y el artículo *Algunas propiedades de la independencia condicionada* de Marmolejo y Muñoz [11].

Índice general

Introducción

Sean $(\Omega,\mathcal{F},P)$ un espacio de probabilidad y $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos: $\mathcal{E}_i \subseteq \mathcal{F}$. Se dice que la familia $\{\mathcal{E}_i\}_{i\in I}$ es independiente; en símbolos: $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$, si para cada subconjunto finito $\{i_1, i_2, \cdots, i_k\}$ de I; $k \geq 2$, y cada escogencia $E_{i_1} \in \mathcal{E}_{i_1}$, $E_{i_2} \in \mathcal{E}_{i_2}, \cdots, E_{i_k} \in \mathcal{E}_{i_k}$ se cumple

$$P(\cap_{j=1}^{k} E_{i_j}) = \prod_{j=1}^{k} P(E_{i_j}). \tag{1}$$

Si $B \in \mathcal{F}$ es un evento de probabilidad positiva. La familia $\{\mathcal{E}_i\}_{i\in I}$ es localmente independiente sobre B (o B-independiente), en símbolos: $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp B$, si en (1) se cambia $P(\cdot)$ por $P(\cdot \mid B)$, esto es,

$$P(\cap_{j=1}^{k} E_{i_j} \mid B) = \prod_{j=1}^{k} P(E_{i_j} \mid B). \tag{2}$$

Sea ahora $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Una familia $\{\mathcal{E}_i\}_{i\in I}$ es independiente dada $\mathcal{G}$ (o $\mathcal{G}$-independiente); en símbolos, $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$, si para cada subconjunto finito $\{i_1, i_2, \cdots, i_k\}$ de I; $k \geq 2$, y cada escogencia $E_{i_1} \in \mathcal{E}_{i_1}, E_{i_2} \in \mathcal{E}_{i_2}, \cdots, E_{i_k} \in \mathcal{E}_{i_k}$, se tiene

$$P(\cap_{j=1}^{k} E_{i_j} \mid \mathcal{G}) = \prod_{j=1}^{k} P(E_{i_j} \mid \mathcal{G}) \quad (c.t.p.). \tag{3}$$

La anterior relación juega un papel importante en la probabilidad y en la estadística. En particular, la relación de $\mathcal{G}$-independencia de dos sub-σ-álgebras $\mathcal{F}_1$ y $\mathcal{F}_2$ de $\mathcal{F}$, interviene en la definición y estudio de las familias de σ-álgebras markovianas, de los procesos markovianos (ver Capítulo XIV de Loeve [8] y Gill [6]) y los procesos recíprocos o campos de Markov (ver Jamison [7]). En estadística, la $\mathcal{G}$-independencia de dos sub-σ-álgebras es útil en el estudio de la relación entre los conceptos de suficiencia e invarianza (ver Nogales y Oyola [12] y Nogales et al. [13]).

En el caso de la $\mathcal{G}$-independencia de dos σ-álgebras $\mathcal{F}_1$ y $\mathcal{F}_2$; $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$, Van Putten y Van Schuppen [17] dieron algunas condiciones necesarias o suficientes para que al efectuar ciertos cambios en uno de los tres objetos que intervienen en la definición de $\mathcal{G}$-independencia, se preserve la relación de independencia condicionada. Ahora al considerar una tercera σ-álgebra $\mathcal{F}_3$, puede suceder que se verifiquen las relaciones de $\mathcal{G}$-independencia dos a dos, es decir que $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$; $\mathcal{F}_1, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$ y $\mathcal{F}_2, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$, pero que no se verifique la relación $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$.

Por otra parte, en la última década se han reportado versiones condicionadas de algunos resultados clásicos de la teoría de probabilidad que involucran la independencia de una sucesión de variables aleatorias; El lema de Borel Cantelli, la desigualdad de Kolmogorov, la desigualdad de Hajek-Renyi, la ley de los grandes números, entre otros (Ver Majerek et al. [9] y Prakasa [15]).

Estos antecedentes, fueron el motivo para dedicar este trabajo al estudio de las propiedades de la independencia de una familia $\{\mathcal{F}_1\}_{i\in I}$ de σ-álgebras, en lugar de sólo dos sub-σ-álgebras $\mathcal{F}_1$ y $\mathcal{F}_2$, buscando generalizar algunos resultados de Van Putten y Van Schuppen y generalizar algunos resultados conocidos sobre familias de eventos independientes (caso $\mathcal{G} = \{\Omega, \emptyset\}$). También motivó a estudiar algunas versiones condicionadas del lema de Borel-Cantelli, mostrando que la versión condicionada del teorema de Feng et al. [5], es más general que la establecida por Yan [19] y que la establecida por Majerek et al. [9].

El trabajo se distribuye de la siguiente manera: En el Capítulo 1 se presenta la definición de independencia para eventos, familia de clases de eventos y variables aleatorias, mostrando algunas de sus propiedades (Secciones 1.1 y 1.3); también algunos tipos de convergencia de sucesiones de variables aleatorias (Sección 1.2); dedicando la Sección 1.4 a la probabildad y esperanza condicionada. El Capítulo 2 aborda el problema de la invarianza de la relación de independencia condicionada, considerando los cambios en la familia de clases de eventos $\{\mathcal{E}_i\}_{i\in I}$ (Sección 2.1), los cambios en la σ-álgebra condicionante $\mathcal{G}$ (Sección 2.2) y los cambios en la medida de probabilidad $P(\cdot)$ (Sección 2.3); en las dos últimas Secciones de este Capítulo se da un vistazo a la independencia condicionada de variables aleatorias y a las familas de σ-álgebras recíprocas y markovianas respectivamente. Finalmente en el Capítulo 3, se presentan versiones condicionadas de algunos resultados clásicos de la teoría de probabilidad, comenzando con algunas desigualdades condicionadas (Sección 3.1), continuando con algunas versiones condicionadas de el lema de Borel-Cantelli (Sección 3.2), luego con la desigualdad generalizada de Kolmogorov (Sección 3.3), terminando con la ley fuerte de los grandes números condicionada (Sección 3.4).

Los aportes fundamentales de este trabajo están constituidos por la mayor parte del Capítulo 2 y por la Sección 3.2 del Capítulo 3. Los resultados aparecen en el artículo *Versión condicionada de una generalización del lema de Borel-Cantelli* de Marmolejo et al. [10] y el artículo *Algunas propiedades de la independencia condicionada* de Marmolejo y Muñoz [11].

Capítulo 1

Preliminares

El objeto de este Capítulo es el de presentar las definiciones y resultados necesarios para el desarrollo del resto del trabajo, haciendo seguimiento de lo hecho en Bauer [2], Shiryaev [16], Williams [18], Ash [1] y Billingsley [3]. En la Sección 1.1 se da la definición de independencia de una familia de eventos y se establecen algunas de sus propiedades. En la Sección 1.2 se revisan las nociones de convergencia en probabilidad y convergencia en casi todas partes de una sucesión de variables aleatorias. La Sección 1.3 se dedica a la independencia de variables aleatorias, aquí se destaca la Ley Débil de los Grandes Números. En la Sección 1.4 se presenta la noción de esperanza condicionada de una variable aleatoria con respecto a un evento, luego con respecto a una descomposición finita de espacio muestral y por último con respecto a una σ-álgebra. Se destacan las propiedades más importantes de la esperanza condicionada y se acompañan de ejemplos ilustrativos.

1.1. Independencia

Dado $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, se dice que dos eventos $A, B \in \mathcal{F}$ son independientes, si $P(A \cap B) = P(A)P(B)$. La siguiente definición es una generalización de lo anterior.

Definición 1.1.1 *Sea $\{\mathcal{E}_i\}_{i\in I}$ es una familia de clases de eventos; $\mathcal{E}_i \subseteq \mathcal{F}$. Se dice que la familia $\{\mathcal{E}_i\}_{i\in I}$ es independiente; en símbolos $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$, si para cada subconjunto finito $\{i_1, i_2, \cdots, i_k\}$ de I; $k \geq 2$, y cada escogencia $E_{i_1} \in \mathcal{E}_{i_1}, \cdots, E_{i_k} \in \mathcal{E}_{i_k}$ se cumple*

$$P(\cap_{j=1}^{k} E_{ij}) = \prod_{j=1}^{k} P(E_{ij}).$$

En particular, dos sub-σ-álgebras $\mathcal{G}_1$ y $\mathcal{G}_2$ de una σ-álgebra $\mathcal{F}$, son independientes si para cada $A \in \mathcal{G}_1$ y $B \in \mathcal{G}_2$ se tiene que A y B son eventos independientes.

Un ejemplo clásico de una sucesión de eventos independientes es el siguiente (ver Bauer [16]).

Ejemplo 1.1.2 *Dado $([0,1), \mathcal{B}([0,1), \lambda))$ el espacio de medida de Lebesgue-Borel uno dimensional. Si para todo $i = 1, 2, \cdots$, se define*

$$A_i = \left[0, \frac{1}{2^i}\right) \cup \left[\frac{2}{2^i}, \frac{3}{2^i}\right) \cup \cdots \cup \left[\frac{2^i-2}{2^i}, \frac{2^i-1}{2^i}\right),$$

entonces $P(A_i) = \frac{1}{2}$, además para todo $n \in \mathbb{N}$ se tiene

$$P(A_{i_1} \cap \cdots \cap A_{i_n}) = \frac{1}{2} P(A_{i_1} \cap \cdots \cap A_{i_{n-1}}).$$

Por tanto para todo subconjunto finito no vacío $\{i_1, i_2, \cdots i_n\}$ de $\mathbb{N}$ se tiene que

$$P(A_{i_1} \cap \cdots \cap A_{i_n}) = P(A_{i_1} \cap \cdots \cap A_{i_{n-1}}) P(A_{i_n}) = P(A_{i_1}) \cdots P(A_{i_{n-1}}) P(A_{i_n}),$$

es decir, la sucesión de eventos $\{A_n\}_{n \in \mathbb{N}}$ es independiente.

Observación 1.1.3 *Sean $\mathcal{E}_1$ y $\mathcal{E}_2$ dos clases de eventos tales que $\mathcal{E}_1 \subseteq \mathcal{E}_2$. Entonces $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp$ si y sólo si $P(A) \in \{0,1\}$, para todo $A \in \mathcal{E}_1$.*

En efecto, si $\mathcal{E}_1$ y $\mathcal{E}_2$ son independientes y $A \in \mathcal{E}_1 \subseteq \mathcal{E}_2$, entonces $P(A) = P(A \cap A) = P(A)P(A) = P^2(A)$. De aquí que $P(A) \in \{0,1\}$. Recíprocamente; si para todo $A \in \mathcal{E}_1$ se tiene $P(A) \in \{0,1\}$, entonces $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp$; pues para cada $A \in \mathcal{E}_1$ y cada $B \in \mathcal{E}_2$, $P(A \cap B) = P(A)P(B)$.

A continuación se definen los sistemas de Dynkin y los π-sistemas, los cuales facilitan el trabajo con las σ-álgebras.

Definición 1.1.4 *Un sistema $\mathbf{D}$ de subconjuntos de Ω es un sistema de Dynkin, si satisface las siguientes condiciones:*

(1) $\Omega \in \mathbf{D}$.

(2) Si $A, B \in \mathbf{D}$ y $B \subseteq A$, entonces $A - B \in \mathbf{D}$.

(3) Si $A_1, A_2, \cdots \in \mathbf{D}$ y $A_i \cap A_j = \emptyset$ para todo $i \neq j$ entonces $\cup_{n=1}^{\infty} A_n \in \mathbf{D}$.

Dada una colección $\mathcal{E}$ de subconjuntos de Ω, el sistema de Dynkin generada por ella, denotado por $\mathbf{D}(\mathcal{E})$, es el menor sistema de Dynkin que contiene a $\mathcal{E}$.

Definición 1.1.5 *Un sistema de eventos Υ es un π-sistema, si $A, B \in \Upsilon$ implica $A \cap B \in \Upsilon$.*

La demostración del siguiente teorema se encuentra en la Sección 1.2 del Capítulo I del libro de Bauer [2].

Teorema 1.1.6 *Un sistema de Dynkin $\mathbf{D}$ es una σ-álgebra si y sólo si $\mathbf{D}$ es un π-sistema.*

Teorema 1.1.7 *Sea $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos. Si $\mathbf{D}(\mathcal{E}_i)$ es el sistema de Dynkin generado por $\mathcal{E}_i$, entonces $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$ si y sólo si $\{\mathbf{D}(\mathcal{E}_i)\}_{i\in I} \perp\!\!\!\perp$.*

Demostración. ($\Rightarrow$) Por la definición de independencia se puede suponer que I es finito; $I = \{1, 2, \cdots, n\}$. Si $\mathbb{D}_1 := \{E \in \mathcal{F} : \{E\}, \mathcal{E}_2, \cdots, \mathcal{E}_n \perp\!\!\!\perp\}$, $\{i_1, i_2, \cdots, i_k\}$ es un subconjunto no vacío de $\{2, \cdots, n\}$ y $E_{i_j} \in \mathcal{E}_{i_j}$, $j = 2, \cdots, k$, entonces $\mathbb{D}_1$ es un sistema de Dynkin. En efecto:

(1) $\Omega \in \mathbb{D}_1$, pues cumple

$$P(\Omega \cap E_{i_1} \cap \cdots \cap E_{i_k}) = P(E_{i_1} \cap \cdots \cap E_{i_k}) = P(\Omega)P(E_{i_1}) \cdots P(E_{i_k}).$$

(2) Si $E, F \in \mathbb{D}_1$ con $E \subseteq F$ entonces $F - E \in \mathbb{D}_1$; ya que se verifica

$$\begin{aligned} P((F-E) \cap E_{i_1} \cap \cdots \cap E_{i_k}) &= P(F \cap E_{i_1} \cap \cdots \cap E_{i_k} - E \cap E_{i_1} \cap \cdots \cap E_{i_k}) \\ &= P(F)P(E_{i_1}) \cdots P(E_{i_k}) - P(E)P(E_{i_1}) \cdots P(E_{i_k}) \\ &= (P(F) - P(E))P(E_{i_1}) \cdots P(E_{i_k}) \\ &= P(F-E)P(E_{i_1}) \cdots P(E_{i_k}). \end{aligned}$$

(3) Dada $\{E_m\}_{m\in\mathbb{N}} \subseteq \mathbb{D}_1$, una sucesión disjunta, entonces $(\uplus_{m=1}^{\infty} E_m) \in \mathbb{D}_1$. En efecto

$$\begin{aligned} P((\uplus_{m=1}^{\infty} E_m) \cap E_{i_1} \cap \cdots \cap E_{i_k}) &= P(\uplus_{m=1}^{\infty}(E_m \cap E_{i_1} \cap \cdots \cap E_{i_k})) \\ &= \sum_{m=1}^{\infty} P(E_m \cap E_{i_1} \cap \cdots \cap E_{i_k}) \\ &= \sum_{m=1}^{\infty} P(E_m)P(E_{i_1}) \cdots P(E_{i_k}) \\ &= P(\uplus_{m=1}^{\infty} E_m)P(A_{i_1}) \cdots P(A_{i_k}). \end{aligned}$$

Como $\mathcal{E}_1 \subseteq \mathbb{D}_1$, entonces $\mathbf{D}(\mathcal{E}_1) \subseteq \mathbb{D}_1$ y $\mathbf{D}(\mathcal{E}_1), \mathcal{E}_2, \cdots \mathcal{E}_n \perp\!\!\!\perp$. Repitiendo el argumento $n-1$ veces se concluye que $\mathbf{D}(\mathcal{E}_1), \mathbf{D}(\mathcal{E}_2), \cdots, \mathbf{D}(\mathcal{E}_n) \perp\!\!\!\perp$.

($\Leftarrow$) Basta tener presente que $\mathcal{E}_i \subseteq \mathbf{D}(\mathcal{E}_i)$ para cada $i \in I$.$\square$

Corolario 1.1.8 *Sea $\{\mathcal{E}_i\}_{i\in I}$ una familia de π-sistemas, $\mathcal{E}_i \subseteq \mathcal{F}$. Si $\sigma(\mathcal{E}_i)$ es la σ-álgebra generada por $\mathcal{E}_i$, entonces $\{\sigma(\mathcal{E}_i)\}_{i\in I} \perp\!\!\!\perp$ si y sólo si $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$.*

Demostración. Como $\mathcal{E}_i$ es un π-sistema, entonces por el teorema 1.1.6 $\sigma(\mathcal{E}_i) = \mathbf{D}(\mathcal{E}_i)$ y el resultado se obtiene a partir del teorema anterior.$\square$

El siguiente ejemplo muestra la importancia de que $\mathcal{E}_i$ sea un π-sistema en el corolario anterior.

Ejemplo 1.1.9 *Dado el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ donde $\Omega = \{1,2,3,4\}$, $\mathcal{F} = \wp(\Omega)$ y $P(\{i\}) = \frac{1}{4}$; $i = 1,2,3,4$. Sea $\mathcal{E}_1 = \{\{1,2\}\}$ y $\mathcal{E}_2 = \{\{1,3\},\{1,4\}\}$. Es claro que $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp$. También del teorema 1.1.7 $\boldsymbol{D}(\mathcal{E}_1), \boldsymbol{D}(\mathcal{E}_2) \perp\!\!\!\perp$. Sin embargo $\mathcal{E}_2$ no es un π-sistema y tampoco $\sigma(\mathcal{E}_1)$ y $\sigma(\mathcal{E}_2)$ son independientes, pues al tomar los eventos $\{1,2\} \in \sigma(\mathcal{E}_1)$ y $\{3\} \in \sigma(\mathcal{E}_2)$, éstos no son independientes.*

Corolario 1.1.10 *Dados $\{\mathcal{E}_i\}_{i\in I}$ una familia de π-sistemas $\mathcal{E}_i \subseteq \mathcal{F}$ tal que $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$ y $\{I_j\}_{j\in J}$ una partición de I; esto es, $I_j \neq \emptyset$ y $\uplus_{j\in J} I_j = I$. Si $\mathcal{F}_j := \sigma(\cup_{i\in I_j} \mathcal{E}_i)$, entonces $\{\mathcal{F}_j\}_{j\in J} \perp\!\!\!\perp$.*

Demostración. Para cada $j \in J$, sea $\mathcal{E}_j^*$ el sistema de eventos $E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}$, donde $\{i_1, i_2, \cdots, i_k\} \subseteq I_j$ y $E_{i_r} \in \mathcal{E}_{i_r}$, $r = 1,2,\cdots,k$. Entonces $\mathcal{E}_j^*$ es un π-sistema tal que $\sigma(\mathcal{E}_j^*) = \mathcal{F}_j$. Ahora, como $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$, por la definición de independencia se tiene $\{\mathcal{E}_j^*\}_{j\in J} \perp\!\!\!\perp$ y por el Corolario 1.1.8, $\{\mathcal{F}_j\}_{j\in J} \perp\!\!\!\perp$.$\square$

Definición 1.1.11 *Dada $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ una sucesión de σ-álgebras en $\mathcal{F}$ y $\mathcal{I}_n = \sigma(\cup_{m=n}^{\infty} \mathcal{F}_m)$, entonces $\mathcal{I}_\infty := \cap_{n=1}^{\infty} \mathcal{I}_n$ es llamada la σ-álgebra de eventos terminales de la sucesión $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$.*

Ejemplo 1.1.12 *Si $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias, y $J_n = \sigma(X_n, X_{n+1}, \cdots)$, entonces para todo $n \geq 1$*

$$A = \left\{ \sum_{i=1}^{\infty} \frac{X_i}{i} < \infty \right\} = \left\{ \sum_{i=n}^{\infty} \frac{X_i}{i} < \infty \right\} \in J_n;$$

esto es, $A \in \cap_{n=1}^{\infty} J_n = \cap_{n=1}^{\infty} \sigma(\cup_{m=n}^{\infty} \{X_m^{-1}(B), B \in \mathcal{B}(\mathbb{R})\})$*, es decir* A *es un evento terminal de la sucesión de* σ*-álgebras* $\{\sigma(X_n)\}_{n\in\mathbb{N}}$.

Teorema 1.1.13 (Ley cero-uno de Kolmogorov) *Sea* $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ *una sucesión de* σ*-álgebras independientes en* $\mathcal{F}$*, entonces* $P(A) = 0$ *ó* $P(A) = 1$ *para todo evento terminal* A *de la sucesión.*

Demostración. Sean $A \in \mathcal{I}_\infty$ y $\mathbb{D}$ el conjunto de todos los eventos independientes de A, es decir, todos los $D \in \mathcal{F}$ tal que $P(A \cap D) = P(A)P(D)$. Es fácil verificar que $\mathbb{D}$ es un sistema de Dynkin. Como del corolario 1.1.10 $\mathcal{I}_{n+1} = \sigma(\cup_{m=n+1}^{\infty} \mathcal{F}_m)$ es independiente de $\mathcal{U}_n = \sigma(\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n)$ y como $A \in \mathcal{I}_\infty$, entonces $\mathcal{U}_n \subset \mathbb{D}$ para todo $n \in \mathbb{N}$; lo que implica $\mathcal{U}_0 \equiv \bigcup_{n=1}^{\infty} \mathcal{U}_n \subset \mathbb{D}$. Puesto $\mathcal{U}_0$ es un π-sistema se tiene por corolario 1.1.8 $\sigma(\mathcal{U}_0) = \mathbf{D}(\mathcal{U}_0) \subset \mathbb{D}$. Como $\mathcal{I}_n \subset \sigma(\mathcal{U}_0)$ para todo $n \in \mathbb{N}$, entonces $A \in \mathcal{I}_\infty \subset \sigma(\mathcal{U}_0)$. Tomando $D = A$, se obtiene $P(A) = P(A \cap A) = P(A)P(A) = P(A)^2$ y por tanto $P(A) = 0$ ó $P(A) = 1$.$\square$

Ejemplo 1.1.14 *Sea* $\{A_n\}_{n\in\mathbb{N}}$ *una sucesión de eventos independientes, entonces* $P(\limsup A_n) = 0$ *ó* $P(\limsup A_n) = 1$*. En efecto, si se toma* $\mathcal{U}_n = \{\emptyset, A_n, A_n^c, \Omega\}$ *la* σ*-álgebra generada por* A_n*, entonces por el corolario 1.1.8 la sucesión de* σ*-álgebras* $\{\mathcal{U}_n\}_{n\in\mathbb{N}}$ *es independiente. Como* $\mathcal{I}_n = \sigma(\bigcup_{m=n}^{\infty} \mathcal{U}_m)$ *es una sucesión decreciente de* σ*-álgebras, entonces* $\limsup A_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m \in \mathcal{I}_n$ *para todo* n*, es decir* $\limsup A_n \in \mathcal{I}_\infty$*. Basta aplicar el teorema 1.1.13.*

En el ejemplo anterior, para $\{A_n\}_{n\in\mathbb{N}}$ suceción de eventos independientes se dedujo que $P(\limsup A_n)$ es cero ó uno. El lema de Borel-Cantelli, enunciado a continuación, va más alla y da un criterio específico en términos de la convergencia ó divergencia de la serie $\sum_n P(A_n)$.

Lema 1.1.15 (de Borel-Cantelli) *Dada una sucesión de eventos* $\{A_n\}_{n\in\mathbb{N}}$*, se tiene:*

(a) *Si* $\sum_n P(A_n) < \infty$*, entonces* $P(\limsup A_n) = 0$.

(b) *Si* $\{A_n\}_{n\in\mathbb{N}}$ *es una sucesión de eventos independientes y* $\sum_n P(A_n) = \infty$*, entonces* $P(\limsup A_n) = 1$.

Demostración.
(a) $P(\limsup A_n) = P(\cap_{n=1}^{\infty} \cup_{k\geq n} A_k) = \lim_n P(\cup_{k\geq n} A_k) \leq \lim_n \sum_{k\geq n} P(A_k) = 0.$

(b) Para probar lo requerido es suficiente con ver que $P(\liminf A_n^c) = 0$; para esto basta probar que $P(\cap_{k=n}^{\infty} A_k^c) = 0$ para todo n. Dado que $\{A_n\}_{n\in\mathbb{N}}$ es una sucesión de eventos independientes, entonces del corolario 1.1.8 $\{A_n^c\}_{n\in\mathbb{N}}$ son independientes; como $\sum_n P(A_n) = \infty$ y $1 - x \leq e^{-x}$, se tiene

$$P(\cap_{k=n}^{n+j} A_k^c) = \prod_{k=n}^{n+j}(1 - P(A_k)) \leq \prod_{k=n}^{n+j} e^{-P(A_k)} \to 0,$$

cuando $j \to \infty$, es decir $P(\cap_{k=n}^{\infty} A_k^c) = \lim_j P(\cap_{k=n}^{n+j} A_k^c) = 0.\square$

1.2. Tipos de convergencia

En la teoría de probabildad como en el análisis, se necesita de varios tipos de convergencia, en este caso sobre variables aleatorias, los siguientes dos tipos de convergencia son los que se usaran en las leyes de los grandes números.

- Convergencia en probabildad.
- Convergencia en casi todas partes o de probabildad uno.

En un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ se dice que una propiedad es valida en casi todas partes (c.t.p.) si la propiedad se cumple excepto en un evento $A \in \mathcal{F}$ de probabilidad nula. En lo que sigue, $X, X_1, X_2 \cdots$ son variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$.

Definición 1.2.1 *La sucesión $\{X_n\}_{n\in\mathbb{N}}$ converge en probabildad a X, $X_n \xrightarrow{P} X$, si para todo $\epsilon > 0$*

$$P(|X_n - X| > \epsilon) \to 0, \quad n \to \infty.$$

Definición 1.2.2 *Una sucesión $\{X_n\}_{n\in\mathbb{N}}$ converge en casi todas partes a X, $X_n \xrightarrow{c.t.p.} X$, si*

$$P(\omega : X_n \nrightarrow X) = 0.$$

Además como en el análisis se necesitara del concepto de sucesión de Cauchy, aquí llamada sucesión fundamental, las cuales se definen según sea el caso.

Definición 1.2.3 *Una sucesión $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión fundamental en probabildad, si para $m, n \to \infty$ y todo $\epsilon > 0$*

$$P(|X_n - X_m| > \epsilon) \to 0,$$

y es fundamental en casi todas partes, si $\{X_n(\omega)\}_{n\in\mathbb{N}}$ es fundamental para casi todo $\omega \in \Omega$.

Teorema 1.2.4

(a) $X_n \xrightarrow{c.t.p.} X$ *si sólo si para todo $\epsilon > 0$*

$$P\left(\sup_{k\geq n} |X_k - X| \geq \epsilon\right) \to 0, \quad n \to \infty.$$

(b) $\{X_n\}_{n\in\mathbb{N}}$ *es fundamental en casi todas partes si y sólo si para todo $\epsilon > 0$*

$$P\left(\sup_{k,l\geq n} |X_k - X_l| \geq \epsilon\right) \to 0, \quad n \to \infty. \tag{1.1}$$

O equivalentemente

$$P\left(\sup_{k\geq 0} |X_{n+k} - X_n| \geq \epsilon\right) \to 0, \quad n \to \infty. \tag{1.2}$$

Demostración. (a) Sean $A_n^\epsilon = \{\omega : |X_n - X| \geq \epsilon\}$ y $A^\epsilon = \limsup A_n^\epsilon = \cap_{n=1}^\infty \cup_{k\geq n} A_k^\epsilon$. Como $\omega_0 \in \{\omega : X_n \nrightarrow X\}$ equivale a que $\lim_{n\to\infty} X_n(\omega_0) \neq X(\omega_0)$; existe $\epsilon > 0$ tal que para todo $n \in \mathbb{N}$ existe $k \geq n$ donde $|X_k(\omega_0) - X(\omega_0)| \geq \epsilon$, entonces

$$\{\omega : X_n \nrightarrow X\} = \cup_{\epsilon>0} A^\epsilon = \cup_{m=1}^\infty A^{\frac{1}{m}},$$

por tanto $X_n \xrightarrow{c.t.p.} X$ equivale a $P(\cup_{m=1}^\infty A^{\frac{1}{m}}) = 0$, lo cual se cumple si y sólo si $P(A^{\frac{1}{m}}) = 0$, $m \geq 1$; siendo igual a decir que para todo $\epsilon > 0$, $P(A^\epsilon) = 0$. Además como $\cup_{k\geq n} A_k^\epsilon \downarrow A^\epsilon$ implica $P(A^\epsilon) = \lim_{n\to\infty} P(\cup_{k\geq n} A_k^\epsilon)$; se tiene que para todo $\epsilon > 0$

$$\lim_{n\to\infty} P(\cup_{k\geq n} A_k^\epsilon) = 0,$$

es decir que para todo $\epsilon > 0$

$$P\left(\sup_{k\geq n} |X_k - X| \geq \epsilon\right) \to 0 \quad n \to \infty.$$

(b) Dado $B^{\epsilon}_{k,l} = \{\omega : |X_k - X_l| \geq \epsilon\}$ y $B_\epsilon = \cap_{n=1}^{\infty} \cup_{k,l\geq n} B^{\epsilon}_{k,l}$, se procede de la misma forma que en (a) y se obtiene el resultado. La equivalencia entre (1.1) y (1.2) se desprende de la siguiente desigualdad

$$\sup_{k\geq 0} |X_{n+k} - X_n| \leq \sup_{k,l\geq 0} |X_{n+k} - X_{n+l}| \leq 2 \sup_{k\geq 0} |X_{n+k} - X_n|.\square$$

Corolario 1.2.5 *Si para todo $\epsilon > 0$ se satisface $\sum_{k=1}^{\infty} P(|X_k - X| \geq \epsilon) < \infty$ entonces $X_n \xrightarrow{c.t.p.} X$.*

Demostración. El resultado es consecuencia inmediata de la siguiente desigualdad

$$P\left(\sup_{k\geq n} |X_k - X| \geq \epsilon\right) = P(\cup_{k\geq n} |X_k - X| \geq \epsilon) \leq \sum_{k\geq n} P(|X_k - X| \geq \epsilon).\square$$

Teorema 1.2.6 *Si $X_n \xrightarrow{c.t.p.} X$ entonces $X_n \xrightarrow{P} X$.*

Demostración. Si $X_n \xrightarrow{c.t.p.} X$ del teorema 1.2.4 se tiene $P\left(\sup_{k\geq n} |X_k - X| \geq \epsilon\right) \to 0$ cuando $n \to \infty$, lo cual equivale a $P(\cup_{k\geq n}\{\omega : |X_k - X| \geq \epsilon\}) \to 0$ cuando $n \to \infty$, entonces $P(\{\omega : |X_k - X| > \epsilon\}) \to 0$, cuando $n \to \infty$, es decir $X_n \xrightarrow{P} X.\square$

Ejemplo 1.2.7 *$X_n \xrightarrow{P} X$ no implica que $X_n \xrightarrow{c.t.p.} X$; ya que al tomar $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$ y P la medida de Lebesgue. Si $A^i_n = [\frac{i-1}{n}, \frac{i}{n}]$, $X^i_n = I_{A^i_n}(\omega)$, $1 \leq i \leq n$; $n \geq 1$, entonces la sucesión*

$$\{X^1_1, X^1_2, X^2_2, X^1_3, X^2_3, X^3_3, \cdots\} = \{I_{[0,1]}, I_{[0,\frac{1}{2}]}, I_{[\frac{1}{2},1]}, I_{[0,\frac{1}{3}]}, I_{[\frac{1}{3},\frac{2}{3}]}, I_{[\frac{2}{3},1]}, \cdots\}$$

es convergente en probabilidad a $X = 0$, pues de la desigualdad de Chebyshev para todo $\epsilon > 0$

$$\lim_{n\to\infty} P(|X^i_n| > \epsilon) < \lim_{n\to\infty} \frac{E|X^i_n|}{\epsilon} = \lim_{n\to\infty} \frac{P([\frac{i-1}{n}, \frac{i}{n}])}{\epsilon} = \lim_{n\to\infty} \frac{1}{n\epsilon} = 0.$$

Y no converge en casi todas partes, pues al tomar algun $\omega_0 \in [0,1]$ la sucesión de números reales $\{X^1_1(\omega_0), X^1_2(\omega_0), X^2_2(\omega_0), X^1_3(\omega_0), X^2_3(\omega_0), X^3_3(\omega_0), \cdots\}$ toma valores ceros y unos.

1.3. Independencia de variables aleatorias

Si una familia de eventos $\{E_i\}_{i\in I}$ es independiente entonces del corolario 1.1.8 la familia $\{\mathcal{F}_i\}_{i\in I}$ de σ-álgebras $\mathcal{F}_i = \{\emptyset, E_i, E^c_i, \Omega\}$ es independiente. Pero $\mathcal{F}_i$ es la σ-álgebra generada por la variable aleatoria indicadora I_{E_i}, entonces la independencia de $\{E_i\}_{i\in I}$ implica la de $\{\sigma(I_{E_i})\}_{i\in I}$. Este resultado justifica la siguiente definición.

Definición 1.3.1 *Una familia $\{X_i\}_{i\in I}$ de variables aleatorias es independiente; si la familia $\{\sigma(X_i)\}_{i\in I}$ de σ-álgebras generadas por las variables aleatorias X_i es independiente.*

Teorema 1.3.2 *Dado $(\Omega, \mathcal{F}, P)$, un espacio de probabilidad, $\{X_i\}_{i\in I}$ una familia de variables aleatorias con imagenes sobre el espacio de medida $(\Omega_i, \mathcal{F}_i)$ y sea $\mathcal{D}_i$ un π-sistema que genera a $\mathcal{F}_i$, con $\Omega_i \in \mathcal{D}_i$, $i \in I$. La familia $\{X_i\}_{i\in I}$ es independiente si y sólo si*

$$P(\cap_{j=1}^{k} X_{i_j}^{-1}(Q_{i_j})) = \prod_{j=1}^{k} P(X_{i_j}^{-1}(Q_{i_j})),$$

para cualquier escogencia finita $X_{i_1}, X_{i_2}, \cdots, X_{i_k}$ de $\{X_i\}_{i\in I}$ y cualquier $Q_{i_j} \in \mathcal{D}_{i_j}$, $j = 1, 2, \cdots, k$.

Demostración. Si $\mathbb{D}_i = \{X_i^{-1}(Q_i) : Q_i \in \mathcal{D}_i\}$, $i \in I$, entonces $\mathbb{D}_i$ es un generador de $\sigma(X_i)$. Como $\mathcal{D}_i$ es un π-sistema, $\Omega_i \in \mathcal{D}_i$, $i \in I$, y cada X_i es medible, se tiene que $\mathbb{D}_i$ es un π-sistema y $\Omega \in \mathbb{D}_i$, $i \in I$. Por lo tanto al aplicar el corolario 1.1.8 se obtiene que $\{\mathbb{D}_i\}_{i\in I} \perp\!\!\!\perp$ si y sólo si $\{\sigma(X_i)\}_{i\in I} \perp\!\!\!\perp$, concluyendo lo deseado.$\square$

Observación 1.3.3 *Sean $X_1, X_2, \cdots, X_n$ variables aleatorias definidas sobre el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y con imagenes sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Dado que el π sistema $\{(-\infty, x] : x \in \mathbb{R}\}$ genera a $\mathcal{B}(\mathbb{R})$, entonces las variables aleatorias $X_1, \cdots, X_n$ son independientes si y sólo si se verifica*

$$P(\cap_{j=1}^{n} X_j^{-1}((-\infty, x_j]) = \prod_{j=1}^{n} P(X_j^{-1}((-\infty, x_j]).$$

O lo que es lo mismo,

$$F_{(X_1,\ldots,X_n)}(x_1, \cdots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n),$$

donde $F_{(X_1,\ldots,X_n)}$ es la función de distribución acumulada n-dimensional del vector aleatorio $(X_1, \cdots, X_n)$ y F_{X_i} es la funcion de distribución de la variable aleatoria X_i.

Teorema 1.3.4 *Sea $\{X_i\}_{i\in I}$ una familia de variables aleatorias con imagenes en el espacio $(\Omega_i, \mathcal{F}_i)$ y sea $Y_i : (\Omega_i, \mathcal{F}_i) \to (\Omega'_i, \mathcal{F}'_i)$ una función medible, para todo $i \in I$, entonces la independencia de $\{X_i\}_{i\in I}$ implica la independencia de $\{(Y_i \circ X_i)\}_{i\in I}$.*

Demostración. Dado que para cualquier $A' \in \mathcal{F}'_i$, se tiene que $(Y_i \circ X_i)^{-1}(A') = X_i^{-1}(Y_i^{-1}(A'))$, entonces $\sigma(Y_i \circ X_i) \subseteq \sigma(X_i)$. Por tanto $\{\sigma(X_i)\}_{i\in I} \perp\!\!\!\perp$ implica $\{\sigma(Y_i \circ X_i)\}_{i\in I} \perp\!\!\!\perp$.□

Los siguientes resultados son clásicos de la teoría de integración, estos serán enunciados en este trabajo por su utilidad y su demostraciones se encuentran en la Sección 6 del Capítulo II del libro de Shiryaev [16].

Proposición 1.3.5 (Desigualdad de Chebyshev) *Sea X una variable aleatoria y sea $g \geq 0$ una función no decreciente tal que $g(X)$ es una variable aleatoria con esperanza finita. Entonces para cualquier $\epsilon > 0$ se tiene*

$$P(X \geq \epsilon) \leq \frac{E(g(X))}{g(\epsilon)}.$$

Corolario 1.3.6 *Si X una variable aleatoria con media μ y varianza $\sigma^2 < \infty$, entonces para cualquier $\epsilon > 0$ se tiene*

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Teorema 1.3.7 (Cambio de variables en la integral de Lebesgue) *Sea $(\Omega, \mathcal{F})$ y $(E, \mathcal{E})$ espacios medibles y $X = X(\omega)$ una función $\mathcal{F}/\mathcal{E}$-medible con valores en E. Sea P una medida de probabilidad sobre $(\Omega, \mathcal{F})$ y P_X la medida de probabilidad sobre $(E, \mathcal{E})$ inducida por $X = X(\omega)$*

$$P_X(A) = P\{\omega : X(\omega) \in A\}, \quad A \in \mathcal{E},$$

entonces

$$\int_A g(x) P_X(dx) = \int_{X^{-1}(A)} g(X(\omega)) P(d\omega), \quad A \in \mathcal{E},$$

para toda función $\mathcal{E}$-medible $g = g(x)$, $x \in E$ (En el sentido que si una integral existe, la otra está bien definida y ambas son iguales).

Teorema 1.3.8 (de Fubini) *Dado el espacio de medida $(\Omega, \mathcal{F}, \mu)$, donde $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ con $\mu = \mu_1 \times \mu_2$ el producto directo de medidas μ_1 y μ_2, es decir la medida sobre $\mathcal{F}$ tal que $\mu(A \times B) = \mu_1 \times \mu_2(A \times B) = \mu_1(A) \times \mu_2(B)$, $A \in \mathcal{F}_1, B \in \mathcal{F}_2$. Si $X = X(\omega_1, \omega_2)$ una función $\mathcal{F}_1 \otimes \mathcal{F}_2$-medible, integrable con respecto a la medida $\mu_1 \times \mu_2$:*

$$\int_{\Omega_1 \times \Omega_2} |X(\omega_1, \omega_2)| \, d(\mu_1 \otimes \mu_2) < \infty.$$

Entonces las integrales $\int_{\Omega_1} X(\omega_1, \omega_2) \mu_1(d\omega_1)$ y $\int_{\Omega_2} X(\omega_1, \omega_2) \mu_2(d\omega_2)$

(1) están definidas para todo ω_1 *y* ω_2*;*

(2) son funciones $\mathcal{F}_2$ *y* $\mathcal{F}_1$ *medibles respectivamente con*

$$\mu_2\left\{\omega_2 : \int_{\Omega_1} |X(\omega_1,\omega_2)|\,\mu_1(d\omega_1) = \infty\right\} = 0,$$

$$\mu_1\left\{\omega_1 : \int_{\Omega_2} |X(\omega_1,\omega_2)|\,\mu_2(d\omega_2) = \infty\right\} = 0;$$

(3)

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} X(\omega_1,\omega_2)d(\mu_1\otimes\mu_2) &= \int_{\Omega_2}(\int_{\Omega_1} X(\omega_1,\omega_2)\mu_1(d\omega_1))\mu_2(d\omega_2)\\ &= \int_{\Omega_1}(\int_{\Omega_2} X(\omega_1,\omega_2)\mu_2(d\omega_2))\mu_1(d\omega_1).\end{aligned}$$

Teorema 1.3.9 *Sean* $X_1,\cdots,X_n$ *variables aleatorias independientes. Si* $E(X_i)$ *es finito para todo* i*, entonces* $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i)$ *y el producto* $\prod_{i=1}^n X_i$ *es integrable.*

Demostración. Si se toma $P = \otimes_{i=1}^n P_{X_i}$. Entonces por la observación 1.3.3, los teoremas 1.3.7 y 1.3.8, se tiene

$$\begin{aligned}E(|\prod_{i=1}^n X_i|) &= \int_\Omega |X_1\cdots X_n|dP = \int_{\mathbb{R}^n} |x_1\cdots x_n|P_{(X_1,\cdots,X_n)}(d(x_1,\cdots x_n))\\ &= \int_{\mathbb{R}^n} |x_1|\cdots|x_n|P_{X_1}d(x_1)\cdots P_{X_n}d(x_n)\\ &= \int_{\mathbb{R}} |x_1|P_{X_1}d(x_1)\cdots\int_{\mathbb{R}} |x_n|P_{X_n}d(x_n) = \prod_{i=1}^n E(|X_i|).\end{aligned}$$

Por 1.3.8 los anteriores cálculos siguen siendo validos si se remueven los valores absolutos. Finalmente si $E(X_i)$ es finita, entonces $E|X_i|$ finita y lo probado muestra que $E(|\prod_{i=1}^n X_i|) = \prod_{i=1}^n E(|X_i|) < \infty$, por lo tanto $E(\prod_{i=1}^n X_i)$ es finita.$\square$

El recíproco del teorema anterior no es válido como lo muestra el siguiente ejemplo

Ejemplo 1.3.10 *Sea* $(\Omega,\mathcal{F},P)$ *un espacio de probabilidad equiprobable, con* $\Omega = \{1,2,3\}$*,* $\mathcal{F} = \wp(\Omega)$*. Definimos* X *y* Y *variables aleatorias sobre* Ω *como:*

$$X(\omega) = \begin{cases} 1 & si\ \omega = 1\\ 0 & si\ \omega = 2\\ -1 & si\ \omega = 3\end{cases} \qquad y \qquad Y(\omega) = \begin{cases} 0 & si\ \omega = 1\ \acute{o}\ \omega = 3\\ 1 & si\ \omega = 2\end{cases}$$

Por lo tanto se satisface $E(XY) = E(X)E(Y) = 0$, puesto que $E(X) = E(XY) = 0$; pero $P(X^{-1}(1) \cap Y^{-1}(1)) = 0$ y $P(X^{-1}(1)) = P(Y^{-1}(1)) = \frac{1}{3}$, es decir que X y Y no son independientes.

Si se agregan más condiciones al recíproco del teorema 1.3.9 se obtiene un resultado distinto del ejemplo anterior, como lo muestra la siguiente proposición.

Proposición 1.3.11 *Sean X, Y variables aleatorias integrables. Entonces $X, Y \perp\!\!\!\perp$ si y sólo si para todo par de funciones borelianas $f(\cdot)$ y $g(\cdot)$ tales que $f(X)$ y $g(Y)$ sean integrables,*

$$E(f(X)g(Y)) = E(f(X))E(g(Y)).$$

Demostración. Si $X, Y \perp\!\!\!\perp$ y $f(\cdot)$, $g(\cdot)$ son borelianas entonces del teorema 1.3.4 $f(X), g(Y) \perp\!\!\!\perp$ y por el teorema 1.3.9 $E(f(X)g(Y)) = E(f(X))E(g(Y))$. Por otra parte si se cumple que $E(f(X)g(Y)) = E(f(X))E(g(Y))$ y se define $f(X) = I_{X^{-1}(A)}$, $g(Y) = I_{Y^{-1}(B)}$, con A y B borelianos, se tiene

$$P(X^{-1}(A) \cap Y^{-1}(B)) = P(X^{-1}(A))P(Y^{-1}(B)),$$

entonces del teorema 1.3.2 se concluye $X, Y \perp\!\!\!\perp$.$\square$

Proposición 1.3.12 *Si $\{X_i\}_{i\in\mathbb{N}}$ son variables aleatorias independientes y si $Var(X_i)$ es finita para todo $i \in \mathbb{N}$, entonces para todo subconjunto finito no vacío $\{i_1, \cdots, i_n\}$ de $\mathbb{N}$ y cada $(\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n$, se cumple*

$$Var(\sum_{j=1}^{n} \alpha_j X_{i_j}) = \sum_{j=1}^{n} \alpha_j^2 Var(X_{i_j}).$$

Demostración. Como las variables aleatorias $\{X_i\}_{i\in\mathbb{N}}$ son independientes, entonces para $X_{i_1}, X_{i_2}, \cdots, X_{i_n}$ se tiene

$$\begin{aligned} Var(\sum_{j=1}^{n} \alpha_j X_{i_j}) &= E[(\sum_{j=1}^{n} \alpha_j X_{i_j} - E(\sum_{j=1}^{n} \alpha_j X_{i_j}))^2] = E[\sum_{j=1}^{n} \alpha_j (X_{i_j} - E(X_{i_j}))]^2 \\ &= E[\sum_{j=1}^{n} \alpha_j^2 (X_{i_j} - E(X_{i_j}))^2 + \sum_{i\neq j=1}^{n} \alpha_k \alpha_j (X_{i_k} - E(X_{i_k}))(X_{i_j} - E(X_{i_j}))] \\ &= \sum_{j=1}^{n} \alpha_j^2 E[(X_{i_j} - E(X_{i_j}))^2] + 0 = \sum_{j=1}^{n} \alpha_j^2 Var(X_{i_j}).\square \end{aligned}$$

Teorema 1.3.13 (Ley débil de los grandes números) *Sean $X_1, X_2, \cdots$ variables aleatorias independientes, tales que cada cada una tiene media $\mu_i < \infty$ y varianzas σ_i^2 uniformemente acotadas por $M < \infty$, entonces para todo $\epsilon > 0$ se tiene*

$$\lim_{n\to\infty} P(|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i)| \geq \epsilon) = 0.$$

Demostración. Como las variables aleatorias $X_1, X_2, \cdots$ son independientes, entonces de la proposicón 1.3.12, se tiene

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i).$$

Luego por el corolario 1.3.6

$$\begin{aligned} P(|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i)| \geq \epsilon) &\leq \frac{1}{\epsilon^2 n^2} Var(\sum_{i=1}^{n} X_i) = \frac{1}{\epsilon^2 n^2}\sum_{i=1}^{n} Var(X_i) \\ &\leq \frac{1}{\epsilon^2 n^2}\sum_{i=1}^{n} M = \frac{1}{\epsilon^2 n} M. \end{aligned}$$

Se toma el límite de n al infinito y se obtiene el resultado.□

En el Capítulo 3 de este trabajo se consideran versiones de la ley fuerte de los grandes números.

1.4. Probabilidad condicionada. Esperanza condicionada

Si $(\Omega, \mathcal{F}, P)$ es un espacio de probabilidad, $B \in \mathcal{F}$ y $P(B) > 0$, la probabilidad condicionada de $A \in \mathcal{F}$ dado B, en símbolos $P(A \mid B)$, se define por

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

Resulta que $P(\cdot \mid B) : \mathcal{F} \to [0, 1]$ es una medida de probabilidad y por tanto $(\Omega, \mathcal{F}, P(\cdot \mid B))$ es un nuevo espacio de probabilidad.

Note que

$$P(A \mid B) = \int_A \frac{I_B}{P(B)} dP,$$

esto es, $P(\cdot \mid B)$ es una medida de probabilidad con densidad (de Radon-Nikodýn) $f(\omega) := \frac{I_B(\omega)}{P(B)}$ con respecto a $P(\cdot)$. Por tanto, si X es una variable aleatoria integrable, entonces se puede considerar la esperanza de X con respecto a la probabilidad $P(\cdot \mid B)$, en símbolos $E(X \mid B)$, como

$$E(X \mid B) := \int_\Omega X(\omega) dP(\omega \mid B) = \int_\Omega X(\omega) \frac{I_B(\omega)}{P(B)} dP(\omega) = \frac{E(XI_B)}{P(B)}.$$

La constante $E(X \mid B)$ se denomina esperanza condicionada de X dado el evento B. Observe que si $X = I_A$; $A \in \mathcal{F}$, entonces $E(X \mid B) \equiv P(A \mid B)$.

Ejemplo 1.4.1 *Si* $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \lambda)$, $B = [\frac{1}{2}, 1]$ *y* $X : \Omega \to \mathbb{R}$ *es la variable aleatoria definida por* $X(\omega) = \omega^2$, *entonces*

$$E(X \mid B) = \frac{E(XI_B)}{P(B)} = 2 \int_{[\frac{1}{2},1]} \omega^2 d\omega = \frac{7}{12}.$$

A continuación se considera la esperanza condicionada con respecto a una descomposición. Con esta visión se definirá el concepto general de esperanza condicionada con respecto a una σ-álgebra $\mathcal{G}$, para luego establer las propiedades elementales de la esperanza condicionada.

Definición 1.4.2 *Si* $(\Omega, \mathcal{F}, P)$ *es un espacio de probabilidad y* $\mathcal{D} = \{D_1, \cdots, D_n\}$ *una descomposición finita de* Ω*; es decir, para* $i, j = 1, \cdots, n$; $D_i \in \mathcal{F}$, $P(D_i) > 0$, $D_i \cap D_j = \emptyset$ *si* $i \neq j$ *y* $\Omega = \uplus_{i=1}^n D_i$. *Entonces para cada* $A \in \mathcal{F}$, *la variable aleatoria simple*

$$P(A \mid \mathcal{D})(\omega) = \sum_{i=1}^n P(A \mid D_i) I_{D_i}(\omega),$$

es llamada la probabilidad condicionada de A *con respecto a la descomposición* $\mathcal{D}$.

La variable aleatoria $P(A \mid \mathcal{D})$ tiene las siguientes propiedades

(a) $P(A \mid \mathcal{D})$ es $\sigma(\mathcal{D})$-medible.

(b) Para todo $D \in \sigma(\mathcal{D})$ se verifica

$$P(A \cap D) = \int_D I_A dP = \int_D P(A \mid \mathcal{D}) dP.$$

La propiedad (a) se sigue de la definición. Para mostrar (b), es suficiente verificar la igualdad para cada D_j

$$\begin{aligned}
\int_{D_j} P(A \mid \mathcal{D})dP &= E(I_{D_j}P(A \mid \mathcal{D})) = E(\sum_{i=1}^{n} P(A \mid D_i)I_{D_i}I_{D_j}) \\
&= E(P(A \mid \ D_j)I_{D_j}) = P(A \mid D_j)P(D_j) \\
&= P(A \cap D_j) = \int_{D_j} I_A dP.
\end{aligned}$$

Sea X una variable aleatoria integrable. La esperanza condicionada de X con respecto a $\mathcal{D}$ es la variable aleatoria

$$E(X \mid \mathcal{D})(\omega) := \sum_{i=1}^{n} E(X \mid D_i)I_{D_i}(\omega),$$

la cual satisface:

(a) $E(X \mid \mathcal{D})$ es $\sigma(\mathcal{D})$-medible.

(b) Para todo $D \in \sigma(\mathcal{D})$ se verifica

$$\int_D XdP = \int_D E(X \mid \mathcal{D})dP.$$

Nuevamente, (a) se sigue de la definición y para verificar (b) es suficiente comprobarlo para cada $D_j \in \mathcal{D}$

$$\begin{aligned}
\int_{D_j} E(X \mid \mathcal{D})dP &= E(I_{D_j}E(X \mid \mathcal{D})) = E(\sum_{i=1}^{n} E(X \mid D_i)I_{D_i}I_{D_j}) \\
&= E(X \mid D_j)P(D_j) = E(XI_{D_j}) = \int_{D_j} XdP.
\end{aligned}$$

Observe que si $X = I_A$; $A \in \mathcal{F}$, entonces $E(X \mid \mathcal{D}) = P(A \mid \mathcal{D})$.

Si $Y : \Omega \to \mathbb{R}$ es una variable aleatoria discreta que toma valores $y_1, y_2, \cdots, y_n$. Entonces $\mathcal{D}_Y := \{\{Y = y_1\}, \{Y = y_2\}, \cdots, \{Y = y_n\}\}$ es una descomposición de Ω. En este caso la variable aleatoria $E(X \mid \mathcal{D}_Y)$ se denota por $E(X \mid Y)$ y se le llama esperanza condicionada de la variable aleatoria X dada la variable aleatoria Y.

Ejemplo 1.4.3 *Sean* $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \lambda)$ *y* $X, Y : \Omega \to \mathbb{R}$ *las variables aleatorias definidas por*

$$X(\omega) = 3\omega^2 + \omega; \quad Y(\omega) = 3I_{[0,\frac{1}{4})}(\omega) + (-2)I_{[\frac{1}{4},1]}(\omega).$$

Aquí $\mathcal{D}_Y = \{\{Y = 3\}, \{Y = -2\}\}$ *y por tanto*

$$E(X \mid Y)(\omega) = E(X \mid Y = 3)I_{[0,\frac{1}{4})}(\omega) + E(X \mid Y = -2)I_{[\frac{1}{4},1]}(\omega).$$

Como

$$E(X \mid Y = 3) = \frac{E(XI_{\{Y=3\}})}{P(\{Y = 3\})} = \frac{\int_0^{\frac{1}{4}}(3\omega^2 + \omega)d\omega}{\frac{1}{4}} = \frac{3}{16}$$

y

$$E(X \mid Y = -2) = \frac{E(XI_{\{Y=-2\}})}{P(\{Y = -2\})} = \frac{\int_{\frac{1}{4}}^{1}(3\omega^2 + \omega)d\omega}{\frac{3}{4}} = \frac{93}{48},$$

entonces

$$E(X \mid Y)(\omega) = \frac{3}{16}I_{[0,\frac{1}{4})}(\omega) + \frac{93}{48}I_{[\frac{1}{4},1]}(\omega).$$

A continuación se da la definición de esperanza condicionada dada una σ-álgebra. Para una revisión de esta definición ver Billingsley [3], Shiryaev [16] o Williams [18].

Definición 1.4.4 *Sea* $(\Omega, \mathcal{F}, P)$ *un espacio de probabilidad,* X *una variable aleatoria integrable y* $\mathcal{G}$ *una sub-*σ*-álgebra de* $\mathcal{F}$*. Entonces existe una variable aleatoria* Y *que es* $\mathcal{G}$*-medible y tal que para todo* $G \in \mathcal{G}$ *(equivalentemente, para todo* G *en un* π*-sistema que contenga a* Ω *y genere a* $\mathcal{G}$*) se verifica*

$$\int_G Y dP = \int_G X dP.$$

Más aún, si Y^* *es otra variable aleatoria con estas propiedades, entonces* $Y = Y^*$ *(c.t.p.). Una variable aleatoria con las propiedades anteriores es llamada una versión de la esperanza condicionada de* X *dada* $\mathcal{G}$ *y se escribe* $Y = E(X \mid \mathcal{G})$*.*

Si $A \in \mathcal{F}$, la probabilidad condicionada de A dada la σ-álgebra $\mathcal{G}$ se denota por $P(A \mid \mathcal{G})$ y se define por $P(A \mid \mathcal{G}) := E(I_A \mid \mathcal{G})$.

Si $(E, \mathcal{E})$ es un espacio medible $Y : \Omega \to E$ es una función $\mathcal{F} - \mathcal{E}$ medible y $\mathcal{G}_Y = \sigma(Y)$, entonces la esperanza condicionada de la variable aleatoria X dada Y se denota por $E(X \mid Y)$ y se define por

$$E(X \mid Y) := E(X \mid \mathcal{G}_Y).$$

El siguiente par de resultados serán enunciados por su utilidad y sus demostraciones se encuentran en el Capítulo II Sección 4 de Shiryaev [16].

Teorema 1.4.5 *Dada X una variable aleatoria. La variable aleatoria Y es una función $\mathcal{F}_X$-medible si y sólo si existe una función boreliana f tal que $Y(\omega) = f(X(\omega))$, para cada $\omega \in \Omega$.*

Teorema 1.4.6 *Sea $\mathcal{D} = \{D_1, D_2, \cdots, \}$ una descomposición finita ó contable de Ω. Si $X = X(\omega)$ es una variable aleatoria $\sigma(\mathcal{D})$-medible, entonces X se puede representar en la forma*

$$X(\omega) = \sum_{i \geq 1} x_i I_{D_i}(\omega),$$

donde $x_i \in \mathbb{R}$. $i = 1, 2, \cdots$; esto es, $X(\omega)$ es constante sobre cada elemento D_i de $\mathcal{D}$.

Ejemplo 1.4.7 *Sea $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \lambda)$. Para calcular $E(Y \mid \sigma(X))(\omega)$, donde $Y(\omega) = 2\omega^2$, $X(\omega) = 2I_{[0,\frac{1}{2})}(\omega) + \omega I_{[\frac{1}{2},1]}(\omega)$ se procede así.*

Puesto que $E(Y \mid \sigma(X))$ es $\sigma(X)$-medible, por el teorema 1.4.5 existe una función boreliana f tal que $E(Y \mid \sigma(X)) = f(X)$, esto es,

$$E(Y \mid \sigma(X))(\omega) = f(2)I_{[0,\frac{1}{2})}(\omega) + f(\omega)I_{[\frac{1}{2},1]}(\omega).$$

Para todo $G \in \sigma(X)$, $\int_G E(Y \mid \sigma(X))dP = \int_G Y dP$, tomando $G = X^{-1}(\{2\}) = [0, \frac{1}{2})$ obtenemos

$$\frac{1}{12} = \int_{[0,\frac{1}{2})} 2\omega^2 d\omega = \int_{[0,\frac{1}{2})} Y dP = \int_{[0,\frac{1}{2})} E(Y \mid \sigma(X))dP = \int_{[0,\frac{1}{2})} f(2)d\omega = \frac{1}{2}f(2).$$

De aquí que $f(2) = \frac{1}{6}$. De otra parte, para cualquier conjunto de Borel $B \subseteq [\frac{1}{2}, 1]$ tenemos $X^{-1}(B) = B$, y por lo tanto

$$\int_B 2\omega^2 d\omega = \int_B Y dP = \int_B E(Y \mid \sigma(X))dP = \int_B f(\omega)d\omega.$$

De aquí que $f(\omega) = 2\omega^2$ (c.t.p.) para $\omega \in [\frac{1}{2}, 1]$. En resumen

$$E(Y \mid \sigma(X)) = \frac{1}{6}I_{[0,\frac{1}{2})}(\omega) + 2\omega^2 I_{[\frac{1}{2},1]}(\omega).$$

La siguiente proposición contiene algunas propiedades importantes de la esperanza condicionada, de las cuales se hará uso posteriormente. Para las demostraciones de todas ellas se puede consultar el Capítulo 9 de Williams [18] y la Sección 7 del Capítulo II de Shiryaev [16].

Proposición 1.4.8 *Dada $\mathcal{G} \subseteq \mathcal{F}$ una sub-σ-álgebra, la esperanza condicional tiene las siguientes propiedades:*

(E1) *Si c es constante y $X = c$ (c.t.p.), entonces $E(X \mid \mathcal{G}) = c$ (c.t.p.).*

(E2) *(Orden) Si $X \leq Y$ (c.t.p.), entonces $E(X \mid \mathcal{G}) \leq E(Y \mid \mathcal{G})$ (c.t.p.).*

(E3) *$|E(X \mid \mathcal{G})| \leq E(|X| \mid \mathcal{G})$ (c.t.p.).*

(E4) *(Linealidad) Si a,b son constantes y $aE(X) + bE(Y)$ existe, entonces*

$$E(aX + bY \mid \mathcal{G}) = aE(X \mid \mathcal{G}) + bE(Y \mid \mathcal{G}) \quad \text{(c.t.p.)}.$$

(E5) *Si $\mathcal{F}_* = \{\emptyset, \Omega\}$ es la σ-álgebra trivial, entonces $E(X \mid \mathcal{F}_*) = E(X)$ (c.t.p.).*

(E6) *Si X es $\mathcal{F}$-medible, entonces $E(X \mid \mathcal{F}) = X$ (c.t.p.).*

(E7) *(Propiedad de la torre) Si $\mathcal{G}_1 \subseteq \mathcal{G}_2$, entonces $E(E(X \mid \mathcal{G}_2) \mid \mathcal{G}_1) = E(X \mid \mathcal{G}_1)$ (c.t.p.). Si $\mathcal{G}_1 \supseteq \mathcal{G}_2$ entonces $E(E(X \mid \mathcal{G}_2) \mid \mathcal{G}_1) = E(X \mid \mathcal{G}_2)$ (c.t.p.).*

(E8) *(Doble esperanza) $E(E(X \mid \mathcal{G})) = E(X)$.*

(E9) *(Papel de la independencia) Dada una variable aleatoria X integrable e independiente de la σ-álgebra $\mathcal{G}$, es decir independiente de I_B con $B \in \mathcal{G}$, entonces $E(X \mid \mathcal{G}) = E(X)$ (c.t.p.).*

(E10) *(Convergencia dominada) Dadas $Y, X_1, X_2, \cdots$ variables aleatorias tales que $|X_n| \leq Y$, $E(Y) < \infty$ y $X_n \to X$ (c.t.p.) entonces $E(X_n \mid \mathcal{G}) \to E(X \mid \mathcal{G})$ (c.t.p.) y $E(|X_n - X| \mid \mathcal{G}) \to 0$ (c.t.p.).*

(E11) *(Convergencia monótona) Dadas $Y, X_1, X_2, \cdots$ variables aleatorias*

(a) *Si $X_n \geq Y$, $E(Y) > -\infty$ y $X_n \uparrow X$ (c.t.p.), entonces*

$$E(X_n \mid \mathcal{G}) \uparrow E(X \mid \mathcal{G}) \quad \text{(c.t.p.)}.$$

(b) *Si $X_n \leq Y$, $E(Y) < \infty$ y $X_n \downarrow X$ (c.t.p.), entonces*

$$E(X_n \mid \mathcal{G}) \downarrow E(X \mid \mathcal{G}) \quad \text{(c.t.p.)}.$$

(E12) *(Lema de Fatou) Dadas $Y, X_1, X_2, \cdots$ variables aleatorias*

(a) Si $X_n \geq Y$, $E(Y) > -\infty$, *entonces*

$$E(\liminf X_n \mid \mathcal{G}) \leq \liminf E(X_n \mid \mathcal{G}) \quad (c.t.p.).$$

(b) Si $X_n \leq Y$, $E(Y) < \infty$, *entonces*

$$\limsup E(X_n \mid \mathcal{G}) \leq E(\limsup X_n \mid \mathcal{G}) \quad (c.t.p.).$$

(E13) (Serie positiva) Si $X_n \geq 0$, *entonces* $E(\sum_{n=1}^{\infty} X_n \mid \mathcal{G}) = \sum_{n=1}^{\infty} E(X_n \mid \mathcal{G})$ *(c.t.p.). En particular* $E(\sum_{n=1}^{\infty} X_n) = \sum_{n=1}^{\infty} E(X_n)$.

(E14) (Extrayendo lo medible) Sea Y *una variable aleatoria* $\mathcal{G}$*-medible,* $E|X| < \infty$ *y* $E|XY| < \infty$, *entonces* $E(XY \mid \mathcal{G}) = Y E(X \mid \mathcal{G})$ *(c.t.p.).*

(E15) (Desigualdad de Jensen) Sea $g : \mathbb{R} \to \mathbb{R}$ *una función convexa tal que* $E|g(X)| < \infty$, *entonces* $E(g(X) \mid \mathcal{G}) \geq g(E(X \mid \mathcal{G}))$ *(c.t.p.).*

Ejemplo 1.4.9 *(Adaptado de Brzeźniak y Zastawniak [4])*
Sean $(\Omega, \mathcal{F}, P) = ((0,1), \mathcal{B}((0,1)), \lambda)$ *y las variables aleatorias* $X(\omega) = 2\omega^2$ *y* $Y(\omega) = 2\omega I_{(0,\frac{1}{2})}(\omega) + (2\omega - 1) I_{[\frac{1}{2},1)}(\omega)$. *Para determinar* $E(X \mid \sigma(Y))(\omega)$*; primero se observa que*

$$X = \frac{Y^2}{2} I_{(0,\frac{1}{2})} + \frac{(Y+1)^2}{2} I_{[\frac{1}{2},1)}.$$

Entonces de la propiedad (E14) de la proposición 1.4.8 se tiene

$$E(X \mid \sigma(Y)) = \frac{Y^2}{2} E(I_{(0,\frac{1}{2})} \mid \sigma(Y)) + \frac{(Y+1)^2}{2} E(I_{[\frac{1}{2},1)} \mid \sigma(Y)).$$

Ahora, la σ*-álgebra inducida por* Y *consiste de conjuntos de la forma* $B \cup (B + \frac{1}{2})$ *para algún conjunto boreliano* $B \subseteq (0, \frac{1}{2})$. *Puesto que*

$$P(B) = \int_{B \cup (B+\frac{1}{2})} I_{(0,\frac{1}{2})} dP = \int_{B \cup (B+\frac{1}{2})} E(I_{(0,\frac{1}{2})} \mid \sigma(Y)) dP$$

y

$$P(B) = P(B + \frac{1}{2}) = \int_{B \cup (B+\frac{1}{2})} I_{[\frac{1}{2},1)} dP = \int_{B \cup (B+\frac{1}{2})} E(I_{[\frac{1}{2},1)} \mid \sigma(Y)) dP,$$

entonces $E(I_{(0,\frac{1}{2})} \mid \sigma(Y)) = E(I_{[\frac{1}{2},1)} \mid \sigma(Y))$. *Como*

$$1 = E(1 \mid Y) = E(I_{(0,\frac{1}{2})} + I_{[\frac{1}{2},1)} \mid Y) = E(I_{(0,\frac{1}{2})} \mid Y) + E(I_{[\frac{1}{2},1)} \mid Y) = 2E(I_{[\frac{1}{2},1)} \mid Y),$$

entonces $E(I_{[\frac{1}{2},1)} \mid Y) = \frac{1}{2}$. *En consecuencia,* $E(X \mid Y) = \frac{1}{4}(Y^2 + (Y+1)^2)$, *esto es,*

$$E(X \mid Y)(\omega) = \begin{cases} \omega^2 + (\omega + \frac{1}{2})^2 & si\ \omega \in (0, \frac{1}{2}) \\ (\omega - \frac{1}{2})^2 + \omega^2 & si\ \omega \in [\frac{1}{2}, 1). \end{cases}$$

De acuerdo a lo dicho al inicio de la Sección, se sabe que $E(X|Y)$ es $\mathcal{G}_Y$-medible, entonces del teorema 1.4.5 existe una función boreliana $m = m(y)$ de $\mathbb{R}$ en $\mathbb{R}$ tal que para cada $\omega \in \Omega$ se tiene $m(Y(\omega)) = E(X|Y)(\omega)$, donde $m(y) := E(X|Y = y)$, la cual es llamada la esperanza condicional de X con respecto al evento $\{Y = y\}$. De otra parte, por el teorema de cambio de variable de la Sección 6 del Capítulo II de Shiryaev [16], se tiene

$$\int_{\{\omega:Y\in B\}} XdP = \int_{\{\omega:Y\in B\}} m(Y)dP = \int_B m(y)P_Y(dy),$$

donde $B \in \mathcal{B}(\mathbb{R})$, P_Y es la distribución de probabilidad de Y y $\{\omega : Y(\omega) \in B\} \in \mathcal{G}_Y$; todo lo cual lleva a la siguiente definición.

Definición 1.4.10 *Sean X y Y variables aleatorias tal que X es integrable. La esperanza condicional de la variable aleatoria X con respecto a $Y = y$ es cualquier función $B \in \mathcal{B}(\mathbb{R})$-medible $m = m(y)$, para la cual*

$$\int_{\{\omega:Y\in B\}} XdP = \int_B m(y)P_Y(dy), \quad B \in \mathcal{B}(\overline{\mathbb{R}}).$$

Que exista tal función sigue del teorema de Radon-Nikodým si se observa que

$$Q(B) = \int_{\{\omega:Y\in B\}} XdP,$$

es una medida signada absolutamente continua con respecto a la medida P_Y (ver Sección 6.7 Capítulo II de shiryaev [16]).

Definición 1.4.11 *La probabilidad condicional de el evento $A \in \mathcal{F}$ dada la condición $Y = y$, en símbolos $P(A|Y = y)$, es $E(I_A|Y = y)$.*

Ejemplo 1.4.12 *Si Y es una variable aleatoria discreta con $P(Y = y_k) > 0$, $\sum_{k=1}^{\infty} P(Y = y_k) = 1$ y X es otra variable aleatoria, tal que $E(X)$ existe. Entonces para $k \geq 1$*

$$P(A|Y = y_k) = \frac{P(A \cap \{Y = y_k\})}{P(Y = y_k)} \quad y \quad E(X|Y = y_k) = \frac{1}{P(Y = y_k)} \int_{\{\omega:Y=y_k\}} XdP.$$

Si $y \notin \{y_1, y_2, \cdots\}$, la probabilidad condicionada $P(A|Y = y)$ y la esperanza condicionada $E(X|Y = y)$, pueden ser definidas de cualquier manera; en particular como cero.

Si $(\Omega, \mathcal{F}, P)$ es un espacio de probabilidad con $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$ se tiene que:

(a) $0 \leq P(A|\mathcal{G}) \leq 1$ (c.t.p.), para todo $A \subseteq \mathcal{F}$.

(b) $P(\cup_{n=1}^{\infty} A_n|\mathcal{G}) = \sum_{n=1}^{\infty} P(A_n|\mathcal{G})$ (c.t.p.), para $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de eventos disjuntos de $\mathcal{F}$.

(c) $P(B - A|\mathcal{G}) = P(B|\mathcal{G}) - P(A|\mathcal{G})$ (c.t.p.), para $A, B \in \mathcal{F}$ y $A \subseteq B$.

(d) Si $\{A_n\}_{n\in\mathbb{N}}$ es una sucesión de eventos en $\mathcal{F}$ y $A_n \uparrow A = \cup_{n=1}^{\infty} A_n$, entonces $P(A_n|\mathcal{G}) \uparrow P(A|\mathcal{G})$ (c.t.p.).

(e) Si $\{A_n\}_{n\in\mathbb{N}}$ es una sucesión de eventos en $\mathcal{F}$ y $A_n \downarrow A = \cap_{n=1}^{\infty} A_n$, entonces $P(A_n|\mathcal{G}) \downarrow P(A|\mathcal{G})$ (c.t.p.).

Se podría considerar de las anteriores propiedades $P(\cdot|\mathcal{G})(\omega) : \mathcal{F} \to [0,1]$, como una medida de probabildad para ω fijo y pensar que existe $\mathcal{N} \in \mathcal{F}$ con $P(\mathcal{N}) = 0$, donde para $\omega \in \mathcal{N}^c$, $P(\cdot|\mathcal{G})(\omega)$ es una medida de probabilidad; sin embargo este no es el caso en general, como lo muestran los siguientes ejemplos.

Ejemplo 1.4.13 *Si $\mathcal{D} = \{D_1, D_2, \cdots, D_n\}$ es una descomposición finita de Ω, $\mathcal{G} = \sigma(D_1, \cdots, D_n)$; $P(D_1) = 0$ y $P(D_i) > 0$, para $i = 2, 3, \cdots, n$. Una versión de $P(A|\mathcal{G})$ es:*

$$P(A|\mathcal{G})(\omega) = \begin{cases} 0 & si\ \omega \in D_1, \\ \dfrac{P(A \cap D_i)}{P(D_i)} & si\ \omega \in D_i,\ i = 2, \cdots, n. \end{cases}$$

En este caso la función $A \to P(A|\mathcal{G})(\omega)$ es una probabilidad, si $\omega \in \Omega - D_1$.
Si se define

$$P(A|\mathcal{G})(\omega) = \begin{cases} P(A) & si\ \omega \in D_1, \\ \dfrac{P(A \cap D_i)}{P(D_i)} & si\ \omega \in D_i,\ i = 2, \cdots, n, \end{cases}$$

entonces para cada $\omega \in \Omega$, $A \to P(A|\mathcal{G})(\omega)$ es una probabilidad.

Ejemplo 1.4.14 *Dado el espacio de probabilidad $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1], \lambda))$; si $\mathcal{G} = \mathcal{F}$, para cada $A \in \mathcal{F}$*

$$f(A) = \begin{cases} \sup A & si\ A \neq \emptyset \\ 0 & si\ A = \emptyset. \end{cases}$$

Y se define $Q : \Omega \times \mathcal{F} \to [0,1]$ *como*

$$Q(\omega, A) := I_A(\omega) + I_{\{f(A)\}}(\omega).$$

Entonces

(a) *Para cada* $A \in \mathcal{F}$*,* $Q(\omega, A)$ *es una versión de* $P(A|\mathcal{G})$*.*

(b) *No existe* $\mathcal{N} \in \mathcal{F}$*, con* $P(\mathcal{N}) = 0$*, tal que* $A \to Q(\omega, A)$ *es una probablidad para cada* $\omega \in \mathcal{N}^c$*.*

En efecto; (a) se cumple, pues si $G \in \mathcal{G}$*, entonces para cada* $A \in \mathcal{F}$

$$\int_G P(A|\mathcal{G})dP = \int_G I_A dP + 0 = \int_G I_A dP + \int_G I_{\{\sup A\}} dP = \int_G (I_A + I_{\{\sup A\}})dP.$$

Además si $x_0 \in (0,1]$*, entonces* $A \to Q(x_0, A)$ *no es una medida de probabilidad. Para mostrarlo; sean* $0 < x_n \nearrow x_0$*, donde* $x_i \neq x_j$*, para* $i \neq j$ *y* $A_n = \{x_n\}$*. En este caso* $\sup \cup_{n\in\mathbb{N}} A_n = x_0$ *y*

(i) $Q(x_0, \cup_{n\in\mathbb{N}} A_n) = I_{\cup_{n\in\mathbb{N}} A_n}(x_0) + I_{\{\sup \cup_{n\in\mathbb{N}} A_n\}}(x_0) = 0 + 1 = 1.$

(ii) $Q(x_0, A_n) = I_{A_n}(x_0) + I_{\{\sup A_n\}}(x_0) = 0 + 0 = 0.$

Definición 1.4.15 *Una función* $P(\omega; B)$*,* $\omega \in \Omega$*,* $B \in \mathcal{F}$*, es una probabilidad condicional regular con respecto a la* σ*-álgebra* $\mathcal{G}$*, si:*

(a) *Para cada* $\omega \in \Omega$ *fijo,* $P(\omega; \cdot) : \mathcal{F} \to [0,1]$ *es una medida de probabilidad.*

(b) *Para cada* $B \in \mathcal{F}$ *fijo, la función* $P(\omega; B)$ *es una versión de la probabilidad condicional* $P(B|\mathcal{G})(\omega)$*, es decir* $P(\omega; B) = P(B|\mathcal{G})(\omega)$ *(c.t.p.).*

Ejemplo 1.4.16 *Si* $\mathcal{G}_1$*,* $\mathcal{G}_2$ *son dos sub-*σ*-álgebras independientes en* $\mathcal{F}$ *entonces la función* $P(\omega; B) := P(B)$*, con* $B \in \mathcal{G}_1$*, es una medida de probabilidad para cada* $\omega \in \Omega$ *y es una versión de la probabilidad condicional* $P(B|\mathcal{G}_2)(\omega)$ *ya que de la independencia*

$$P(B|\mathcal{G}_2) = E(I_B|\mathcal{G}_2) = E(I_B) = P(B) \quad (c.t.p.).$$

Por lo tanto $P(\omega; B)$ *es una probabilidad condicional regular con respecto a la sub-*σ*-álgebra* $\mathcal{G}_2$*.*

Teorema 1.4.17 *Dada $P(\omega; B)$ una probabilidad condicional regular con respecto a $\mathcal{G}$ y X una variable aleatoria integrable. Entonces*

$$E(X|\mathcal{G})(\omega) = \int_{\Omega} X(s)P(\omega; ds) \quad (c.t.p.).$$

Demostración. Para $X = I_B$, $B \in \mathcal{F}$ por el parte (b) de la definición, se tiene

$$E(X|\mathcal{G})(\omega) = P(B|\mathcal{G})(\omega) = P(\omega; B) = \int_{\Omega} I_B(s)P(\omega, ds) \quad (c.t.p.).$$

Para X una variable aleatoria simple es análogo, sólo hay que tener en cuenta la linealidad; para $X \geq 0$ existe una sucesión de variables simples $\{X_n\}_{n\in\mathbb{N}}$ tales que $X_n \uparrow X$ y por lo tanto de la propiedad *(E11)* de la propiedad 1.4.8 $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ (c.t.p.) por tanto del teorema de convergencia monotona para integrales y para la medida $P(\omega, \cdot)$, se tiene

$$E(X|\mathcal{G})(\omega) = \lim_{n\to\infty} E(X_n|\mathcal{G})(\omega) = \lim_{n\to\infty} \int_{\Omega} X_n(s)P(\omega, ds) = \int_{\Omega} X(s)P(\omega, ds).$$

Para X arbitraria; como $X = X^+ - X^-$, se utiliza lo anterior.$\square$

Corolario 1.4.18 *Sean X, Y variables aleatorias, $\mathcal{G}_Y$ la σ-álgebra generada por Y donde (X, Y) tiene función de distribución con densidad $f_{XY}(X, Y)$. Si $E|g(X)| < \infty$, entonces*

$$E(g(X)|Y = y) = \int_{\infty}^{-\infty} g(x)f_{X|Y}(x|y)dx,$$

donde $f_{X|Y}(x|y)$ es la densidad de la distribución condicional.

Demostración. Dado que $E|g(X)| < \infty$. Entonces del teorema de cambio de variable de la Sección 6 del Capítulo II de Shiryaev [16] y de el teorema anterior

$$\begin{aligned} E(g(X)|Y = y) &= \int_{\Omega} g(X(s))P(\omega, ds) = \int_{\Omega} g(X(s))dP(X(s)|Y = y)(\omega) \\ &= \int_{-\infty}^{\infty} g(x)f_{X|Y}(x|y)dx.\square \end{aligned}$$

Definición 1.4.19 *Dados $(\Omega, \mathcal{F})$ y $(E, \mathcal{E})$ espacios medibles, se dice que la aplicación $X = X(\omega)$ definida sobre Ω y que toma valores en E, es $\mathcal{F}/\mathcal{E}$-medible o es un elemento aleatorio; si $\{\omega : X(\omega) \in B\} \in \mathcal{F}$, para todo $B \in \mathcal{E}$.*

Definición 1.4.20 *Sean $(\Omega, \mathcal{F})$ y $(E, \mathcal{E})$ espacios medibles, $X = X(\omega)$ un elemento aleatorio con valores en E y $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Una función $Q(\omega, B)$, con $\omega \in \Omega$ y $B \in \mathcal{G}$ es una distribución condicional regular de X dada $\mathcal{G}$, si*

(a) $Q(\omega; B)$ *es una medida de probabilidad para cada* $\omega \in \Omega$.

(b) *Para todo* $B \in \mathcal{E}$, $Q(\omega; B) = P(X \in B|\mathcal{G})(\omega)$ *(c.t.p.)*.

Definición 1.4.21 *Dada* $X = X(\omega)$ *una variable aleatoria. Una función* $F = F(\omega, x)$, $\omega \in \Omega$, $x \in \mathbb{R}$ $B \in \mathcal{G}$ *es una función de distribución condicional regular con respecto a* $\mathcal{G}$, *si*

(a) *Para cada* $\omega \in \Omega$, $F(\omega; x)$ *es una función de distribución sobre* $\mathbb{R}$.

(b) *Para cada* $x \in \mathbb{R}$, $F(\omega; x) = P(X \leq x|\mathcal{G})(\omega)$ *(c.t.p.)*.

Ejemplo 1.4.22 *Para* $(\Omega, \mathcal{F}, P)$ *un espacio de probabilidad,* $\mathcal{G}$ *una sub-*σ*-álgebra de* $\mathcal{F}$ *y* X *una variable aleatoria.*

(1) *Si* X *es* $\mathcal{G}$*-medible. Entonces*

$$P(X \leq x|\mathcal{G}) = E(I_{\{X \leq x\}}|\mathcal{G}) = I_{\{X \leq x\}} \quad \text{(c.t.p.)},$$

de lo cual

$$P(X \leq x|\mathcal{G})(\omega) = \begin{cases} 1 & si\ X(\omega) \leq x, \\ 0 & si\ X(\omega) > x. \end{cases}$$

Por lo tanto $F(\omega; x) := I_{\{X \leq x\}}(\omega)$ *es una función de distribución regular, ya que cumple las condiciones dadas en la definición anterior.*

(2) *Si* X *es independiente de* $\mathcal{G}$. *Entonces*

$$P(X \leq x|\mathcal{G}) = E(I_{\{X \leq x\}}|\mathcal{G}) = E(I_{\{X \leq x\}}) = P(X \leq x) = F_X(x) \quad \text{(c.t.p.)},$$

Por lo tanto $F(\omega; x) := F_X(x)$ *es como en el anterior caso una función de distribución regular.*

El siguiente resultado, es de existencia de distribuciones condicionales regulares y su demostración se encuentra en la Sección 7 del Capítulo II de Shiryaev [16].

Teorema 1.4.23 *Si* $X = X(\omega)$ *es un elemento aleatorio con valores en un espacio metrico, completo y separable* $(E, \mathcal{G})$. *Entonces existe una distribución condicional regular de* X *con respecto a* $\mathcal{G}$. *En particular tal distribución existe para el espacio* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Capítulo 2

Propiedades de invarianza de la independencia condicionada.

En este Capítulo se establecen algunas condiciones necesarias y/o suficientes para que la relación de independencia condicionada se preserve cuando se hacen cambios en uno de los tres objetos que intervienen en ésta, ya sea en la familia de clases de eventos, la sub-σ-álgebra condicionante o la medida de probabilidad. Estos resultados, que constituyen el mayor aporte de este trabajo, se pueden ver como generalizaciones de los establecidos por Van Putten y Van Schuppen [17] para dos sub-σ-álgebras y de los conocidos para clases de eventos independientes (ver Secciones 5.1 y 5.2 de Bauer [2]). Los resultados más importantes aparecen en Marmolejo y Muñoz [11]. El plan de este capítulo es como sigue. En la Sección 2.1 se revisan las definiciones de independencia condicionada dado un evento e independencia condicionada dada una σ-álgebra, en la Sección 2.2 se consideran los cambios en la familia $\{\mathcal{E}_i\}_{i\in I}$, en la Sección 2.3 los cambios en la σ-álgebra condicionante $\mathcal{G}$ y en la Sección 2.4 los cambios en la medida de probabilidad P. Como aplicación del contenido de las secciones 2.1 a 2.4, en la Sección 2.5 se establecen algunos resultados sobre la independencia condicionada de una familia de variables aleatorias. Finalmente, en la Sección 2.6 se hace una breve presentación de las familias de σ-álgebras recíprocas y markovianas.

2.1. Independencia condicionada.

Dado el espacio de probabilidad $(\Omega, \mathcal{F}, P)$; si $B \in \mathcal{F}$ es un evento tal que $P(B) > 0$ y $P(A|B) = \frac{P(A\cap B)}{P(B)}$ indica la probabilidad condicionada del evento A dado B, entonces

$(\Omega, \mathcal{F}, P(\cdot \mid B))$ es también un espacio de probabilidad. La independencia en este último espacio se denomina independencia condicionada por el evento B o independencia local sobre B. Específicamente se tiene la siguiente definición.

Definición 2.1.1 *Sea $B \in \mathcal{F}$ un evento tal que $P(B) > 0$. Se dice que los eventos $A_1, A_2, \cdots, A_n \in \mathcal{F}$ son independientes condicionalmente dado el evento B, en símbolos: $A_1, A_2, \cdots, A_n \perp\!\!\!\perp B$, si para cualquier subconjunto $\{i_1, \cdots, i_k\}$ de $\{1, 2, \cdots, n\}$; $k \geq 2$, se cumple*

$$P(\cap_{j=1}^{k} A_{i_j} \mid B) = \prod_{j=1}^{k} P(A_{i_j} \mid B).$$

El siguiente ejemplo ilustra que la independencia condicionada no implica independencia y viceversa.

Ejemplo 2.1.2 *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad tal que $\Omega = \{1, 2, 3, 4\}$, $\mathcal{F} = \wp(\Omega)$ y $P(\{i\}) = \frac{1}{4}$; $i = 1, 2, 3, 4$.*

(a) Los eventos $A_1 = \{1, 2\}$ y $A_2 = \{1, 3\}$ son independientes, pero no son condicionalmente independientes dado $B = \{2, 3\}$, pues $P(A_1 \cap A_2) = \frac{1}{4} = P(A_1)P(A_2)$ y $P(A_1 \cap A_2 | B) = 0 \neq \frac{1}{4} = P(A_1|B)P(A_2|B)$.

(b) Los eventos $A_1 = \{1, 2, 3\}$ y $A_2 = \{2, 3, 4\}$ son condicionalmente independientes dado $B = \{2, 3\}$, pero no son independientes. En efecto $P(A_1 \cap A_2 | B) = 1 = P(A_1|B)P(A_2|B)$ y $P(A_1 \cap A_2) = \frac{1}{2} \neq (\frac{3}{4})(\frac{3}{4}) = P(A_1)P(A_2)$.

La siguiente definición extiende el concepto de independencia dado un evento a la independencia condicionada por una σ-álgebra.

Definición 2.1.3 *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$ y $\{\mathcal{E}_i\}_{i \in I}$ una familia de clases de eventos $\mathcal{E}_i \subseteq \mathcal{F}$. Se dice que $\{\mathcal{E}_i\}_{i \in I}$ es independiente dada $\mathcal{G}$, en símbolos $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$, si para cada subconjunto finito $\{i_1, i_2, \cdots, i_k\}$ de I; $k \geq 2$, y cada escogencia $E_{i_1} \in \mathcal{E}_{i_1}, E_{i_2} \in \mathcal{E}_{i_2}, \cdots, E_{i_k} \in \mathcal{E}_{i_k}$, se tiene*

$$P(\cap_{j=1}^{k} E_{i_j} \mid \mathcal{G}) = \prod_{j=1}^{k} P(E_{i_j} \mid \mathcal{G}) \quad (c.t.p.). \tag{2.1}$$

La $\mathcal{G}$-independencia de una sucesión finita de clases de eventos $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_n$; $n \geq 2$, también se denota por $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_n \perp\!\!\!\perp \mathcal{G}$.

Observación 2.1.4 *Las definiciones de independencia y de $\mathcal{G}$-independencia coinciden cuando $P(G) \in \{0,1\}$ para todo $G \in \mathcal{G}$ (en particular cuando $\mathcal{G} = \{\emptyset, \Omega\}$); pues para cada evento $G \in \mathcal{G}$, se cumple $P(A \cap G) = P(A)P(G)$.*

En el siguiente ejemplo se muestra la importancia de verificar (2.1) sobre cada subconjunto $\{i_1, i_2, \cdots, i_k\}$ de $\{1, 2, \cdots, n\}$; $k \geq 2$.

Ejemplo 2.1.5 *Sea $(\Omega, \mathcal{F}, P)$ el espacio de probabilidad del ejemplo 2.1.2. Si se toman $E_1 = \{1,2\}$, $E_2 = \{2,3\}$, $E_3 = \{1,3\}$, $\mathcal{F}_1 = \sigma(E_1)$, $\mathcal{F}_2 = \sigma(E_2)$, $\mathcal{F}_3 = \sigma(E_3)$ y $\mathcal{G} = \{\emptyset, \Omega\}$. Entonces se tiene que*

$$P(E_i \cap E_j \mid \mathcal{G}) = P(E_i \cap E_j) = \frac{1}{4} = P(E_i)P(E_j) = P(E_i \mid \mathcal{G})P(E_j \mid \mathcal{G}),$$

$$P(E_i \cap E_j^c \mid \mathcal{G}) = P(E_i \cap E_j^c) = \frac{1}{4} = P(E_i)P(E_j^c) = P(E_i \mid \mathcal{G})P(E_j^c \mid \mathcal{G}),$$

es decir, $\mathcal{F}_i, \mathcal{F}_j \perp\!\!\!\perp \mathcal{G}$ para $i \neq j$ con $1 < i, j < 3$. Sin embargo $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ no son independientes dada la σ-álgebra $\mathcal{G}$; pues $P(E_1 \cap E_2 \cap E_3) = 0 \neq \frac{1}{8} = P(E_1)P(E_2)P(E_3)$, es decir

$$P(E_1 \cap E_2 \cap E_3 \mid \mathcal{G}) \neq P(E_1 \mid \mathcal{G})P(E_2 \mid \mathcal{G})P(E_3 \mid \mathcal{G}).$$

Proposición 2.1.6 *Sea $\{\mathcal{E}_i\}_{i \in I}$ una familia de clases de eventos.*

(a) *$\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$ si y sólo si para todo $J \subseteq I$, $J \neq \emptyset$; se tiene $\{\mathcal{E}_j\}_{j \in J} \perp\!\!\!\perp \mathcal{G}$.*

(b) *Si $\mathcal{E}_i \subseteq \mathcal{G}$ para cada $i \in I$, entonces $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$. En particular, siempre se cumple $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \sigma(\cup_{i \in I} \mathcal{E}_i)$ y $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{F}$.*

(c) *Si $\mathcal{E}_1 \subseteq \mathcal{G}$ o $\mathcal{E}_2 \subseteq \mathcal{G}$ entonces $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp \mathcal{G}$. En especial, siempre se cumplen las relaciones $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp \sigma(\mathcal{E}_1)$, $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp \sigma(\mathcal{E}_2)$ y $\mathcal{E}_1, \mathcal{E}_2 \perp\!\!\!\perp \sigma(\mathcal{E}_1, \mathcal{E}_2)$.*

(d) *Si $\mathcal{F}$, $\mathcal{G}$ son independientes y $\mathcal{E}_i \subseteq \mathcal{F}$, para cada $i \in I$, entonces $\{\mathcal{E}_i\}_{i \in I}$ es independiente si y sólo si $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$.*

(e) *Si $\mathcal{E}_i \sqcap \mathcal{G} := \{E_i \cap G : E_i \in \mathcal{E}_i, G \in \mathcal{G}\}$, entonces $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$ si y sólo si $\{\mathcal{E}_i \sqcap \mathcal{G}\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$.*

(f) *Si $\{\mathcal{E}_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$, entonces para cada subconjunto finito no vacío $\{i_1, i_2, \cdots, i_k\}$ de I, cada escogencia $E_{i_1} \in \mathcal{E}_{i_1}, \cdots, E_{i_k} \in \mathcal{E}_{i_k}$ y cada $(\alpha_1, \alpha_2, \cdots, \alpha_k) \in \mathbb{R}^k$ se cumple*

$$Var(\sum_{j=1}^{k} \alpha_j I_{E_{i_j}} \mid \mathcal{G}) = \sum_{j=1}^{k} \alpha_j^2 Var(I_{E_{i_j}} \mid \mathcal{G}),$$

donde $Var(I_E \mid \mathcal{G}) = E[(I_E - E(I_E))^2 \mid \mathcal{G}]$.

Demostración. (a), (b) y (c) se obtienen fácilmente de la definición de independencia condicinada y de la propiedad *(E14)* de la proposición 1.4.8.

(d) Si $\mathcal{F}$ y $\mathcal{G}$ son independientes, para todo $A \in \mathcal{F}$ se cumple $P(A \mid \mathcal{G}) = P(A)$, por tanto para $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ se tiene

$$P(\cap_{j=1}^k E_{i_j}) = P(\cap_{j=1}^k E_{i_j} \mid \mathcal{G}) = \prod_{j=1}^k P(E_{i_j} \mid \mathcal{G}) = \prod_{j=1}^k P(E_{i_j}) \quad (c.t.p.),$$

para cada $E_{i_1} \in \mathcal{E}_{i_1} \subseteq \mathcal{F},\ i_j \in \{i_1, i_2, \cdots, i_k\} \subseteq I$; es decir, $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp$.

(e) por la propiedad *(E14)* de la proposición 1.4.8, para $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$, se tiene

$$P(\cap_{j=1}^k E_{i_j} \cap G \mid \mathcal{G}) = I_G E(I_{\cap_{j=1}^k E_{i_j}} \mid \mathcal{G}) = I_G \prod_{j=1}^k E(I_{E_{i_j}} \mid \mathcal{G}) = \prod_{j=1}^k P(E_{i_j} \cap G \mid \mathcal{G}) \quad (c.t.p.),$$

para todo $E_{i_1} \in \mathcal{E}_{i_1} \subseteq \mathcal{F}$, $i_j \in \{i_1, i_2, \cdots, i_k\} \subseteq I$ y cada $G \in \mathcal{G}$. Por tanto como $E_{i_j} \cap G \in \mathcal{E}_{i_j} \sqcap \mathcal{G}$, se obtiene $\{\mathcal{E}_i \sqcap \mathcal{G}\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$.

(f) Aquí solo hay que tener en cuenta la propiedad *(E4)* de la proposición 1.4.8 y actuar como en la proposición 1.3.12, tomando $I_{E_{i_j}}$ en vez de X_{i_j}.$\square$

De aquí en adelante, se darán condiciones necesarias y/o suficientes para que se preserve la relación de independencia condicionada cuando hay cambios en uno los objetos que intervienen en la definición de independencia condicionada, como son $\{\mathcal{E}_i\}_{i\in I}$, $\mathcal{G}$ y $P(\cdot)$; generalizando así algunas propiedades de la $\mathcal{G}$-independencia de dos σ-álgebras presentadas en Van putten y Van Schuppen [17].

2.2. El problema de cambio en $\{\mathcal{E}_i\}_{i\in I}$.

Es claro que si $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ y si $\mathcal{A}_i \subseteq \mathcal{E}_i$, para cada $i \in I$, entonces $\{\mathcal{A}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$. El siguiente teorema establece que la relación de independencia condicionada se preserva cuando cambiamos cada clase $\mathcal{E}_i$ por el sistema de Dynkin generado por ella.

Teorema 2.2.1 *Sea $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos $\mathcal{E}_i \subseteq \mathcal{F}$. Si $\boldsymbol{D}(\mathcal{E}_i)$ es el sistema de Dynkin generado por $\mathcal{E}_i$, entonces $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ si y sólo si $\{\boldsymbol{D}(\mathcal{E}_i)\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$.*

Demostración. Si $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$, la demostración de que $\{\mathbf{D}(\mathcal{E}_i)\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ es análoga a lo hecho en el teorema 1.1.7, solo hay que tener en cuenta que en este caso se utilizan las propiedades de esperanza condicionada. Si $\{\mathbf{D}(\mathcal{E}_i)\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ es inmediato de la definición de la $\mathcal{G}$-independencia que $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$.$\square$

Corolario 2.2.2 *Sea $\{\mathcal{E}_i\}_{i\in I}$ una familia de π-sistemas; $\mathcal{E}_i \subseteq \mathcal{F}$, y sea $\sigma(\mathcal{E}_i)$ la σ-álgebra generada por $\mathcal{E}_i$. Entonces $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ si y sólo si $\{\sigma(\mathcal{E}_i)\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$.*

Demostración. Aplicando los teoremas 1.1.6 y 2.2.1 se obtiene el resultado.$\square$

En el corolario anterior no se puede omitir que $\mathcal{E}_i$ es un π-sistema como se muestra en el siguiente ejemplo.

Ejemplo 2.2.3 *Sean el espacio de probabilidad $(\Omega,\ \mathcal{F},\ P)$, $\mathcal{E}_1$ y $\mathcal{E}_2$ como en el ejemplo 1.1.9, y $\mathcal{G} = \{\emptyset, \Omega\}$. Entonces de lo hecho en el ejemplo 1.1.9 no es posible $\sigma(\mathcal{E}_1), \sigma(\mathcal{E}_2) \perp\!\!\!\perp \mathcal{G}$; pues $P(E \mid \mathcal{G}) = P(E)$, para todo $E \in \sigma(\mathcal{E}_i)$, $i = 1, 2$.*

Corolario 2.2.4 *Sean $\{\mathcal{F}_i\}_{i\in I}$ es una familia de sub-σ-álgebras de $\mathcal{F}$ y $\mathcal{F}_i \vee \mathcal{G}$ la σ-álgebra generada por $\mathcal{F}_i \cup \mathcal{G}$. Entonces $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ si y sólo si $\{\mathcal{F}_i \vee \mathcal{G}\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$.*

Demostración. Como $\mathcal{F}_i \sqcap \mathcal{G} := \{F_i \cap G : F_i \in \mathcal{F}_i, G \in \mathcal{G}\}$ es un π-sistema, tal que $\mathcal{F}_i \vee \mathcal{G} = \sigma(\mathcal{F}_i \sqcap \mathcal{G})$; entonces por (e) de la proposición 2.1.6 y el corolario anterior, se obtiene lo deseado.$\square$

Corolario 2.2.5 *Sea $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos, donde cada $\mathcal{E}_i$ es un π-sistema y tal que $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$. Sea $\{I_j\}_{j\in J}$ una partición de I. Si $\mathcal{F}_j := \sigma(\cup_{i\in I_j}\mathcal{E}_i)$, entonces $\{\mathcal{F}_j\}_{j\in J} \perp\!\!\!\perp \mathcal{G}$.*

Demostración. La demostración es igual a la hecha en el corolario 1.1.10, sólo que aquí se utiliza el corolario 2.2.2.$\square$

Teorema 2.2.6 *Sean $\{\mathcal{F}_i\}_{i\in I}$ y $\mathcal{G}$ una familia de sub-σ-álgebras de $\mathcal{F}$. Para $J \subseteq I$ sea $\mathcal{F}^J := \sigma(\cup_{I-J}\mathcal{F}_i)$. Las siguientes afirmaciones son equivalentes:*

(a) $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$.

(b) *Para cada $J = \{i_1, i_2, \cdots, i_k\} \subseteq I$; $k \geq 1$, y cada escogencia $F_{i_j} \in \mathcal{F}_{i_j}$, $j = 1, \cdots, k$, se cumple*

$$P(\cap_{j=1}^{k} F_{i_j} \mid \mathcal{G}) = P(\cap_{j=1}^{k} F_{i_j} \mid \mathcal{G} \vee \mathcal{F}^J) \quad (c.t.p.).$$

(c) Para cada $J = \{i_1, i_2, \cdots, i_k\} \subseteq I$; $k \geq 1$, *y cada variable aleatoria integrable* X *que sea* $\mathcal{F}_{i_1} \vee \mathcal{F}_{i_2} \vee \cdots \vee \mathcal{F}_{i_k}$*-medible se cumple*

$$E(X \mid \mathcal{G}) = E(X \mid \mathcal{G} \vee \mathcal{F}^J) \quad (c.t.p.).$$

(d) Para cada subconjunto finito no vacío $J = \{i_1, i_2, \cdots, i_k\}$ *de* I *y cada conjunto de variables aleatorias integrables* $X_{i_1}, X_{i_2}, \cdots, X_{i_k}$ *tal que* X_{i_j} *sea* $\mathcal{F}_{i_j}$*-medible;* $j = 1, 2, \cdots, k$ *y* $\prod_{j=1}^k X_{i_j}$ *sea integrable, se cumple*

$$\prod_{j=1}^{k} E(X_{i_j} | \mathcal{G}) = E(\prod_{j=1}^{k} X_{i_j} | \mathcal{G}) \quad (c.t.p.).$$

Demostración. $(a) \Rightarrow (b)$ Del corolario anterior, $\mathcal{F}_{i_1}, \mathcal{F}_{i_2}, \cdots, \mathcal{F}_{i_k}, \mathcal{F}^J \perp\!\!\!\perp \mathcal{G}$. Si $\mathcal{E} := \{F \cap G : F \in \mathcal{F}^J, G \in \mathcal{G}\}$, entonces $\mathcal{E}$ es un π-sistema que contiene a Ω tal que $\sigma(\mathcal{E}) = \mathcal{F}^J \vee \mathcal{G}$, además

$$\begin{aligned}
P((\cap_{j=1}^k F_{i_j} \cap F) \cap G) &= \int_G I_{\cap_{j=1}^k F_{i_j} \cap F} dP = \int_G P(\cap_{j=1}^k F_{i_j} \cap F \mid \mathcal{G}) dP \\
&= \int_G P(F \mid \mathcal{G}) \prod_{j=1}^{k} P(F_{i_j} \mid \mathcal{G}) dP \\
&= \int_G E(I_F \mid \mathcal{G}) \prod_{j=1}^{k} E(I_{F_{i_j}} \mid \mathcal{G}) dP \\
&= \int_G E(I_F \prod_{j=1}^{k} E(I_{F_{i_j}} \mid \mathcal{G}) \mid \mathcal{G}) \\
&= \int_G I_F \prod_{j=1}^{k} E(I_{F_{i_j}} \mid \mathcal{G}) dP \\
&= \int_{G \cap F} \prod_{j=1}^{k} E(I_{F_{i_j}} \mid \mathcal{G}) dP \\
&= \int_{G \cap F} \prod_{j=1}^{k} P(F_{i_j} \mid \mathcal{G}) = \int_{G \cap F} P(\cap_{j=1}^k F_{i_j} \mid \mathcal{G}) dP.
\end{aligned}$$

De otro lado, como $F \cap G \in \mathcal{F}^J \vee \mathcal{G}$,

$$P(\cap_{j=1}^k F_{i_j} \cap (F \cap G)) = \int_{G \cap F} \prod_{j=1}^{k} I_{F_{i_j}} dP = \int_{G \cap F} P(\cap_{j=1}^k F_{i_j} \mid \mathcal{G} \vee \mathcal{F}^J) dP$$

Por tanto, para todo $F \cap G \in \mathcal{E}$ se tiene

$$\int_{G\cap F} P(\cap_{j=1}^{k} F_{i_j} \mid \mathcal{G})dP = \int_{G\cap F} P(\cap_{j=1}^{k} F_{i_j} \mid \mathcal{G} \vee \mathcal{F}^J)dP,$$

entonces $P(\cap_{j=1}^{k} F_{i_j} \mid \mathcal{G}) = P(\cap_{j=1}^{k} F_{i_j} \mid \mathcal{G} \vee \mathcal{F}^J)$ (c.t.p.).

$(b) \Rightarrow (c)$ Por la hipótesis (c) vale para variables aleatorias de la forma $Y = I_{F_{i_1}\cap F_{i_2}\cap\cdots\cap F_{i_k}}$ donde $F_{i_j} \in \mathcal{F}_{i_j}$, $j = 1, 2\cdots, k$, como $\{F_{i_1} \cap F_{i_2} \cap \cdots \cap F_{i_k} : F_{i_j} \in \mathcal{F}_{i_j}, \quad j = 1, 2, \cdots, k\}$ es un π-sistema que genera a $\mathcal{F}_{i_1} \vee \mathcal{F}_{i_2} \vee \cdots \vee \mathcal{F}_{i_k}$, entonces (c) vale para las variables aleatorias de la forma $Y = I_F$, $F \in \mathcal{F}_{i_1} \vee \mathcal{F}_{i_2} \vee \cdots \vee \mathcal{F}_{i_k}$. Por *(E4)* de la proposición 1.4.8, (c) vale para variables aleatorias simples $Y = \sum_{i=1}^{m} \alpha_i F_i$ que sean $\mathcal{F}_{i_1} \vee \mathcal{F}_{i_2} \vee \cdots \vee \mathcal{F}_{i_k}$-medibles. Para Y arbitraria que sea $\mathcal{F}_{i_1} \vee \mathcal{F}_{i_2} \vee \cdots \vee \mathcal{F}_{i_k}$-medible, existe una sucesión de variables aleatorias simples $Y_1, Y_2, \cdots$ tales que cada Y_n es $\mathcal{F}_{i_1} \vee \mathcal{F}_{i_2} \vee \cdots \vee \mathcal{F}_{i_k}$-medible, $|Y_n| \leq |Y|$ y $Y_n(\omega) \to Y(\omega)$ para todo $\omega \in \Omega$. Entonces por *(E10)* de la proposición 1.4.8, (c.t.p.) se tiene:

$$E(Y|\mathcal{G}) = \lim_{n\in\mathbb{N}} E(Y_n|\mathcal{G}) = \lim_{n\in\mathbb{N}} E(Y_n|\mathcal{G} \vee \mathcal{F}^J) = E(Y|\mathcal{G} \vee \mathcal{F}^J).$$

$(c) \Rightarrow (d)$ Si se toma $\mathcal{F}^J := \sigma(\cup_{I-J}\sigma(X_{i_j}))$, por *(E7)* de la proposición 1.4.8 y por hipótesis, se tiene

$$\begin{aligned}
\prod_{j=1}^{k} E(X_{i_j} \mid \mathcal{G}) &= E(X_{i_1} \mid \mathcal{G})E(X_{i_2} \mid \mathcal{G}) \prod_{j=3}^{k} E(X_{i_j} \mid \mathcal{G}) \\
&= E(X_{i_1} E(X_{i_2} \mid \mathcal{G}) \mid \mathcal{G}) \prod_{j=3}^{k} E(X_{i_j} \mid \mathcal{G}) \\
&= E(X_{i_1} E(X_{i_2} \mid \mathcal{G} \vee \mathcal{F}^{\{i_2\}}) \mid \mathcal{G}) \prod_{j=3}^{k} E(X_{i_j} \mid \mathcal{G}) \\
&= E(E(X_{i_1} X_{i_2} \mid \mathcal{G} \vee \mathcal{F}^{\{i_2\}}) \mid \mathcal{G}) \prod_{j=3}^{k} E(X_{i_j} \mid \mathcal{G}) \\
&= E(X_{i_1} X_{i_2} \mid \mathcal{G}) \prod_{j=3}^{k} E(X_{i_j} \mid \mathcal{G}) \quad (c.t.p.).
\end{aligned}$$

Continuando con este proceso, se llega a

$$\prod_{j=1}^{k} E(X_{i_j} \mid \mathcal{G}) = E(\prod_{j=1}^{k} X_{i_j} \mid \mathcal{G}) \quad (c.t.p.).$$

$(d) \Rightarrow (a)$ Basta tomar $X_{i_j} = I_{F_{i_j}}$, donde $F_{i_j} \in \mathcal{F}_{i_j}$.$\square$

Corolario 2.2.7 *Sean $\{\mathcal{F}_i\}_{i\in I}$ una familia de sub-σ-álgebras de $\mathcal{F}$ y $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Si $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$, entonces para cada $i, j \in I$; $i \neq j$, se cumple $\mathcal{F}_i \cap \mathcal{F}_j \subseteq \mathcal{G}$. Más aún se tiene que $\cap_{i\in I}\mathcal{F}_i \subseteq \mathcal{G}$.*

Demostración. Dados $i \neq j$ en I, de la definición de la independencia dada $\mathcal{G}$, es claro que $\mathcal{F}_i, \mathcal{F}_j \perp\!\!\!\perp \mathcal{G}$. Si $A \in \mathcal{F}_i \cap \mathcal{F}_j$ del teorema 2.2.6 se concluye :

$$P(A \mid \mathcal{G}) = P(A \mid \mathcal{G} \vee \mathcal{F}_j) = E(I_A \mid \mathcal{G} \vee \mathcal{F}_j) = I_A \quad (c.t.p),$$

entonces I_A es $\mathcal{G}$-medible, es decir $A \in \mathcal{G}$.□

2.3. El problema de cambio en $\mathcal{G}$.

Dado $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$, $\mathcal{G}$, $\mathcal{G}_1$ sub-σ-álgebras de $\mathcal{F}$. Los siguientes ejemplos y propiedades muestran algunos de los posibles efectos que se tienen al variar por lo cual se condiciona.

Ejemplo 2.3.1 *Dado $\Omega = \{1, 2, 3, 4\}$, $\mathcal{F} = \wp(\Omega)$ y $P(\{i\}) = \frac{1}{4}$, para cualquier $i \in \Omega$.*

(1) Si $F_1 = \{1, 2, 3\}$, $F_2 = \{2, 3, 4\}$, $G = \{2, 3\}$ y $\mathcal{G} = \sigma(G)$ entonces $F_1, F_2 \perp\!\!\!\perp G$ pero no se cumple que $F_1, F_2 \perp\!\!\!\perp G^c$, por tanto tampoco se da que $F_1, F_2 \perp\!\!\!\perp \mathcal{G}$.

(2) Sean $F_1 = \{1, 2\}$, $F_2 = \{1, 3\}$ y $G = \{2, 3\}$. Si $\mathcal{F}_1 = \sigma(\{F_1\})$, $\mathcal{F}_2 = \sigma(\{F_2\})$, $\mathcal{G}_1 = \{\emptyset, \Omega\}$ y $\mathcal{G}_2 = \sigma(\{G\})$; como los eventos F_1 y F_2 son independientes, entonces $\mathcal{F}_1$ y $\mathcal{F}_2$ son $\mathcal{G}_1$-independientes. No obstante, $\mathcal{F}_1$ y $\mathcal{F}_2$ no son $\mathcal{G}_2$-independientes, pues

$$P(F_1 \cap F_2|\mathcal{G}_2) = \frac{1}{2}I_{G^c} \neq (\frac{1}{2})(\frac{1}{2}) = P(F_1|\mathcal{G}_2)P(F_2|\mathcal{G}_2).$$

(3) Sean $F_1 = \{1\}$, $F_2 = \{3\}$ y $G = \{1, 2\}$. Sean ahora $\mathcal{F}_1 = \sigma(\{F_1\})$, $\mathcal{F}_2 = \sigma(\{F_2\})$, $\mathcal{G}_1 = \{\emptyset, \Omega\}$ y $\mathcal{G}_2 = \sigma(\{G\})$; como los eventos F_1 y F_2 no son independientes, entonces $\mathcal{F}_1$ y $\mathcal{F}_2$ no son $\mathcal{G}_1$-independientes. Sin embargo, $\mathcal{F}_1$ y $\mathcal{F}_2$ son $\mathcal{G}_2$-independientes, para mostrarlo, es suficiente con tener la igualdad

$$P(F_1 \cap F_2|\mathcal{G}_2) = 0 = (\frac{1}{2}I_G)(\frac{1}{2}I_{G^c}) = P(F_1|\mathcal{G}_2)P(F_2|\mathcal{G}_2).$$

(4) Si $F_1 = \{1, 2\}$, $F_2 = \{3, 4\}$, $F_3 = \{2, 3\}$ y $G = \{4\}$, entonces $P(F_1 \cap (F_2 \cup F_3)|G) = P(\{2\}|G) = 0 = P(F_1|G)P(F_2 \cup F_3|G)$, $P(F_1 \cap F_2|G) = 0 = P(F_1|G)P(F_2|G)$ y $P(F_1 \cap F_3|G \cup F_2) = P(\{2\}|\{3, 4\}) = 0 = P(F_1|G \cup F_2)P(F_3|G \cup F_2)$. Es decir que $F_1, F_2 \cup F_3 \perp\!\!\!\perp G$; $F_1, F_2 \perp\!\!\!\perp G$ y $F_1, F_3 \perp\!\!\!\perp G \cup F_2$.

En el ejemplo anterior se mostró que en general al hacer más fina ó más gruesa la σ-álgebra condicionante $\mathcal{G}$, la relación de independencia condicionada no se preserva; es decir, si $\mathcal{G}_1$, $\mathcal{G}_2$ y $\mathcal{F}_i$ son sub-σ-álgebras de $\mathcal{F}$; $i \in I$,

(i) $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}_1$ y $\mathcal{G}_1 \subseteq \mathcal{G}_2$ no implican $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}_2$.

(ii) $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}_2$ y $\mathcal{G}_1 \subseteq \mathcal{G}_2$ no implican $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}_1$.

Sin embargo, la $\mathcal{G}$-independencia de dos sub-σ-álgebras se preserva cuando hay cierta relación entre la σ-álgebras en cuestión (ver Van Putten y Van Schuppen [17]). En especial se tienen las propiedades siguientes.

Proposición 2.3.2 *Si* $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$ *y* $\mathcal{G} \subseteq \mathcal{G}_1 \subseteq \mathcal{F}_1$ *entonces* $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}_1$.

Demostración. Como $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$ entonces del teorema 2.2.6, para todo $F_1 \in \mathcal{F}_1$ y $F_2 \in \mathcal{F}_2$, se tiene $P(F_2|\mathcal{F}_1) = P(F_2|\mathcal{G} \vee \mathcal{F}_1) = P(F_2|\mathcal{G})$ y también

$$P(F_2|\mathcal{G}_1) = P(P(F_2|\mathcal{F}_1)|\mathcal{G}_1) = P(P(F_2|\mathcal{G})|\mathcal{G}_1) = P(F_2|\mathcal{G}),$$

de lo cual $P(F_2|\mathcal{G}_1) = P(F_2|\mathcal{F}_1 \vee \mathcal{G}_1)$ y nuevamente de 2.2.6 se tiene $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}_1$.□

Otro resultado el cual sera útil adelante y es una generalización de el ejemplo (4) de 2.3.1 es el siguiente.

Proposición 2.3.3 $\mathcal{F}_1, \mathcal{F}_2 \vee \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$ *si y sólo si* $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$ *y* $\mathcal{F}_1, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G} \vee \mathcal{F}_2$.

Demostración. Dado que los conjuntos $\mathcal{E}_1 = \{F_2 \cap F_3 : F_2 \in \mathcal{F}_2, F_3 \in \mathcal{F}_3\}$ y $\mathcal{E}_2 = \{G \cap F_2 : G \in \mathcal{G}, F_2 \in \mathcal{F}_2\}$ son π-sistemas que contienen a Ω, tal que $\sigma(\mathcal{E}_1) = \mathcal{F}_2 \vee \mathcal{F}_3$ y $\sigma(\mathcal{E}_2) = \mathcal{G} \vee \mathcal{F}_2$, se tiene:

($\Leftarrow$) Si $\mathcal{F}_1, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G} \vee \mathcal{F}_2$, $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$ y $G \in \mathcal{G}$, entonces del teorema 2.2.6 parte (b) y de $(E14)$ de la proposición 1.4.8, se cumple:

$$\begin{aligned}
\int_G P(F_1 \cap (F_2 \cap F_3)|\mathcal{G})dP &= \int_G I_{F_1\cap F_2\cap F_3} dP = \int_{G\cap F_2} I_{F_1\cap F_3} dP \\
&= \int_{G\cap F_2} P(F_1 \cap F_3|\mathcal{G} \vee \mathcal{F}_2)dP \\
&= \int_{G\cap F_2} P(F_1|\mathcal{G} \vee \mathcal{F}_2)P(F_3|\mathcal{G} \vee \mathcal{F}_2)dP \\
&= \int_{G\cap F_2} E(I_{F_3}P(F_1|\mathcal{G} \vee \mathcal{F}_2)|\mathcal{G} \vee \mathcal{F}_2)dP
\end{aligned}$$

$$
\begin{aligned}
&= \int_{G\cap F_2} I_{F_3} P(F_1|\mathcal{G}\vee\mathcal{F}_2)dP\\
&= \int_{G} I_{F_2\cap F_3} P(F_1|\mathcal{G})dP\\
&= \int_{G} E(I_{F_2\cap F_3} P(F_1|\mathcal{G})|\mathcal{G})dP\\
&= \int_{G} P(F_1|\mathcal{G})P(F_2\cap F_3|\mathcal{G})dP,
\end{aligned}
$$

de lo cual $P(F_1\cap(F_2\cap F_3)|\mathcal{G}) = P(F_1|\mathcal{G})P(F_2\cap F_3|\mathcal{G})$ (c.t.p.). Entonces al tomar en cuenta lo supuesto el principio de la demostración y de lo dicho en 1.4.4, se tiene que $\mathcal{F}_1, \mathcal{F}_2\vee\mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$.

($\Rightarrow$) Si $\mathcal{F}_1, \mathcal{F}_2\vee\mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$, como $\mathcal{F}_2\subseteq\mathcal{F}_2\vee\mathcal{F}_3$, es claro que $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$. Por tanto al tener en cuenta esto, lo dicho en al inicio de la demostración y la definición 1.4.4, se tiene

$$
\begin{aligned}
\int_{G\cap F_2} P(F_1\cap F_3|\mathcal{G}\vee\mathcal{F}_2)dP &= \int_{G\cap F_2} I_{F_1\cap F_3}dP = \int_G P(F_1\cap(F_2\cap F_3)|\mathcal{G})dP\\
&= \int_G P(F_1|\mathcal{G})P(F_2\cap F_3|\mathcal{G})dP\\
&= \int_G E(I_{F_2\cap F_3}P(F_1|\mathcal{G})|\mathcal{G})dP\\
&= \int_{G\cap F_2} I_{F_3}P(F_1|\mathcal{G})dP = \int_{G\cap F_2} I_{F_3}P(F_1|\mathcal{G}\vee\mathcal{F}_2)dP\\
&= \int_{G\cap F_2} E(I_{F_3}P(F_1|\mathcal{G}\vee\mathcal{F}_2)|\mathcal{G}\vee\mathcal{F}_2)dP\\
&= \int_{G\cap F_2} P(F_1|\mathcal{G}\vee\mathcal{F}_2)P(F_3|\mathcal{G}\vee\mathcal{F}_2)dP,
\end{aligned}
$$

que equivale a $P(F_1\cap F_3|\mathcal{G}\vee\mathcal{F}_2) = P(F_1|\mathcal{G}\vee\mathcal{F}_2)P(F_3|\mathcal{G}\vee\mathcal{F}_2)$ (c.t.p.), es decir que $\mathcal{F}_1, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}\vee\mathcal{F}_2$.$\square$

Teorema 2.3.4 *Sean $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ y $\mathcal{G}$ sub-σ-álgebras de $\mathcal{F}$.*

(a) Si $\mathcal{F}_1, \mathcal{F}_2\vee\mathcal{G} \perp\!\!\!\perp$ entonces $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$.

(b) Si $\mathcal{F}_1\vee\mathcal{F}_2\vee\mathcal{G}, \mathcal{F}_3 \perp\!\!\!\perp$. Entonces $\mathcal{F}_1, \mathcal{F}_2\vee\mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$ si y sólo si $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$.

Demostración. (a) Dado que $\mathcal{F}_1, \mathcal{F}_2\vee\mathcal{G} \perp\!\!\!\perp$, entonces en particular $\mathcal{F}_1, \mathcal{G} \perp\!\!\!\perp$, por tanto del teorema 2.2.6 y para $F_1\in\mathcal{F}_1$ se tiene que $P(F_1|\mathcal{G}) = P(F_1|\{\emptyset,\Omega\}\vee\mathcal{G}) = P(F_1)$ y $P(F_1|\mathcal{F}_2\vee\mathcal{G}) = P(F_1|\{\emptyset,\Omega\}\vee\mathcal{F}_2\vee\mathcal{G}) = P(F_1)$. Entonces $P(F_1|\mathcal{F}_2\vee\mathcal{G}) = P(F_1|\mathcal{G})$ y nuevamente de 2.2.6 se obtiene lo requerido.

(b) $(\Rightarrow)$ Como $\mathcal{F}_2 \subseteq \mathcal{F}_2 \vee \mathcal{F}_3$ entonces es claro que $\mathcal{F}_1, \mathcal{F}_2 \vee \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$ implica $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$.

$(\Leftarrow)$ De (a) se tiene que $\mathcal{F}_1 \vee \mathcal{F}_2 \vee \mathcal{G}, \mathcal{F}_3 \perp\!\!\!\perp$ implica que $\mathcal{F}_1, \mathcal{F}_3 \perp\!\!\!\perp \mathcal{F}_2 \vee \mathcal{G}$ y como $\mathcal{F}_1, \mathcal{F}_2 \perp\!\!\!\perp \mathcal{G}$ entonces de la proposición 2.3.3 se tiene $\mathcal{F}_1, \mathcal{F}_2 \vee \mathcal{F}_3 \perp\!\!\!\perp \mathcal{G}$.$\square$

En general se tienen los siguientes resultados.

Teorema 2.3.5 *Sean $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos y $\mathcal{D} = \{D_1, D_2, \cdots\}$ una descomposición de Ω. Si $\mathcal{G} = \sigma(\mathcal{D})$, entonces $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$ si y sólo si $\{\mathcal{E}_i\}_{i\in I} \perp\!\!\!\perp D_n$ para $n = 1, 2, \cdots$.*

Demostración. Para cada $A \in \mathcal{F}$ se tiene que $P(A \mid \mathcal{G})$ es $\sigma(\mathcal{D})$-medible, entonces $P(A \mid \mathcal{G}) = \sum_{n\geq 1} P(A \mid D_n) I_{D_n}$ (c.t.p.), así para cada $\{i_1, i_2, \cdots, i_k\} \subseteq I$; $k \geq 2$, y cada escogencia $E_{i_j} \in \mathcal{E}_{i_j}$, $j = 1, 2, \cdots, k$,

$$P(\cap_{j=1}^k E_{i_j} \mid \mathcal{G}) = \sum_{n\geq 1} P(\cap_{j=1}^k E_{i_j} \mid D_n) I_{D_n} \quad (c.t.p.),$$

y

$$\prod_{j=1}^k P(E_{i_j} \mid \mathcal{G}) = \sum_{n\geq 1} (\prod_{j=1}^k P(E_{i_j} \mid D_n)) I_{D_n} \quad (c.t.p.),$$

por lo tanto $P(\cap_{j=1}^k E_{i_j} \mid \mathcal{G}) = \prod_{j=1}^k P(E_{i_j} \mid \mathcal{G})$ (c.t.p.) equivale a $P(\cap_{j=1}^k E_{i_j} \mid D_n) = \prod_{j=1}^k P(E_{i_j} \mid D_n)$ (c.t.p.).$\square$

Teorema 2.3.6 *Sean $\{\mathcal{F}_i\}_{i\in I}$ una familia de sub-σ-álgebras de $\mathcal{F}$ y $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Si $\{\mathcal{F}_i\}_{i\in I} \perp\!\!\!\perp \mathcal{G}$, entonces para cada $J := \{i_1, i_2, \cdots, i_k\} \subseteq I$; $k \geq 2$, se cumple $\mathcal{F}_{i_1}, \mathcal{F}_{i_2}, \cdots, \mathcal{F}_{i_k} \perp\!\!\!\perp \mathcal{G} \vee \mathcal{F}^J$, donde $\mathcal{F}^J := \sigma(\cup_{I-J} \mathcal{F}_i)$.*

Demostración. Para $F_{i_j} \in \mathcal{F}_{i_j}$, $j = 1, 2, \cdots, k$, del teorema 2.2.6 se tienen $(c.t.p.)$ las siguientes igualdades

$$\begin{aligned} P(\cap_{j=1}^k F_{i_j} \mid \mathcal{G} \vee \mathcal{F}^J) &= P(\cap_{j=1}^k F_{i_j} \mid \mathcal{G}) = \prod_{j=1}^k P(F_{i_j} \mid \mathcal{G}) = \prod_{j=1}^k E(I_{F_{i_j}} \mid \mathcal{G}) \\ &= \prod_{j=1}^k E(E(I_{F_{i_j}} \mid \mathcal{G}) \mid \mathcal{G} \vee \mathcal{F}^J) \\ &= \prod_{j=1}^k E(E(I_{F_{i_j}} \mid \mathcal{G} \vee \mathcal{F}^{\{i_j\}}) \mid \mathcal{G} \vee \mathcal{F}^J) \end{aligned}$$

$$= \prod_{j=1}^{k} E(I_{F_{i_j}} \mid \mathcal{G} \vee \mathcal{F}^J) = \prod_{j=1}^{k} P(F_{i_j} \mid \mathcal{G} \vee \mathcal{F}^J).\square$$

Corolario 2.3.7 *Si $\{\mathcal{F}_i\}_{i\in I}$ es una familia independiente de sub-σ-álgebras de $\mathcal{F}$, entonces para cada $J := \{i_1, i_2, \cdots, i_k\} \subseteq I$; $k \geq 2$, se cumple $\mathcal{F}_{i_1}, \mathcal{F}_{i_2}, \cdots, \mathcal{F}_{i_k} \perp\!\!\!\perp \mathcal{F}^J$.*

Demostración. Sólo basta tomar $\mathcal{G} = \{\emptyset, \Omega\}$ en el teorema anterior.$\square$

Observación 2.3.8 *El caso recíproco del corolario anterior no es cierto, pues si $\mathcal{F}_i = \mathcal{F}$ para todo $i \in I$, entonces de la $\mathcal{F}$-medibilidad se tiene que para $\{i_1, \cdots, i_k\} \subseteq I$; $k \geq 2$, se cumple $\mathcal{F}_{i_1}, \mathcal{F}_{i_2}, \cdots, \mathcal{F}_{i_k} \perp\!\!\!\perp \mathcal{F}^J$, ya que $\mathcal{F}^J = \mathcal{F}$; pero la familia $\{\mathcal{F}_i\}_{i\in I}$ no es independiente.*

2.4. El problema de cambio en P.

De ahora en adelante, para P y Q medidas de probablidad del espacio medible $(\Omega, \mathcal{F})$, $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$ y $\{\mathcal{E}_i\}_{i\in I}$ una familia de clases de eventos $\mathcal{E}_i \subseteq \mathcal{F}$; la esperanza con respecto a estas medidas se denotará por $E_P(\cdot)$ y $E_Q(\cdot)$ respectivamente. Además la $\mathcal{G}$-independencia de $\{\mathcal{E}_i\}_{i\in I}$ con respecto a estas medidas se simboliza por $\{\mathcal{E}_i\}_{i\in I} \underset{P}{\perp\!\!\!\perp} \mathcal{G}$ y $\{\mathcal{E}_i\}_{i\in I} \underset{Q}{\perp\!\!\!\perp} \mathcal{G}$.

Teorema 2.4.1 *Si Q es absolutamente continua con respecto a P ($Q << P$) y $dQ = Y dP$, entonces para cada variable aleatoria X que sea Q-integrable se cumple*

$$E_P(XY \mid \mathcal{G}) = E_Q(X \mid \mathcal{G}) E_P(Y \mid \mathcal{G}), \quad (P - c.t.p.).$$

Demostración. Para cada $G \in \mathcal{G}$, por definición y propiedades *(E7)*, *(E14)* de la proposición 1.4.8, se tiene:

$$\begin{aligned}\int_G E_P(XY \mid \mathcal{G})dP &= \int_G XY dP = \int_G X dQ = \int_G E_Q(X \mid \mathcal{G}) dQ = \int_G E_Q(X \mid \mathcal{G}) Y dP \\ &= \int_G E_P(E_Q(X \mid \mathcal{G})Y \mid \mathcal{G})dP = \int_G E_Q(X \mid \mathcal{G}) E_P(Y \mid \mathcal{G}) dP.\end{aligned}$$

Entonces $E_P(XY \mid \mathcal{G}) = E_Q(X \mid \mathcal{G}) E_P(Y \mid \mathcal{G})$, $(P - c.t.p.)$.$\square$

Observación 2.4.2 *Si en el teorema anterior se toma $Y = \frac{I_B}{P(B)}$ y $X = I_A$, donde $A, B \in \mathcal{F}$ y $P(B) > 0$, entonces se obtiene*

$$P(A \cap B|\mathcal{G}) = Q(A|\mathcal{G}) P(B|\mathcal{G}), \quad (P - c.t.p.).$$

Proposición 2.4.3 *Si $Q << P$ y $dQ = YdP$, donde Y es una variable aleatoria $\mathcal{G}$-medible y $\{\mathcal{E}_i\}_{i\in I}$ es una familia de clases de eventos, entonces $\{\mathcal{E}_i\}_{i\in I} \underset{P}{\perp\!\!\!\perp} \mathcal{G}$ implica $\{\mathcal{E}_i\}_{i\in I} \underset{Q}{\perp\!\!\!\perp} \mathcal{G}$.*

Demostración. Por el teorema anterior y la $\mathcal{G}$-medibilidad de Y, para cada $A \in \mathcal{F}$ las siguientes igualdades se cumplen $(P - c.t.p.)$

$$YE_P(I_A \mid \mathcal{G}) = E_P(YI_A \mid \mathcal{G}) = E_P(Y \mid \mathcal{G})E_Q(I_A \mid \mathcal{G}) = YE_Q(I_A \mid \mathcal{G})$$

Por tanto, sobre el evento $\{Y \neq 0\} \in \mathcal{G}$ se tiene $E_P(I_A \mid \mathcal{G}) = E_Q(I_A \mid \mathcal{G})$ $(Q - c.t.p.)$.□

Corolario 2.4.4 *Si $dQ = \frac{I_G}{P(G)}dP$, donde $G \in \mathcal{G}$, entonces $\{\mathcal{E}_i\}_{i\in I} \underset{P}{\perp\!\!\!\perp} \mathcal{G}$ implica $\{\mathcal{E}_i\}_{i\in I} \underset{Q}{\perp\!\!\!\perp} \mathcal{G}$.*

Demostración. Esto se obtiene, sólo tomando $Y = \frac{I_G}{P(G)}$ en la proposición anterior.□

En el corolario anterior no se puede omitir que $G \in \mathcal{G}$ como se muestra en la parte (2) del siguiente ejemplo.

Ejemplo 2.4.5

(1) Si en el corolario anterior tomamos $\mathcal{G} = \{\emptyset, \Omega\}$ entonces $\{\mathcal{E}_i\}_{i\in I} \underset{P}{\perp\!\!\!\perp}$ implica $\{\mathcal{E}_i\}_{i\in I} \underset{Q}{\perp\!\!\!\perp}$.

(2) Si $\Omega = \{1, 2, 3, 4\}$, $\mathcal{F} = \wp(\Omega)$, $\mathcal{G} = \{\emptyset, \Omega\}$, $A_1 = \{1, 2\}$, $A_2 = \{1, 3\}$, $G = \{2, 3\}$ y $P(\{i\}) = \frac{1}{4}$, entonces es claro que $G \notin \mathcal{G}$ y además el anterior corolario no se cumple, pues $P(A_1 \cap A_2 \mid \mathcal{G}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A_2 \mid \mathcal{G})P(A_2 \mid \mathcal{G})$; pero

$$Q(A_1 \cap A_2 \mid \mathcal{G}) \neq Q(A_2 \mid \mathcal{G})Q(A_2 \mid \mathcal{G}),$$

puesto que

$$Q(A_1 \cap A_2 \mid \mathcal{G}) = \int_{A_1 \cap A_2} dQ = \frac{1}{P(G)} \int_{A_1 \cap A_2 \cap G} dP = 2P(A_1 \cap A_2 \cap G) = 0$$

y

$$Q(A_1 \mid \mathcal{G}) = \frac{1}{P(G)} \int_{A_1 \cap G} dP = 2 \cdot \frac{1}{4} = \frac{1}{2} = Q(A_2 \mid \mathcal{G}).$$

Proposición 2.4.6 *Si* $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_n \underset{P}{\perp\!\!\!\perp} \mathcal{G}$ *y* $dQ = \frac{I_B}{P(B)} dP$*, donde* $B = \cap_{i=1}^n B_i$ *con* $B_i \in \mathcal{E}_i$ *para* $i = 1, 2, \cdots, n$*, entonces* $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_n \underset{Q}{\perp\!\!\!\perp} \mathcal{G}$*.*

Demostración. Por la hipótesis, $P(B|\mathcal{G}) = P(\cap_{i=1}^n B_i|\mathcal{G}) = \prod_{i=1}^n P(B_i|\mathcal{G})$. De acuerdo con la observación 2.4.2, para cada $E_i \in \mathcal{E}_i$, $i = 1, 2, \cdots, n$,

$$Q(E_i|\mathcal{G}) = \frac{P(E_i \cap B|\mathcal{G})}{P(B|\mathcal{G})} = \frac{P(E_i \cap B_i|\mathcal{G})}{P(B_i|\mathcal{G})}.$$

Ahora, si $\{i_1, i_2, \cdots, i_k\} \subseteq \{1, 2, \cdots, n\}$; $k \geq 2$ y $E_{i_j} \in \mathcal{E}_{i_j}$, para $j = 1, 2, \cdots k$, entonces

$$Q(\cap_{j=i}^k E_{i_j}|\mathcal{G}) = \frac{P((\cap_{j=1}^k E_{i_j}) \cap B|\mathcal{G})}{P(B|\mathcal{G})} = \prod_{j=1}^k \frac{P(E_{i_j} \cap B_{i_j}|\mathcal{G})}{P(B_{i_j}|\mathcal{G})} = \prod_{j=1}^k Q(E_{i_j}|\mathcal{G}).\square$$

2.5. Independencia condicionada de variables aleatorias.

En el corolario 2.2.2 se consideró que una familia de eventos $\{E_i\}_{i \in I}$ es independiente dada una sub-σ-álgebra $\mathcal{G}$ si y sólo si la familia de σ-álgebras $\{\sigma(I_{E_i})\}_{i \in I}$ es independiente dada la misma $\mathcal{G}$. En la siguiente definición se establece que sucede para variables aleatorias en general.

Definición 2.5.1 *La familia de variables aleatorias* $\{X_i\}_{i \in I}$ *definidas sobre* $(\Omega, \mathcal{F}, P)$ *es* $\mathcal{G}$*-independiente (*$\mathcal{G}$ *una sub-*σ*-álgebra de* $\mathcal{F}$*) si* $\{\sigma(X_i)\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$*; lo que se simboliza como* $\{X_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$*.*

Ejemplo 2.5.2 *Sean* $(\Omega, \mathcal{F}, P)$ *un espacio de probabilidad y* $\{X_n\}_{n \in \mathbb{N}}$ *una sucesión de variables aleatorias independientes, entonces del corolario 2.3.7 para la* σ*-álgebra* $\mathcal{G} := \sigma(X_{k+1}, X_{k+2}, \cdots)$ *se tiene* $X_1, \cdots, X_k \perp\!\!\!\perp \mathcal{G}$*. Ahora si* $\mathcal{T}_\infty := \bigcap_{n \in \mathbb{N}} \mathcal{T}_n$*, con* $\mathcal{T}_n := \sigma(X_n, X_{n+1}, \cdots)$*, es la* σ*-álgebra de eventos terminales asociados a* $\{X_n\}_{n \in \mathbb{N}}$ *entonces de la ley cero-uno de kolmogorov para todo* $G \in \mathcal{T}_\infty$*,* $P(G) = 0$ *ó* $P(G) = 1$*, de lo cual por proposición 1.1.3 se concluye que* $\sigma(X_1, X_2, \cdots)$ *es independiente de* $\mathcal{T}_\infty$ *y de la parte (d) de la proposición 2.1.6 se concluye que la familia de clases de eventos* $\sigma(X_n) \subseteq \sigma(X_1, X_2, \cdots)$*,* $n \in \mathbb{N}$*, es* $\mathcal{T}_\infty$*-independiente, es decir* $\{X_n\}_{n \in \mathbb{N}} \perp\!\!\!\perp \mathcal{T}_\infty$*.*

Teorema 2.5.3 *Una familia de objetos aleatorios* $\{X_i\}_{i \in I}$*, con imagenes sobre los espacios de medida* $(\Omega_i, \mathcal{F}_i)$ *donde* $\mathcal{D}_i$ *es el* π*-sistema que genera a* $\mathcal{F}_i$ *y* $\Omega_i \in \mathcal{D}_i$*, es independiente si y sólo si*

$$P(\cap_{j=1}^k X_{i_j}^{-1}(Q_{i_j})|\mathcal{G}) = \prod_{j=1}^k P(X_{i_j}^{-1}(Q_{i_j})|\mathcal{G}) \quad (c.t.p.)$$

para cualquier escogencia finita $X_{i_1}, X_{i_2}, \cdots, X_{i_k}$ *de* $\{X_i\}_{i \in I}$ *y cualquier* $Q_{i_j} \in \mathcal{D}_{i_j}$, $j = 1, \cdots, k$.

Demostración. La demostración es análoga a la del teorema 1.3.2, sólo que en este caso se utiliza el corolario 2.2.2.$\square$

Observación 2.5.4

(1) Como $\pi(X_i) = \{\{X_i \leq x\}, x \in \mathbb{R}\}$ *es un* π*-sistema que genera a* $\sigma(X_i)$*, para todo* $i \in I$*, entonces del teorema anterior* $\{X_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$ *si y sólo si* $\{\pi(X_i)\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$.

(2) Para $Y_i = g_i(X_i)$*, con* $g_i : \Omega \to \mathbb{R}$ *es una función boreliana para cada* $i \in I$ *y* $\{X_i\}_{i \in I}$ *una familia de variables aleatorias, se tiene que* $\sigma(Y_i) \subseteq \sigma(X_i)$ *y por tanto* $\{X_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$ *implica* $\{Y_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$.

Proposición 2.5.5 *Sea* $\{X_i\}_{i \in I}$ *una familia de variables aleatorias integrables. Entonces*

(1) $\{E(X_i|\mathcal{G})\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$.

(2) $\{X_i\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$ *si y sólo si* $\{X_i - E(X_i|\mathcal{G})\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$.

Demostración. (1) Esto se sigue de (b) de la proposición 2.1.6, puesto que $E(X_i|\mathcal{G})$ por definición 1.4.4 es $\mathcal{G}$-medible.

(2)($\Rightarrow$) Por la hipótesis y por el corolario 2.2.4, $\{\sigma(X_i) \vee \mathcal{G}\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$. El resultado se sigue de la inclusión $\sigma(X_i - E(X_i|\mathcal{G})) \subseteq \sigma(X_i) \vee \mathcal{G}$.

($\Leftarrow$) Por la hipótesis y por el corolario 2.2.4, $\{\sigma(X_i - E(X_i|\mathcal{G})) \vee \mathcal{G}\}_{i \in I} \perp\!\!\!\perp \mathcal{G}$. El resultado se sigue de la inclusión $\sigma(X_i) \subseteq \sigma(X_i - E(X_i|\mathcal{G})) \vee \mathcal{G}$.$\square$

Corolario 2.5.6 *Si* X_1 *y* X_2 *son variables aleatorias integrables* $\mathcal{G}$*-independientes, entonces* $Cov(X_1, X_2|\mathcal{G}) = 0$ *(c.t.p.).*

Demostración. El resultado se obtiene de aplicar (2) de la proposición 2.5.5 y (2) de la observación 2.5.4.$\square$

2.6. Familia de σ-álgebras recíprocas y markovianas.

Es sabido que $\{X_t\}_{t\in T}$ es un proceso de Markov si para $B \in \mathcal{B}(\mathbb{R})$ y cada $s, t \in T$ con $s < t$ se tiene

$$P(X_t \in B \mid \mathcal{F}_s^i) = P(X_t \in B \mid X_s) \quad (c.t.p.),$$

donde $\mathcal{F}_t^i = \sigma(X_s, s \leq t)$. En el libro de Loève [8] se dice que una familia de variables aleatorias $\{X_t\}_{t\in T}$ es un proceso de Markov, si la familia de σ-álgebras $\{\sigma(X_t)\}_{t\in T}$ forma una familia markoviana, por esta razón en esta Sección se hará uso de todo el desarrollo hecho hasta ahora, para exponer algunas propiedades y definiciones equivalentes que tienen este tipo de familias, que inducen este tipo tan importante de procesos.

Definición 2.6.1 *Dado $(\Omega,\mathcal{F},P)$ un espacio de probabilidad. Una filtración es una familia no decreciente $\{\mathcal{F}_t\}_{t\in I}$ de sub-σ-álgebras de $\mathcal{F}$, es decir $\mathcal{F}_s \subseteq \mathcal{F}_t$, si $s < t$.*

Ejemplo 2.6.2 *Dada una sucesión de variables aleatorias $X = \{X_t\}_{t\geq 0}$ la familia de sub-σ-álgebras $\mathcal{F}_t^X = \sigma(X_r : 0 \leq r \leq t)$ es una filtración, para la cual se cumple que X_t es $\mathcal{F}_t^X$-medible.*

Definición 2.6.3 *Dado un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y $G := \{\mathcal{G}_t\}_{t\in\mathbb{R}^+}$ una familia de sub-σ-álgebras de $\mathcal{F}$; $\mathcal{F}_I := \sigma(\mathcal{G}_t : t \in I)$, $I \subseteq \mathbb{R}^+$.*

(a) G es una familia markoviana, si para cada $t \in \mathbb{R}^+$ se verifica

$$\mathcal{F}_{[0,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{t\}}.$$

(b) G es una familia recíproca, si para cada $0 \leq s \leq t$ se verifica

$$\mathcal{F}_{[s,t]}, \mathcal{F}_{[0,s]} \vee \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}}.$$

Ejemplo 2.6.4 *Si $G := \{\mathcal{F}_t\}_{t\in I}$ es una filtración, entonces G es una familia markoviana. Esto es cierto, pues para todo $A \in \mathcal{F}_{[0,t]}$ y $B \in \mathcal{F}_{[t,\infty)}$, se tiene*

$$P(A \cap B|\mathcal{F}_t) = E(I_A I_B|\mathcal{F}_t) = I_A E(I_B|\mathcal{F}_t) = P(A|\mathcal{F}_t)P(B|\mathcal{F}_t),$$

Más aún, la familia G es una familia recíproca por el corolario del próximo teorema.

Teorema 2.6.5 *Una familia G es markoviana si y sólo si para $0 \leq s \leq t$, se cumple*

$$\mathcal{F}_{[0,s]}, \mathcal{F}_{[s,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}}.$$

Demostración. ($\Rightarrow$) Si G es markoviana entonces $\mathcal{F}_{[0,i]}, \mathcal{F}_{[i,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{i\}}$, para $i = t$ ó $i = s$, como $\mathcal{F}_{\{s\}} \subseteq \mathcal{F}_{\{s,t\}} \subseteq \mathcal{F}_{[s,\infty)}$, entonces para cada $A \in \mathcal{F}_{[0,s]}$

$$P(A|\mathcal{F}_{[s,\infty)}) = P(A|\mathcal{F}_{[s,\infty)} \vee \mathcal{F}_{\{s\}}) = P(A|\mathcal{F}_{\{s\}}),$$

de lo cual

$$P(A|\mathcal{F}_{[s,\infty)} \vee \mathcal{F}_{\{s,t\}}) = P(A|\mathcal{F}_{[s,\infty)}) = P(A|\mathcal{F}_{\{s\}}),$$

por tanto

$$\begin{aligned} P(A|\mathcal{F}_{\{s,t\}}) &= P(P(A|\mathcal{F}_{[s,\infty)})|\mathcal{F}_{\{s,t\}}) = P(P(A|\mathcal{F}_{\{s\}})|\mathcal{F}_{\{s,t\}}) \\ &= P(A|\mathcal{F}_{\{s\}}) = P(A|\mathcal{F}_{[s,\infty)} \vee \mathcal{F}_{\{s,t\}}) \end{aligned}$$

y por el teorema 2.2.6, se tiene que $\mathcal{F}_{[0,s]}, \mathcal{F}_{[s,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}}$. Análogamente como $\mathcal{F}_{\{t\}} \subseteq \mathcal{F}_{\{s,t\}} \subseteq \mathcal{F}_{[0,t]}$ entonces $\mathcal{F}_{[0,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}}$. Por tanto para $A \in \mathcal{F}_{[0,s]}$, $B \in \mathcal{F}_{[s,t]}$, $C \in \mathcal{F}_{[t,\infty]}$; se tiene que $A \cap B \in \mathcal{F}_{[0,t]}$, $B \in \mathcal{F}_{[s,\infty]}$ y

$$P(A\cap B\cap C|\mathcal{F}_{\{s,t\}}) = P(A\cap B|\mathcal{F}_{\{s,t\}})P(C|\mathcal{F}_{\{s,t\}}) = P(A|\mathcal{F}_{\{s,t\}})P(B|\mathcal{F}_{\{s,t\}})P(C|\mathcal{F}_{\{s,t\}}),$$

concluyendo que $\mathcal{F}_{[0,s]}, \mathcal{F}_{[s,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}}$.

($\Leftarrow$) Basta tomar $t = s$ en la hipótesis.$\square$

Corolario 2.6.6 *Si G es una familia markoviana, entonces G es una familia recíproca.*

Demostración. Si G es una familia markoviana donde $0 \leq s \leq t$, entonces del teorema anterior y el corolario 2.2.5 $\mathcal{F}_{[s,t]}, \mathcal{F}_{[0,s]} \vee \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}}$; obteniendo lo requerido.$\square$

El siguiente ejemplo muestra que el recíproco del corolario anterior es falso.

Ejemplo 2.6.7 *Dado* $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}[0,1], \lambda)$ *el espacio de medida de Lebesgue sobre los borelianos* $[0,1]$ *con*

$$\mathcal{G}_t = \begin{cases} \sigma(\{[0,\frac{1}{2}]\}) & si\ t \in (0,1) \\ \mathcal{B}([0,1]) & si\ t \in \{0\} \cup [1,\infty), \end{cases}$$

entonces de la definición de $\mathcal{F}_{[0,s]}, \mathcal{F}_{[s,t]}, \mathcal{F}_{[t,\infty)}$ *y de (E14) de la proposición 1.4.8 es claro que* $G = \{\mathcal{G}_t\}_{t\in\mathbb{R}^+}$ *es una familia recíproca. Por otro lado en el caso* $0 < s \leq t < 1$, *se tiene que* $\mathcal{F}_{[0,s]}$, $\mathcal{F}_{[s,t]}$, $\mathcal{F}_{[t,\infty)}$ *no son independientes dado* $\mathcal{F}_{\{s,t\}}$ *entonces del teorema 2.6.5* $G = \{\mathcal{G}\}_{t\in\mathbb{R}^+}$ *no es una familia markoviana.*

Proposición 2.6.8 *$G := \{\mathcal{G}_t\}_{t\in\mathbb{R}^+}$ es una familia recíproca si y sólo si*

$$\begin{array}{ccc} \mathcal{F}_{[s,t]}, \mathcal{F}_{[0,s]} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}} & y & \mathcal{F}_{[s,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}} \vee \mathcal{F}_{[0,s]} \\ & o & \\ \mathcal{F}_{[s,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}} & y & \mathcal{F}_{[s,t]}, \mathcal{F}_{[0,s]} \perp\!\!\!\perp \mathcal{F}_{\{s,t\}} \vee \mathcal{F}_{[t,\infty)}. \end{array}$$

Demostración. Es consecuencia inmediata de la definición de familia recíproca y la proposición 2.3.3.□

Teorema 2.6.9 *Las siguientes afirmaciones son equivalentes:*

(a) *$G := \{\mathcal{G}_t\}_{t\in\mathbb{R}^+}$ es una familia markoviana.*

(b_1) *Para $0 \le s \le t$ se cumple $\mathcal{F}_{[0,s]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s\}}$.*

(b_2) *Para $0 \le s \le t$ y $A \in \mathcal{F}_{[t,\infty)}$ se cumple $P(A|\mathcal{F}_{[0,s]}) = P(A|\mathcal{F}_{\{s\}})$.*

(b_3) *Para $0 \le s \le t$ y X una variable aleatoria integrable $\mathcal{F}_{[t,\infty)}$-medible se cumple $E(X|\mathcal{F}_{[0,s]}) = E(X|\mathcal{F}_{\{s\}})$.*

(c_1) *Para $0 \le s_1 < \cdots < s_n = s \le t = t_1 < \cdots < t_m$ se cumple $\mathcal{F}_{\{s_1,\cdots,s_n\}}, \mathcal{F}_{\{t_1,\cdots,t_m\}} \perp\!\!\!\perp \mathcal{F}_{\{s\}}$.*

(c_2) *Para $0 \le s_1 < \cdots < s_n = s \le t = t_1 < \cdots < t_m$ y $A \in \mathcal{F}_{\{t_1,\cdots,t_m\}}$ se cumple $P(A|\mathcal{F}_{\{s_1,\cdots,s_n\}}) = P(A|\mathcal{F}_{\{s\}})$.*

(c_3) *Para $0 \le s_1 < \cdots < s_n = s \le t = t_1 < \cdots < t_m$ y X una variable aleatoria integrable $\mathcal{F}_{\{t_1,\cdots,t_m\}}$-medible se cumple $E(X|\mathcal{F}_{\{s_1,\cdots,s_n\}}) = E(X|\mathcal{F}_{\{s\}})$.*

(d_1) *Para $0 \le s_1 < \cdots < s_n = s \le t$ se cumple $\mathcal{F}_{\{s_1,\cdots,s_n\}}, \mathcal{F}_{\{t\}} \perp\!\!\!\perp \mathcal{F}_{\{s\}}$.*

(d_2) *Para $0 \le s_1 < \cdots < s_n = s \le t$ y $A \in \mathcal{F}_{\{t_1,\cdots,t_m\}}$ se cumple $P(A|\mathcal{F}_{\{s_1,\cdots,s_n\}}) = P(A|\mathcal{F}_{\{s\}})$.*

(d_3) *Para $0 \le s_1 < \cdots < s_n = s \le t$ y X una variable aleatoria integrable $\mathcal{F}_{\{t\}}$-medible se cumple $E(X|\mathcal{F}_{\{s_1,\cdots,s_n\}}) = E(X|\mathcal{F}_{\{s\}})$.*

Demostración. (a) equivale a (b_1) por la definición de familia markoviana y el teorema 2.6.5. Las equivalencias de (b_1) con (b_2) y con (b_3); de (c_1) con (c_2) y con (c_3); (d_1) con (d_2) y con (d_3) están garantizadas por el teorema 2.2.6. Las implicaciones

$(b_1) \Rightarrow (c_1) \Rightarrow (d_1)$ son consecuencia de la definición de independencia condicionada.

$(c_1) \Rightarrow (b_1)$ Como $\prod_{[0,s]} := \cap_{J\subseteq[o,s]}\mathcal{F}_J$, donde J recorre todos los subconjuntos finitos de $[o,s]$ de la forma $\{s_1,\cdots,s_n\}$ con $0 \leq s_1 < \cdots < s_n = s$, es un π-sistema tal que $\sigma(\prod_{[0,s]}) = \mathcal{F}_{[0,s]}$ y $\prod_{[t,\infty)} := \cap_{I\subseteq[t,\infty)}\mathcal{F}_I$, donde I recorre todos los subconjuntos finitos de $[t,\infty]$ de la forma $\{t_1,\cdots,t_m\}$ con $t = t_1 < \cdots < t_m$, es también un π-sistema tal que $\sigma(\prod_{[t,\infty]}) = \mathcal{F}_{[t,\infty]}$. Entonces de la hipótesis y el corolario 2.2.2 se obtiene $\mathcal{F}_{[0,s]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{s\}}$.

$(d_3) \Rightarrow (c_2)$ Sean $0 \leq s_1 < \cdots < s_n = s \leq t = t_1 < t_2 < \cdots < t_m$ y $F_{t_j} \in \mathcal{F}_{t_j}$, $j = 1,2,\cdots,m$. Como $\prod_{\{t_1,\cdots,t_m\}} = \{\cap_{j=1}^m F_{t_j} : F_{t_j} \in \mathcal{F}_{t_j}; j = 1,\cdots,m\}$ es un π-sistema que contiene a Ω y que genera a $\mathcal{F}_{\{t_1,\cdots,t_m\}}$, entonces es suficiente hacer la demostración para los elementos de $\prod_{\{t_1,\cdots,t_m\}}$

$$\begin{aligned}
P(\cap_{j=1}^m F_{t_j}|\mathcal{F}_{\{s_1,\cdots,s_n\}}) &= E(E(\prod_{j=1}^{m} I_{F_{t_j}}|\mathcal{F}_{\{s_1,\cdots,s_n,t_1,\cdots,t_{m-1}\}})|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E(\prod_{j=1}^{m-1} I_{F_{t_j}} E(I_{F_{t_m}}|\mathcal{F}_{\{s_1,\cdots,s_n,t_1,\cdots,t_{m-1}\}})|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E(\prod_{j=1}^{m-1} I_{F_{t_j}} E(I_{F_{t_m}}|\mathcal{F}_{t_{m-1}})|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E((\prod_{j=1}^{m-2} I_{F_{t_j}})X_{t_{m-1}}|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E(E(\prod_{j=1}^{m-2} I_{F_{t_j}} X_{t_{m-1}}|\mathcal{F}_{\{s_1,\cdots,s_n,t_1,\cdots,t_{m-2}\}})|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E(\prod_{j=1}^{m-2} I_{F_{t_j}} E(X_{t_{m-1}}|\mathcal{F}_{\{s_1,\cdots,s_n,t_1,\cdots,t_{m-2}\}})|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E(\prod_{j=1}^{m-2} I_{F_{t_j}} E(X_{t_{m-1}}|\mathcal{F}_{\{t_{m-2}\}})|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= E((\prod_{j=1}^{m-3} I_{F_{t_j}})X_{t_{m-2}}|\mathcal{F}_{\{s_1,\cdots,s_n\}}) \\
&= \cdots = E(X_{t_1}|\mathcal{F}_{\{s_1,\cdots,s_n\}}) = E(X_{t_1}|\mathcal{F}_{\{s\}}),
\end{aligned}$$

donde es claro que $X_{t_{m-1}} := I_{F_{t_{m-1}}}E(I_{F_{t_m}}|\mathcal{F}_{t_{m-1}})$, $X_{t_{m-2}} := I_{F_{t_{m-2}}}E(X_{t_{m-1}}|\mathcal{F}_{t_{m-2}})$ y $X_{t_1} := I_{F_{t_1}}E(X_{t_2}|\mathcal{F}_{t_1})$ son variables aleatorias $\mathcal{F}_{t_{m-1}}$-medible, $\mathcal{F}_{t_{m-2}}$-medible y $\mathcal{F}_{t_1}$-

medible respectivamente; por lo tanto

$$\begin{aligned} P(\cap_{j=1}^m F_{t_j}|\mathcal{F}_{\{s_1,\cdots,s_n\}}) &= E(X_{t_1}|\mathcal{F}_{\{s\}}) = E(E(X_{t_1}|\mathcal{F}_{\{s\}})|\mathcal{F}_{\{s\}}) \\ &= E(E(\prod_{j=1}^m I_{F_{t_j}}|\mathcal{F}_{\{s_1,\cdots,s_n\}})|\mathcal{F}_{\{s\}}) \\ &= E(\prod_{j=1}^m I_{F_{t_j}}|\mathcal{F}_{\{s\}}) = P(\cap_{j=1}^m F_{t_j}|\mathcal{F}_{\{s\}}).\square \end{aligned}$$

Para el siguiente teorema se tomara en cuenta que para I y T subconjuntos de $\mathbb{R}^+$, $\sigma(\mathcal{F}_I|\mathcal{F}_T)$ representa la σ-álgebra $\sigma(P(A|\mathcal{F}_T) : A \in \mathcal{F}_I)$.

Teorema 2.6.10 *Dada $G = \{\mathcal{G}\}_{t\in\mathbb{R}^+}$ una familia de sub-σ-álgebras de $\mathcal{F}$. Las siguientes afirmaciones son equivalentes:*

(1) G es una familia markoviana.

(2) $\sigma(\mathcal{F}_{[0,t]}|\mathcal{F}_{[t,\infty)}) = \mathcal{F}_{\{t\}}$, $t \in \mathbb{R}^+$.

(3) $\sigma(\mathcal{F}_{[t,\infty)}|\mathcal{F}_{[0,t]}) = \mathcal{F}_{\{t\}}$, $t \in \mathbb{R}^+$.

Demostración. (1) $\Rightarrow$ (2) Por hipótesis para $t \in \mathbb{R}^+$ se tiene $\mathcal{F}_{[0,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{t\}}$, entonces por teorema 2.2.6 para cada $A \in \mathcal{F}_{[0,t]}$ se cumple

$$P(A|\mathcal{F}_{[t,\infty)}) = P(A|\mathcal{F}_{[t,\infty)} \vee \mathcal{F}_{\{t\}}) = P(A|\mathcal{F}_{\{t\}}),$$

entonces $P(A|\mathcal{F}_{[t,\infty)})$ es $\mathcal{F}_{\{t\}}$-medible, luego $\sigma(\mathcal{F}_{[0,t]}|\mathcal{F}_{[t,\infty)}) \subseteq \mathcal{F}_{\{t\}}$. Por otro lado para $A \in \mathcal{F}_{\{t\}}$ se tiene que $A \in \mathcal{F}_{[0,t]}$ y $A \in \mathcal{F}_{[t,\infty)}$ y por lo tanto $I_A = P(A|\mathcal{F}_{[t,\infty)})$, luego $\mathcal{F}_{\{t\}} \subseteq \sigma(\mathcal{F}_{[0,t]}|\mathcal{F}_{[t,\infty)})$.

(2) $\Rightarrow$ (1) Por hipótesis para $A \in \mathcal{F}_{[0,t]}$, $P(A|\mathcal{F}_{[t,\infty)})$ es $\mathcal{F}_t$-medible, entonces por *(E7)* y *(E14)* de la proposición 1.4.8

$$P(A|\mathcal{F}_{\{t\}}) = P(P(A|\mathcal{F}_{[t,\infty)})|\mathcal{F}_{\{t\}}) = P(A|\mathcal{F}_{[t,\infty)}) = P(A|\mathcal{F}_{[t,\infty)} \vee \mathcal{F}_{\{t\}}),$$

lo que por teorema 2.2.6 equivale a $\mathcal{F}_{[0,t]}, \mathcal{F}_{[t,\infty)} \perp\!\!\!\perp \mathcal{F}_{\{t\}}$.

(1) $\Rightarrow$ (3) y (3) $\Rightarrow$ (1) es análogo a lo anterior.$\square$

Capítulo 3

Versiones condicionadas de algunos resultados clásicos de la teoría de probabilidad.

En este Capítulo se presentan versiones condicionadas de algunos resultados clásicos que aparecen en la literatura de la teoría de probabilidad; durante el desarrollo de este apareceran diferentes personajes que han jugado un papel destacado en la teoría de probabilidad y en las matemáticas en general, como Cauchy, Erdös, Kolmogorov, Borel, entre otros. En la Sección 3.1 se presentan algunas desigualdades condicionadas. En la Sección 3.2 se consideran varias versiones condicionadas del Lema de Borel-Cantelli, las cuales aparecen en Majerek et al. [9], Yan [19], Feng et al. [5], Petrov [14]. En la Sección 3.3 se presenta las desigualdades condicionadas de Kolmogorov y la de Hajek-Renyi. Por último la Sección 3.4 se dedica a la ley de los grandes números condicionada.

3.1. Desigualdades Condicionadas

En esta Sección se demostrarán las versiones condicionadas de un par de desigualdades clásicas en la teoría de probabilidad, las cuales son importantes para el desarrollo del resto de este trabajo.

Proposición 3.1.1 (Desigualdad de Cauchy-schwarz condicionada) *Si X y Y son variables aleatorias cuadrado integrables y $\mathcal{G}$ es una sub-σ-álgebra de $\mathcal{F}$. Entonces*

$$(E(XY|\mathcal{G}))^2 \leq E(X^2|\mathcal{G})E(Y^2|\mathcal{G}) \quad (c.t.p.).$$

Demostración. Dado $t \in \mathbb{R}$, se tiene

$$0 \leq E((tX-Y)^2|\mathcal{G}) = t^2E(X^2|\mathcal{G}) - 2tE(XY|\mathcal{G}) + E(Y^2|\mathcal{G}) \quad (c.t.p.),$$

si se toma $a = E(X^2|\mathcal{G})$, $b = -2E(XY|\mathcal{G})$ y $c = E(Y^2|\mathcal{G})$, como $at^2+bt+c = a(t+\frac{b}{2a})^2 + \frac{4ac-b^2}{4a}$ entonces $4ac - b^2 \geq 0$, es decir $4E(X^2|\mathcal{G})E(Y^2|\mathcal{G}) \geq 4(E(XY|\mathcal{G}))^2$ (c.t.p.), obteniendo el resultado.$\square$

Proposición 3.1.2 (Desigualdad de Chung-Erdös condicionada) *Dado un espacio de probabildad* $(\Omega, \mathcal{F}, P)$ *,* $\{A_k\}_{1\leq k\leq n}$ *una sucesión de eventos,* $\mathcal{G}$ *una sub-*σ*-álgebra de* $\mathcal{F}$ *y* $\{X_k\}_{1\leq k\leq n}$ *variables aleatorias* $\mathcal{G}$*-medibles. Entonces*

(a) $P(\cup_{k=1}^n A_k|\mathcal{G}) \geq \dfrac{(\sum_{k=1}^n X_kP(A_k|\mathcal{G}))^2}{\sum_{i,k=1}^n X_iX_kP(A_i\cap A_k|\mathcal{G})} \quad (c.t.p.).$

(b) Para $N \in \{1,\cdots,n\}$*, se tiene en (c.t.p.)*

$$(\sum_{N\neq i=1}^n X_NX_iP(A_N\cap A_i|\mathcal{G}))^2 \leq X_N^2P(A_N|\mathcal{G}) \sum_{N\neq i=1}^n \sum_{N\neq j=1}^n X_iX_jP(A_i\cap A_j|\mathcal{G}).$$

Demostración. (a) Al tomar $X := I_{\cup_{i=1}^n A_i}$ y $Y := \sum_{i=1}^n X_iI_{A_i}$, entonces de las propiedades de esperanza condicionada

$$\begin{aligned} E(XY|\mathcal{G}) &= E(\sum_{i=1}^n X_iI_{A_i}I_{\cup_{i=1}^n A_i}|\mathcal{G}) = \sum_{i=1}^n E(X_iI_{A_i\cap\cup_{i=1}^n A_1}|\mathcal{G}) = \sum_{i=1}^n E(X_iI_{A_i}|\mathcal{G}) \\ &= \sum_{i=1}^n X_iE(I_{A_i}|\mathcal{G}) = \sum_{i=1}^n X_iP(A_i|\mathcal{G}) \quad (c.t.p.). \end{aligned}$$

Como $E(Y^2|\mathcal{G}) = E(\sum_{i,k=1}^n X_iX_kI_{A_i}I_{A_k}|\mathcal{G}) = \sum_{i,k=1}^n X_iX_kP(A_i\cap A_k|\mathcal{G})$ (c.t.p.) y $E(X^2|\mathcal{G}) = E(I_{\cup_{k=1}^n A_k}|\mathcal{G}) = P(\cup_{k=1}^n A_k|\mathcal{G})$ (c.t.p.); al aplicar la desigualdad de Cauchy-shwarz condicionada se tiene

$$(\sum_{i=1}^n X_iP(A_i|\mathcal{G}))^2 \leq P(\cup_{k=1}^n A_k|\mathcal{G}) \sum_{i,k=1}^n X_iX_kP(A_i\cap A_k|\mathcal{G}) \quad (c.t.p.),$$

llegando a lo requerido.

(b) Se procede de la misma forma que en la primera parte; pero tomando $X := X_NI_{A_N}$ y $Y := \sum_{N\neq i=1} X_iI_{A_i}$. $\square$

3.2. Lema de Borel-Cantelli condicionado

El lema 1.1.15 de Borel-Cantelli es uno de los más importantes de la teoría de probabilidad, llamado así en honor de dos eminentes matemáticos Émile Borel y Francesco Paolo Cantelli; quienes lo dieron a conocer en las primeras décadas del siglo XX; este lema juega un papel destacado en la teoría de probabilidad, por ser una herramienta muy usada para demostrar resultados que involucran convergencia casi en toda parte de sucesiones de variables aleatorias. Desde hace más de 50 años aparecen en la literatura trabajos dedicados a la segunda parte de este lema, buscando debilitar la hipótesis de independencia de la sucesión de variables aleatorias o dando una versión lo más general posible, A continuación se mencionarán algunos resultados establecidos por Yan [19], Feng et al. [5], Petrov [14] respectivamente

Teorema 3.2.1 *Sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de eventos en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Si $\sum_{n=1}^{\infty} P(A_n) = \infty$, entonces*

$$P(\limsup A_n) \geq \limsup_{n\in\mathbb{N}} \frac{|\sum_{i=1}^{n} P(A_i)|^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} P(A_i \cap A_j)} = \limsup_{n\in\mathbb{N}} \frac{\sum_{1\leq i<j\leq n} P(A_i)P(A_j)}{\sum_{1\leq i<j\leq n} P(A_i \cap A_j)}.$$

Teorema 3.2.2 *Sean $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de eventos y $\{x_n\}_{n\in\mathbb{N}}$ una sucesión de números reales. Si $\sum_{n=1}^{\infty} x_n P(A_n) = \infty$, entonces*

$$P(\limsup A_n) \geq \limsup_{n\in\mathbb{N}} \frac{|\sum_{i=1}^{n} x_i P(A_i)|^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j P(A_i \cap A_j)}.$$

Obviamente, si $x_n = 1$ para todo $n \in \mathbb{N}$, entonces se obtiene la desigualdad del resultado de Yan 3.2.1.

Teorema 3.2.3 *Sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de eventos en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Si $\sum_{n=1}^{\infty} P(A_n) = \infty$, entonces se verifica para todo $H \in \mathbb{R}$ que*

$$\alpha_H := \liminf_{n\in\mathbb{N}} \frac{\sum_{1\leq i<j\leq n} P(A_i \cap A_j) - HP(A_i)P(A_j)}{[\sum_{i=1}^{n} P(A_i)]^2}$$

implica

$$P(\limsup A_n) \geq \frac{1}{H + 2\alpha_H}.$$

Estos resultados serán demostrados en sus versiones condicionadas, empezando por la versión condicionada del lema de Borel-Cantelli presentada en Majerek et al. [9].

Lema 3.2.4 *Sean $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de eventos de un espacio de probabildad $(\Omega, \mathcal{F}, P)$ y $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$.*

(1) Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, entonces $\sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{G}) < \infty$ (c.t.p.).

(2) Si $A = \{\omega : \sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{G}) < \infty\}$ es tal que $P(A) < 1$, entonces sólo hay una cantidad finita de eventos de la sucesión $\{A_n \cap A\}_{n\in\mathbb{N}}$ que tienen probabilidad uno; es decir, $P(\limsup A_n \cap A) = 0$.

(3) Si $A = \{\omega : \sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{G}) = \infty\}$ y la sucesión de eventos $\{A_n\}_{n\in\mathbb{N}}$ es $\mathcal{G}$-independiente, entonces $P(\limsup A_n) = P(A)$.

Demostración. (1) Las propiedades *(E4)* y *(E7)* de la proposición 1.4.8, el teorema de convergencia monotona y $\sum_{n=1}^{\infty} P(A_n) < \infty$ implican que

$$\begin{aligned}
\sum_{n=1}^{\infty} E(I_{A_n}) &= \sum_{n=1}^{\infty} E(E(I_{A_n}|\mathcal{G})) = \lim_{k\to\infty} \sum_{n=1}^{k} E(E(I_{A_n}|\mathcal{G})) = \lim_{k\to\infty} E(\sum_{n=1}^{k} E(I_{A_n}|\mathcal{G})) \\
&= E(\lim_{k\to\infty} \sum_{n=1}^{k} E(I_{A_n}|\mathcal{G})) = E(\sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{G})),
\end{aligned}$$

obteniendo $\sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{G}) < \infty$ (c.t.p.).

(2) Dado que

$$\begin{aligned}
P(\limsup(A_n \cap A)|\mathcal{G}) &= P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} (A_k \cap A)|\mathcal{G}) = \lim_{n\to\infty} P(\cup_{k=n}^{\infty}(A_k \cap A)|\mathcal{G}) \\
&\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} E(I_{(A_k\cap A)}|\mathcal{G}) = 0 \quad (c.t.p.),
\end{aligned}$$

entonces $P(\limsup(A_n \cap A)) = E[P(\limsup(A_n \cap A)|\mathcal{G})] = 0$.

(3) De las propiedades de esperanza condicionada y como $1 - x \leq e^{-x}$, entonces

$$\begin{aligned}
P(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c|\mathcal{G}) &= \lim_{n\to\infty} P(\cap_{k=n}^{\infty} A_k^c|\mathcal{G}) = \lim_{n\to\infty} (\lim_{k\to\infty} P(\cap_{i=n}^{k} A_i^c|\mathcal{G})) \\
&= \lim_{n\to\infty} (\lim_{k\to\infty} \prod_{i=n}^{k} P(A_i^c|\mathcal{G})) = \lim_{n\to\infty} (\lim_{k\to\infty} \prod_{i=n}^{k} (1 - P(A_i|\mathcal{G}))) \\
&= \lim_{n\to\infty} \prod_{i=n}^{\infty} (1 - P(A_i|\mathcal{G})) \leq \lim_{n\to\infty} e^{-\sum_{i=n}^{\infty} P(A_i|\mathcal{G})} \quad (c.t.p.),
\end{aligned}$$

por tanto para $\omega \in A$, $P(\liminf A_n^c|\mathcal{G})(\omega) \leq \lim_{n\to\infty} e^{-\sum_{i=n}^{\infty} P(A_i|\mathcal{G})(\omega)} = 0$, de lo cual

$$\begin{aligned}
P(\liminf A_n^c) &= \int_{\Omega} P(\liminf A_n^c|\mathcal{G}) dP = \int_{A} P(\liminf A_n^c|\mathcal{G}) dP + \int_{A^c} P(\liminf A_n^c|\mathcal{G}) dP \\
&= \int_{A^c} I_{\liminf A_n^c} dP = P(\liminf A_n^c \cap A^c) \leq P(A^c),
\end{aligned}$$

luego $P(A) \leq P(\limsup A_n)$. Por otro lado de el punto (2) de este teorema, se tiene que $P(\limsup A_k \cap A^c) = 0$ y como $P(\limsup A_k) = P(\limsup A_k \cap A) + P(\limsup A_k \cap A^c)$, entonces $P(\limsup A_k) = P(\limsup A_k \cap A) \leq P(A)$. Obteniendo lo requerido.$\square$

En las siguientes dos versiones condicionadas se busca debilitar la hipótesis de independencia del teorema de Borel-Cantelli. A continuación se mencionará y demostrará un lema que es útil en la demostración de el teorema 3.2.6, el cual es la versión condicionada del que está en Feng et al. [5].

Lema 3.2.5 *Sean $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de eventos, $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias $\mathcal{G}$-medibles y $A := \{\omega : \sum_{n=1}^{\infty} X_n P(A_n|\mathcal{G}) = \infty\}$. Entonces sobre A:*

(a) $\lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} X_i X_j P(A_i \cap A_j|\mathcal{G}) = \infty$ *(c.t.p.).*

(b) Para $m = 1, 2, 3, \cdots$, se cumple (c.t.p.)

$$\limsup_{n\in\mathbb{N}} \frac{\left[\sum_{i=1}^{n} X_i P(A_i|\mathcal{G})\right]^2}{\sum_{i=1}^{n} \sum_{j=1}^{n} X_i X_j P(A_i \cap A_j|\mathcal{G})} = \limsup_{n\in\mathbb{N}} \frac{\left[\sum_{i=m}^{n} X_i P(A_i|\mathcal{G})\right]^2}{\sum_{i=m}^{n} \sum_{j=m}^{n} X_i X_j P(A_i \cap A_j|\mathcal{G})}.$$

Demostración. (a) De la hipótesis y de la parte (a) de la proposición 3.1.2 se tiene que

$$\lim_{n\in\mathbb{N}}\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})\geq\lim_{n\in\mathbb{N}}\frac{\left[\sum_{i=1}^{n}X_iP(A_i|\mathcal{G})\right]^2}{P(\cup_{i=1}^{n}A_i|\mathcal{G})}=\infty\quad(c.t.p.).$$

(b) Lo requerido equivale a demostrar

$$\limsup_{n\in\mathbb{N}}\frac{\left[\sum_{i=m}^{n}X_iP(A_i|\mathcal{G})\right]^2}{\left[\sum_{i=1}^{n}X_iP(A_i|\mathcal{G})\right]^2}=\limsup_{n\in\mathbb{N}}\frac{\sum_{i=m}^{n}\sum_{j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}{\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}\quad(c.t.p.).$$

Como

$$\limsup_{n\in\mathbb{N}}\frac{\left[\sum_{i=m}^{n}X_iP(A_i|\mathcal{G})\right]^2}{\left[\sum_{i=1}^{n}X_iP(A_i|\mathcal{G})\right]^2}=\left[1-\limsup_{n\in\mathbb{N}}\frac{\sum_{i=1}^{m-1}X_iP(A_i|\mathcal{G})}{\sum_{i=1}^{n}X_iP(A_i|\mathcal{G})}\right]^2=1\quad(c.t.p.),$$

sobre A, y como

$$\begin{aligned}&\limsup_{n\in\mathbb{N}}\ \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}{\sum_{i=m}^{n}\sum_{j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}\\&\qquad=1+\limsup_{n\in\mathbb{N}}\frac{2\sum_{i=1}^{m-1}\sum_{i\neq j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})+\sum_{j=1}^{m-1}X_j^2P(A_j|\mathcal{G})}{\sum_{i=m}^{n}\sum_{j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}\quad(c.t.p.),\end{aligned}$$

de la parte (b) de 3.1.2 para cada X_i, $i\in\{1,2,\cdots\}$ fijo y de (a), se tiene:

$$\limsup_{n\in\mathbb{N}}\frac{2\sum_{i=1}^{m-1}\sum_{i\neq j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})+\sum_{j=1}^{m-1}X_j^2P(A_j|\mathcal{G})}{\sum_{i=m}^{n}\sum_{j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}=0\quad(c.t.p.),$$

de lo cual

$$\limsup_{n\in\mathbb{N}}\frac{\sum_{i=m}^{n}\sum_{j=m}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}{\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}=1\quad(c.t.p.),$$

demostrando lo deseado.□

Teorema 3.2.6 *Sean* $\{A_n\}_{n\in\mathbb{N}}$ *una sucesión de eventos en un espacio de probabilidad* $(\Omega,\mathcal{F},P)$ *,* $\mathcal{G}$ *una sub-*σ*-álgebra de* $\mathcal{F}$ *y* $\{X_n\}_{n\in\mathbb{N}}$ *una sucesión de variables aleatorias* $\mathcal{G}$*-medibles. Entonces sobre el evento* $A:=\{\omega:\sum_{n=1}^{\infty}X_nP(A_n|\mathcal{G})=\infty\}$*, se verifica la desigualdad:*

$$P(\limsup A_n|\mathcal{G})\geq\limsup_{n\in\mathbb{N}}\frac{\left[\sum_{i=1}^{n}X_iP(A_i|\mathcal{G})\right]^2}{\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_jP(A_i\cap A_j|\mathcal{G})}\quad(c.t.p.).$$

Demostración. Por la parte (b) del lema anterior, la parte (a) desigualdad de Chung-Erdös condicionada y *(E11)* de la proposición 1.4.8, sobre el evento A se tiene en (c.t.p.):

$$\begin{aligned}
P(\limsup A_n|\mathcal{G}) &= P(\cap_{m=1}^{\infty} \cup_{i=m}^{\infty} A_i|\mathcal{G}) = \lim_{m\to\infty} P(\cup_{i=m}^{\infty} A_i|\mathcal{G}) \\
&= \lim_{m\to\infty} \left[\lim_{n\to\infty} P\left(\cup_{i=m}^{n} A_i|\mathcal{G}\right)\right] \\
&\geq \lim_{m\to\infty} \left[\limsup_{n\in\mathbb{N}} \frac{\left[\sum_{i=m}^{n} X_i P(A_i|\mathcal{G})\right]^2}{\sum_{i=m}^{n}\sum_{j=m}^{n} X_i X_j P(A_i\cap A_j|\mathcal{G})}\right] \\
&= \limsup_{n\in\mathbb{N}} \frac{\left[\sum_{i=1}^{n} X_i P(A_i|\mathcal{G})\right]^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j P(A_i\cap A_j|\mathcal{G})}.\square
\end{aligned}$$

Observación 3.2.7

(1) Si en el teorema anterior se toma $X_n = 1$, para todo $n \in \mathbb{N}$, se obtiene

$$P(\limsup A_n|\mathcal{G}) \geq \limsup_{n\in\mathbb{N}} \frac{\left[\sum_{i=1}^{n} P(A_i|\mathcal{G})\right]^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} P(A_i\cap A_j|\mathcal{G})} \quad (c.t.p.),$$

la cual es la versión condicionada del lema de Borel-Cantelli que aparece en Yan [19].

(2) Si $X_n = 1$ para todo $n \in \mathbb{N}$ y si la sucesión $\{A_n\}_{n\in\mathbb{N}}$ es al menos dos a dos $\mathcal{G}$-independiente (en particular; cuando $\{A_n\}_{n\in\mathbb{N}}$ es una sucesión $\mathcal{G}$-independientes), entonces el teorema anterior establece que (c.t.p.) sobre el evento $A = \{\omega : \sum_{n=1}^{\infty} P(A_n|\mathcal{G}) = \infty\}$ se verifica $P(\limsup A_n|\mathcal{G}) = 1$. De otra parte, (c.t.p.) sobre A^c:

$$P(\limsup A_n|\mathcal{G}) = \lim_{n\to\infty} P(\cup_{m=n}^{\infty} A_m|\mathcal{G}) \leq \lim_{n\to\infty} \sum_{m=n}^{\infty} P(A_n|\mathcal{G}) = 0$$

De esto se sigue que

$$\begin{aligned}
P(\limsup A_n|\mathcal{G}) &= P(\limsup A_n|\mathcal{G}) I_A + P(\limsup A_n|\mathcal{G}) I_{A^c} \\
&= P(\limsup A_n|\mathcal{G}) I_A \\
&= I_A.
\end{aligned}$$

Ahora, la propiedad (E8) de la proposición 1.4.8 conduce a la igualdad:

$$P(\limsup A_n) = E[P(\limsup A_n|\mathcal{G})] = E[I_A] = P(A),$$

que es la parte (3) del lema 3.2.4.

Ejemplo 3.2.8 *Sea* $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \lambda)$*, donde* $\mathcal{B}([0,1])$ *es la* σ*-álgebra de los conjuntos borelianos de* $[0,1]$ *y* λ *es la medida de Lebesgue. Sea* $\mathcal{G}$ *la* σ*-álgebra generada por la descomposición* $\mathcal{D}$ *de* $[0,1]$ *dada por*

$$\mathcal{D} = \{[0,1/4), [1/4,1/2), [1/2,3/4), [3/4,1]\}$$

para los eventos $E_1 := [1/4, 5/8)$ *y* $E_2 := [3/8, 3/4)$ *se verifican:*

$$P(E_1|\mathcal{G}) = I_{[1/4,1/2)} + \frac{1}{2} I_{[1/2,3/4)}, \quad P(E_2|\mathcal{G}) = \frac{1}{2} I_{[1/4,1/2)} + I_{[1/2,3/4)},$$

$$P(E_1 \cap E_2|\mathcal{G}) = \frac{1}{2} I_{[1/4,3/4)} \quad y \quad P(E_1 \cup E_2|\mathcal{G}) = I_{[1/4,3/4)}.$$

Considere ahora la sucesión de eventos $\{A_n\}_{n\in\mathbb{N}}$ *descrita por:*

$$E_1, E_2, E_1 \cap E_2, E_1, E_2, E_1 \cap E_2, E_1, E_2, E_1 \cap E_2, \cdots$$

y la sucesión de variables aleatorias $\{X_n\}_{n\in\mathbb{N}}$ *dada por:*

$$1, 1, -1, 1, 1, -1, 1, 1, -1, \cdots$$

Por el teorema 3.2.6, c.t.p. Sobre $A = \{\omega : \sum_{n=1}^{\infty} X_n P(A_n|\mathcal{G}) = \infty\} = [1/4, 3/4)$ *se verifica la desigualdad*

$$P(\limsup A_n|\mathcal{G}) \geq P(E_1|\mathcal{G}) + P(E_2|\mathcal{G}) - P(E_1 \cap E_2|\mathcal{G}) = P(E_1 \cup E_2|\mathcal{G}).$$

De hecho, $P(\limsup A_n|\mathcal{G}) = P(E_1 \cup E_2|\mathcal{G})$*.*

El siguiente lema es la forma condicionada del lema que está en Petrov [14].

Lema 3.2.9 *Sean* $\{A_n\}_{n\in\mathbb{N}}$ *una sucesión de eventos,* $\mathcal{G}$ *una sub-*σ*-álgebra de* $\mathcal{F}$ *y* $A = \{\omega : \sum_{k=1}^{\infty} P(A_k|\mathcal{G}) = \infty\}$*, con*

$$\beta_H^{(m)} = \liminf_{n\in\mathbb{N}} \frac{\sum_{m\leq i<k\leq n} (P(A_i \cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2},$$

Para todo entero $m \geq 1$*. Entonces sobre el evento* A*, para todo* $H \in \mathbb{R}$ *fijo y todo* m *positivo, se tiene*

$$\beta_H^{(m)} = \liminf_{n\in\mathbb{N}} \frac{\sum_{1\leq i<k\leq n} (P(A_i \cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2} \quad (c.t.p.). \tag{3.1}$$

Demostración. Como $\sum_{k=m}^{n} P(A_k|\mathcal{G}) = \sum_{k=1}^{n} P(A_k|\mathcal{G}) - \sum_{k=1}^{m-1} P(A_k|\mathcal{G})$, se tiene que

$$\lim_{n\to\infty} \frac{\sum_{k=m}^{n} P(A_k|\mathcal{G})(\omega)}{\sum_{k=1}^{n} P(A_k|\mathcal{G})(\omega)} = 1,$$

para todo $\omega \in A$, entonces $\sum_{k=m}^{n} P(A_k|\mathcal{G})(\omega) = \sum_{k=1}^{n} P(A_k|\mathcal{G})(\omega)$ cuando $n \to \infty$ y $\omega \in A$; por tanto la parte derecha de la igualdad (3.1) también se escribe como

$$W = \liminf_{n\in\mathbb{N}} \frac{\sum_{1\leq i<k\leq n}(P(A_i \cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2},$$

sobre el conjunto A. Como para a_{ik} números reales, con $1 \leq i < k \leq n$ y $m < n$ todos números enteros positivos, se tiene

$$\sum_{1\leq i<k\leq n} a_{ik} = \sum_{1\leq i<k\leq m} a_{ik} + \sum_{m\leq i<k\leq n} a_{ik} + \sum_{1\leq i<m<k\leq n} a_{ik},$$

entonces de todo lo anterior

$$\begin{aligned} W \quad = \quad & \liminf_{n\in\mathbb{N}} \frac{\sum_{1\leq i<k\leq m}(P(A_i \cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} \\ & + \liminf_{n\in\mathbb{N}} \frac{\sum_{m\leq i<k\leq n}(P(A_i \cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} \\ & + \liminf_{n\in\mathbb{N}} \frac{\sum_{1\leq i<m<k\leq n}(P(A_i \cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2}. \end{aligned}$$

Como para todo $\omega \in A$ se tiene $P(A_i \cap A_k|\mathcal{G})(\omega) \leq P(A_k|\mathcal{G})(\omega) \leq 1$, entonces

$$\begin{aligned} & \left| \frac{\sum_{1\leq i<m<k\leq n}(P(A_i \cap A_k|\mathcal{G})(\omega) - HP(A_i|\mathcal{G})(\omega)P(A_k|\mathcal{G})(\omega))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G})(\omega))^2} \right| \\ & \leq \frac{m\sum_{m<k\leq n}(P(A_k|\mathcal{G})(\omega) + |H|P(A_k|\mathcal{G})(\omega))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G})(\omega))^2}, \end{aligned}$$

de lo cual

$$\lim_{n\to\infty} \frac{\sum_{1\leq i<m<k\leq n}(P(A_i \cap A_k|\mathcal{G})(\omega) - HP(A_i|\mathcal{G})(\omega)P(A_k|\mathcal{G})(\omega))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G})(\omega))^2} = 0,$$

y como también para $\omega \in A$ se tiene que

$$\lim_{n\to\infty} \frac{\sum_{1\leq i<k\leq m}(P(A_i\cap A_k|\mathcal{G})(\omega) - HP(A_i|\mathcal{G})(\omega)P(A_k|\mathcal{G})(\omega))}{(\sum_{k=m}^{n} P(A_k|\mathcal{G})(\omega))^2} = 0,$$

se concluye lo deseado.$\square$

La siguiente es la versión condicionada del lema de Borel-Cantelli formulada en Petrov [14] y presentada en Prakasa [15].

Teorema 3.2.10 *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad; $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$; $A_1, A_2, \cdots$ una sucesión de eventos; un conjunto $A = \{\omega : \sum_{k=1}^{\infty} P(A_k|\mathcal{G}) = \infty\}$ y H una función $\mathcal{G}$-medible. Si se toma*

$$\alpha_H = \liminf_{n\in\mathbb{N}} \frac{\sum_{1\leq i<k\leq n}(P(A_i\cap A_k|\mathcal{G}) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G}))}{(\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2},$$

entonces sobre el conjunto A se tiene

$$P(\limsup A_n|\mathcal{G}) \geq \frac{1}{H+2\alpha_H} \quad (c.t.p.).$$

Demostración. Como

$$\sum_{i,k=m}^{n} P(A_i\cap A_k|\mathcal{G}) = \sum_{k=m}^{n} P(A_k|\mathcal{G}) + 2\sum_{m\leq i<k\leq n} P(A_i\cap A_k|\mathcal{G}) = \sum_{k=m}^{n} P(A_k|\mathcal{G}) + T_1 + T_2,$$

donde

$$T_1 = 2\sum_{m\leq i<k\leq n}(P(A_i\cap A_k) - HP(A_i|\mathcal{G})P(A_k|\mathcal{G})) \quad y \quad T_2 = 2H\sum_{m\leq i<k\leq n} P(A_i|\mathcal{G})P(A_k|\mathcal{G})$$

entonces al aplicar la desigualdad de Chung-Erdös para $\{X_k = 1\}_{1\leq k\leq n}$ se tiene (c.t.p.)

$$\begin{aligned} P(\cup_{k=m}^{n} A_k|\mathcal{G}) &\geq \frac{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2}{\sum_{i,k=m}^{n} P(A_i\cap A_k|\mathcal{G})} = \left[\frac{\sum_{i,k=m}^{n} P(A_i\cap A_k|\mathcal{G})}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2}\right]^{-1} \\ &= \left[\frac{1}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} + \frac{T_1}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} + \frac{T_2}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2}\right]^{-1}. \end{aligned}$$

Como $\sum_{k=m}^{n}(P(A_k|\mathcal{G}))^2 \leq \sum_{k=m}^{n} P(A_k|\mathcal{G})$ (c.t.p.) esto implica que para todo $\omega \in A$ se tenga

$$\lim_{n\to\infty} \frac{\sum_{k=m}^{n}(P(A_k|\mathcal{G}))^2}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} \leq \lim_{n\to\infty} \frac{1}{\sum_{k=m}^{n} P(A_k|\mathcal{G})} = 0$$

y dado que $(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2 = \sum_{k=m}^{n}(P(A_k|\mathcal{G}))^2 + 2\sum_{m\leq i<k\leq n} P(A_k|\mathcal{G})P(A_i|\mathcal{G})$ (c.t.p.), entonces sobre el conjunto A se tiene

$$\lim_{n\to\infty} \frac{T_2}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} = \lim_{n\to\infty} \frac{H[(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2 - \sum_{k=m}^{n}(P(A_k|\mathcal{G}))^2]}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2} = H,$$

de lo cual, sobre A

$$P(\cup_{k=m}^{\infty} A_k|\mathcal{G}) = \liminf_{n\in\mathbb{N}} P(\cup_{k=m}^{n} A_k|\mathcal{G}) \geq \left[\liminf_{n\in\mathbb{N}} \left(\frac{T_1}{(\sum_{k=m}^{n} P(A_k|\mathcal{G}))^2}\right) + H\right]^{-1} (c.t.p.),$$

luego del lema anterior, para todo $m \geq 1$ se tiene

$$P(\cup_{k=m}^{\infty} A_k|\mathcal{G}) \geq \left[2\beta_H^{(m)} + H\right]^{-1} = [2\alpha_H + H]^{-1} \quad (c.t.p.);$$

por tanto sobre el conjunto A se cumple

$$P(\limsup A_n|\mathcal{G}) = P(\cap_{m=1}^{\infty} \cup_{k=m}^{\infty} A_k|\mathcal{G}) = \lim_{m\to\infty} P(\cup_{k=m}^{\infty} A_k|\mathcal{G}) \geq \frac{1}{2\alpha_H + H} \quad (c.t.p.).\square$$

Observación 3.2.11 *Dado que*

$$2\sum_{1\leq i<k\leq n} P(A_i|\mathcal{G})P(A_k|\mathcal{G}) = \left(\sum_{k=1}^{n} P(A_k|\mathcal{G})\right)^2 - \sum_{k=1}^{n}(P(A_k|\mathcal{G}))^2,$$

$$2\sum_{1\leq i<k\leq n} P(A_i \cap A_k|\mathcal{G}) = \sum_{i,k=1}^{n} P(A_i \cap A_k|\mathcal{G}) - \sum_{k=1}^{n} P(A_k|\mathcal{G})$$

y $(P(A_k|\mathcal{G}))^2 \leq P(A_k|\mathcal{G})$ *(c.t.p.) para todo* $1 \leq k \leq n$*. Entonces*

$$2\alpha_H = \liminf_{n\in\mathbb{N}} \left[\frac{\sum_{i,k=1}^{n} P(A_i \cap A_k|\mathcal{G})}{(\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2} - \frac{1}{\sum_{k=1}^{n} P(A_k|\mathcal{G})} - H + \frac{H\sum_{k=1}^{n}(P(A_k|\mathcal{G}))^2}{(\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2}\right],$$

además sobre $A = \{\omega : \sum_{k=1}^{\infty} P(A_k|\mathcal{G}) = \infty\}$*, se tiene*

$$2\alpha_H = \liminf_{n\in\mathbb{N}} \left[\frac{\sum_{i,k=1}^{n} P(A_i \cap A_k|\mathcal{G})}{(\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2} \right] - H.$$

De la desigualdad de Cauchy-schwarz condicionada se tiene que

$$\begin{aligned}
(\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2 &= (E(\sum_{k=1}^{n} I_{A_k}|\mathcal{G}))^2 \leq E((\sum_{k=1}^{n} I_{A_k})^2|\mathcal{G}) \\
&= E(\sum_{i,k=1}^{n} I_{A_k} I_{A_i}|\mathcal{G}) = \sum_{i,k=1}^{n} P(A_k \cap A_i|\mathcal{G}) \quad (c.t.p.),
\end{aligned}$$

entonces

$$\frac{\sum_{i,k=1}^{n} P(A_k \cap A_i|\mathcal{G})}{\sum_{k=1}^{n} P(A_k|\mathcal{G}))^2} \geq 1 \quad (c.t.p.).$$

Por tanto de todo lo anterior, sobre el conjunto A se cumple que $2\alpha_H \geq 1 - H$, lo cual equivale a que $2\alpha_H + H \geq 1$.

3.3. Desigualdad generalizada de Kolmogorov

Andréi Nikolayevich Kolmogorov matemático ruso que hizo grandes e importantes aportes en los campos de la teoría de la probabilidad y de la topología. Particularmente, estructuró el sistema axiomático de la teoría de probabilidad a partir de la teoría de conjuntos. Entre los primeros trabajos probabilisticos de Kolmogorov estuvo el publicado en 1924 junto a Khinchin sobre la convergencia de series cuyos terminos dependen del azar, aquí aparece la conocida y potente desigualdad de Kolmogorov, la cual sera mencionada más adelante y tiene su demostración en el Capítulo IV Sección 2 del libro de Shiryaev [16]. En este trabajo, la versión condicionada de la desigualdad de Kolmogorov es una herramienta importante en la demostración de la desigualdad de Hájek-Renyi condicionada y para demostrar la ley de los grandes números de Cantelli condicionada.

Teorema 3.3.1 (Desigualdad de Kolmogorov)

(a) Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes tal que para todo $i \leq n$ $E(X_i) = 0$, $E(X_i^2) < \infty$ y $S_k = X_1 + \cdots + X_k$. Entonces para todo $\epsilon > 0$

$$P\left(\max_{1\leq k\leq n} |S_k| \geq \epsilon\right) \leq \frac{E(S_n^2)}{\epsilon^2}.$$

(b) Si también $P(|X_i| \leq c) = 1$*, para todo* $i \leq n$*, entonces*

$$P\left(\max_{1\leq k\leq n} |S_k| \geq \epsilon\right) \geq 1 - \frac{(c+\epsilon)^2}{E(S_n^2)}.$$

Para la versión condicionada de la desigualdad de Kolmogorov se necesitara del corolario que viene a continuación del siguiente lema, el cual es conocido como lema de reemplazo.

Lema 3.3.2 *Sean* $\mathcal{G}$*,* $\mathcal{F}_1$ *dos sub-*σ*-álgebras de* $\mathcal{F}$*;* X*,* Y *dos variables aleatorias tal que* $\sigma(X) \subseteq \mathcal{F}_1, Y \perp\!\!\!\perp \mathcal{G}$ *y* $f : \mathbb{R}^2 \to \mathbb{R}$ *una función boreliana para la cual se cumple* $E[|f(X,Y)|] < \infty$*, entonces*

$$E[f(X,Y)|\mathcal{F}_1 \vee \mathcal{G}] = m(X),$$

donde $m(x) = E[f(x,Y)|\mathcal{G}]$*. Esto es*

$$E[f(X,Y)|\mathcal{F}_1 \vee \mathcal{G}](\omega_0) = E[f(x,Y)|\mathcal{G}](\omega_0),$$

donde $X(\omega_0) = x$*.*

Demostración. Del teorema 1.4.23 para las varibles aleatorias X y Y existen $F_X^{\mathcal{F}_1\vee\mathcal{G}}$ y $F_Y^{\mathcal{F}_1\vee\mathcal{G}}$ funciones de distribución condicional regular respectivamente; además para el vector aleatorio $Z = (X,Y)$ existe una distribución condicional regular $F_Z^{\mathcal{F}_1\vee\mathcal{G}}$, tal que del teorema de Fubini y el teorema 1.4.17, se tiene

$$\begin{aligned} E[f(X,Y)|\mathcal{F}_1 \vee \mathcal{G}](\omega_0) &= \int_{\mathbb{R}^2} f(x,y)dF_Z^{\mathcal{F}_1\vee\mathcal{G}}(\omega_0,x,y) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)dF_X^{\mathcal{F}_1\vee\mathcal{G}}(\omega_0,x)dF_Y^{\mathcal{F}_1\vee\mathcal{G}}(\omega_0,y). \end{aligned}$$

Dado que $F_X^{\mathcal{F}_1\vee\mathcal{G}}(\omega,x) = P(X \leq x|\mathcal{F}_1 \vee \mathcal{G})(\omega) = E(I_{X\leq x}|\mathcal{F}_1 \vee \mathcal{G})(\omega)$ (c.t.p.) y que $\sigma(X) \subseteq \mathcal{F}_1$, entonces $\sigma(X) \subseteq \mathcal{F}_1 \vee \mathcal{G}$ y $F_X^{\mathcal{F}_1\vee\mathcal{G}}(\omega,x) = I_{\{X\leq x\}}(\omega)$ (c.t.p.). Por otra parte como $\mathcal{F}_1, Y \perp\!\!\!\perp \mathcal{G}$ entonces por el teorema 2.2.6

$$F_Y^{\mathcal{F}_1\vee\mathcal{G}}(\omega,y) = P(Y \leq y|\mathcal{F}_1 \vee \mathcal{G})(\omega) = P(Y \leq y|\mathcal{G})(\omega) = F_Y^{\mathcal{G}}(\omega;y) \quad (c.t.p.),$$

con $F_Y^{\mathcal{G}}(\omega; y)$ una función de distribución condicional regular para Y dada $\mathcal{G}$. Luego

$$\begin{aligned} E[f(X,Y)|\mathcal{F}_1 \vee \mathcal{G}](\omega_0) &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y) dI_{\{X \leq x\}}(\omega) dF_Y^{\mathcal{G}}(\omega_0; y) \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x,y) dF_Y^{\mathcal{G}}(\omega_0; y) \right] dI_{\{X \leq x\}}(\omega) \\ &= \int_{\mathbb{R}} E[f(x,Y)|\mathcal{G}](\omega_0) dI_{\{X \leq x\}}(\omega_0) = E[f(x,Y)|\mathcal{G}](\omega_0), \end{aligned}$$

donde $x = X(\omega_0)$.□

Corolario 3.3.3 *Sean X, Y dos variables aleatorias, $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$, $A \in \sigma(X)$ y $r \geq 1$. Si $X, Y \perp\!\!\!\perp \mathcal{G}$ y $E[|X - E(x|\mathcal{G}) + Y - E(Y|\mathcal{G})|^r] < \infty$, entonces*

$$E[|X - E(X|\mathcal{G}) + Y - E(Y|\mathcal{G})|^r I_A|\mathcal{G}] \geq E[|X - E(X|\mathcal{G})|^r I_A|\mathcal{G}] \quad (c.t.p.).$$

Demostración. Dado que $\sigma(X) = \{A \in \mathcal{F} : X^{-1}(B) = A, \quad B \in \mathcal{B}(\mathbb{R})\}$; $X, Y \perp\!\!\!\perp \mathcal{G}$; tomando $X = X(\omega)$, $E(X(\omega)|\mathcal{G}) = E(x|\mathcal{G})$ y utilizando el lema anterior se tiene:

$$\begin{aligned} E[|X - E(X|\mathcal{G}) + Y - &E(Y|\mathcal{G})|^r I_A|\sigma(X) \vee \mathcal{G}](\omega_0) \\ &= E[|x - E(x|\mathcal{G}) + Y - E(Y|\mathcal{G})|^r I_B(x)|\mathcal{G}](\omega_0) \\ &\geq |E[(x - E(x|\mathcal{G}) + Y - E(Y|\mathcal{G}))I_B(x)|\mathcal{G}]|^r(\omega_0) \\ &= |(x - E(x|\mathcal{G}))I_B(x) + E[(Y - E(Y|\mathcal{G}))I_B(x)|\mathcal{G}]|^r(\omega_0) \\ &= |(x - E(x|\mathcal{G}))I_B(x) + E[Y - E(Y|\mathcal{G})|\mathcal{G}]E[I_B(x)|\mathcal{G}]|^r(\omega_0) \\ &= |(x - E(x|\mathcal{G}))I_B(x)|^r(\omega_0) = |(X - E(X|\mathcal{G}))I_A|^r(\omega_0). \end{aligned}$$

Entonces

$$E[E[|X - E(X|\mathcal{G}) + Y - E(Y|\mathcal{G})|^r I_A|\sigma(X) \vee \mathcal{G}]|\mathcal{G}] \geq E[|(X - E(X|\mathcal{G}))I_A|^r|\mathcal{G}] \quad (c.t.p.),$$

que equivale a

$$E[|X - E(X|\mathcal{G}) + Y - E(Y|\mathcal{G})|^r I_A|\mathcal{G}] \geq E[|(X - E(X|\mathcal{G}))I_A|^r|\mathcal{G}] \quad (c.t.p.).\square$$

Teorema 3.3.4 (Desigualdad generalizada de Kolmogorov) *Sean $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$ y $\{X_k\}_{1 \leq k \leq n}$ un conjunto de variables aleatorias $\mathcal{G}$-independientes, donde $E[|X_k|^r|\mathcal{G}] < \infty$ (c.t.p.) para todo $1 \leq k \leq n$ y algún $r \geq 1$. Entonces para una variable aleatoria $\varepsilon > 0$ (c.t.p.) $\mathcal{G}$-medible, $S_k = \sum_{i=1}^k X_i$ y el conjunto $C = \{\max_{1 \leq k \leq n} |S_k - E(S_k|\mathcal{G})| \geq \varepsilon\}$, se tiene*

$$\varepsilon^r P(C|\mathcal{G}) \leq E[|S_n - E(S_n|\mathcal{G})|^r I_C|\mathcal{G}] \leq E[|S_n - E(S_n|\mathcal{G})|^r|\mathcal{G}] \quad (c.t.p.).$$

Demostración. Como $X_1, \cdots, X_n \perp\!\!\!\perp \mathcal{G}$ entonces $S_k = \sum_{i=1}^{k} X_i$ y $R_k = \sum_{i=k}^{n} X_i$ son $\mathcal{G}$-independientes. Además si se toma $\mathcal{T} = \text{ínf}\{k : |S_k - E(S_k|\mathcal{G})| \geq \varepsilon, 1 \leq k \leq n\}$ y $A_k = C \cap \{\mathcal{T} = k\}$, entonces del corolario anterior

$$\begin{aligned} E[|S_n - E(S_n|\mathcal{G})|^r I_{A_k}|\mathcal{G}] &= E[|S_k + R_k - E(S_k|\mathcal{G}) - E(R_k|\mathcal{G})|^r I_{A_k}|\mathcal{G}] \\ &= E[|S_k - E(S_k|\mathcal{G}) + R_k - E(R_k|\mathcal{G})|^r I_{A_k}|\mathcal{G}] \\ &\geq E[|S_k - E(S_k|\mathcal{G})|^r I_{A_k}|\mathcal{G}] \\ &\geq E(\varepsilon^r I_{A_k}|\mathcal{G}) = \varepsilon^r P(A_k|\mathcal{G}) \quad (c.t.p.), \end{aligned}$$

por lo tanto

$$\begin{aligned} E[|S_n - E(S_n|\mathcal{G})|^r I_C|\mathcal{G}] &= \sum_{k=1}^{n} E[|S_n - E(S_n|\mathcal{G})|^r I_{A_k}|\mathcal{G}] \geq \sum_{k=1}^{n} \varepsilon^r P(A_k|\mathcal{G}) \\ &= \varepsilon^r P(\uplus_{k=1}^{n} A_k|\mathcal{G}) = \varepsilon^r P(C|\mathcal{G}) \quad (c.t.p.) \end{aligned}$$

y como $|S_n - E(S_n|\mathcal{G})|^r \geq |S_n - E(S_n|\mathcal{G})|^r I_C$ (c.t.p.), entonces

$$\varepsilon^r P(C|\mathcal{G}) \leq E[|S_n - E(S_n|\mathcal{G})|^r I_C|\mathcal{G}] \leq E[|S_n - E(S_n|\mathcal{G})|^r|\mathcal{G}] \quad (c.t.p.).\square$$

A continuación se enuncia la desigualdad de Hájek-Rényi (ver demostración Sección 6.3 del libro de Bauer [1]) y despues su versión condicionada en la cual su demostración es un objetivo de esta Sección.

Teorema 3.3.5 (Desigualdad de Hájek-Rényi) *Sean $X_1, \cdots, X_n$ variables aleatorias integrables e independientes, $\gamma_1 \geq \cdots \geq \gamma_n > 0$ números reales y $S_i = X_1 - E(X_1) + \cdots + X_i - E(X_i)$, $i = 1, \cdots, n$. Entonces para todo $m = 1, \cdots, n$ y todo $\epsilon > 0$*

$$P\left\{\sup_{m \leq i \leq n} \gamma_i |S_i| \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2}\left[\gamma_m^2 \sum_{i=1}^{m} Var(X_i) + \sum_{i=m+1}^{n} \gamma_i^2 Var(X_i)\right].$$

Teorema 3.3.6 (Desigualdad de Hájek-Rényi condicionada) *Sea $\{X_k\}_{1 \leq k \leq m}$ una familia de variables aleatorias $\mathcal{G}$-independientes con $E(X_k^2|\mathcal{G}) < \infty$; una variable aleatoria $\varepsilon > 0$ (c.t.p.) $\mathcal{G}$-medible; $S_k = \sum_{i=1}^{k} X_i$, $k \geq 1$ y*

$$C = \{\max_{n \leq k \leq m} c_k |S_k - E(S_k|\mathcal{G})| \geq \varepsilon\},$$

donde $\{c_k\}_{1\leq k\leq m}$ *es una sucesión no decreciente de variables aleatorias positivas (c.t.p.). Entonces para todo* $1 \leq n \leq m$ *se tiene*

$$\varepsilon^2 P(C|\mathcal{G}) \leq c_n^2 \sum_{k=1}^{n} E[X_k - E(X_k|\mathcal{G})|\mathcal{G}]^2 + \sum_{n+1}^{m} c_k^2 E[X_k - E(X_k|\mathcal{G})|\mathcal{G}]^2 \quad (c.t.p.).$$

Demostración. Si $E(X_k|\mathcal{G}) = 0,\ 1 \leq k \leq m;\ \varepsilon > 0$ (c.t.p.) es una variable aleatoria $\mathcal{G}$-medible; $A := \{\max_{n\leq k\leq m} c_k S_k \geq \varepsilon\}$ y $\mathcal{T} := \inf\{k : |S_k| \geq \varepsilon, 1 \leq k \leq n\}$, entonces para los conjuntos $A_r = A \cap \{\mathcal{T} = r\}$, $Z = \sum_{k=n}^{m-1} S_k^2(c_k^2 - c_{k+1}^2) + c_m^2 S_m^2$ y la $\mathcal{G}$-independencia de los $\{X_k\}_{1\leq k\leq m}$ se tiene:

$$\begin{aligned}
E(Z|\mathcal{G}) &= \sum_{k=n}^{m-1}(c_k^2 - c_{k+1}^2)E(S_k^2|\mathcal{G}) + c_m^2 E(S_m^2|\mathcal{G}) \\
&= \sum_{k=n}^{m-1}(c_k^2 - c_{k+1}^2)E(\sum_{i,j=1}^{k} X_i X_j|\mathcal{G}) + c_m^2 E(\sum_{i,j=1}^{m} X_i X_j|\mathcal{G}) \\
&= \sum_{k=n}^{m-1}(c_k^2 - c_{k+1}^2)E(\sum_{i=1}^{k} X_i^2 + 2\sum_{1\leq i<j\leq k} X_i X_j|\mathcal{G}) \\
&\quad + c_m^2 E(\sum_{i=1}^{m} X_i^2 + 2\sum_{1\leq i<j\leq m} X_i X_j|\mathcal{G}) \\
&= \sum_{k=n}^{m-1}(c_k^2 - c_{k+1}^2)E(\sum_{i=1}^{k} X_i^2|\mathcal{G}) + \sum_{i=1}^{m} c_m^2 E(X_i^2|\mathcal{G}) \\
&= \sum_{k=n}^{m-1} c_k^2 \sum_{i=1}^{k} E(X_i^2|\mathcal{G}) - \sum_{k=n}^{m-1} c_{k+1}^2 \sum_{i=1}^{k} E(X_i^2|\mathcal{G}) + \sum_{i=1}^{m} c_m^2 E(X_i^2|\mathcal{G}) \\
&= c_n^2 \sum_{i=1}^{n} E(X_i^2|\mathcal{G}) + \sum_{k=n+1}^{m-1} c_k^2 \sum_{i=1}^{k} E(X_i^2|\mathcal{G}) \\
&\quad - \sum_{k=n}^{m-1} c_{k+1}^2 \sum_{i=1}^{k} E(X_i^2|\mathcal{G}) + \sum_{i=1}^{m} c_m^2 E(X_i^2|\mathcal{G}) \\
&= c_n^2 \sum_{i=1}^{n} E(X_i^2|\mathcal{G}) + \sum_{k=n}^{m-1} c_{k+1}^2 \sum_{i=1}^{k+1} E(X_i^2|\mathcal{G}) - \sum_{k=n}^{m-1} c_{k+1}^2 \sum_{i=1}^{k} E(X_i^2\mathcal{G}) \\
&= c_n^2 \sum_{i=1}^{n} E(X_i^2|\mathcal{G}) + \sum_{k=n}^{m-1} c_{k+1}^2 \left[\sum_{i=1}^{k} E(X_i^2|\mathcal{G}) + E(X_{k+1}^2|\mathcal{G})\right] \\
&\quad - \sum_{k=n}^{m-1} c_{k+1}^2 \sum_{i=1}^{k} E(X_i^2\mathcal{G})
\end{aligned}$$

$$\begin{aligned}
&= c_n^2 \sum_{i=1}^{n} E(X_i^2|\mathcal{G}) + \sum_{k=n}^{m-1} c_{k+1}^2 E(X_{k+1}^2|\mathcal{G}) \\
&= c_n^2 \sum_{i=1}^{n} E(X_i^2|\mathcal{G}) + \sum_{k=n+1}^{m} c_k^2 E(X_k^2|\mathcal{G}) \quad (c.t.p).
\end{aligned}$$

Por otro lado como se cumple que $E(Z|\mathcal{G}) \geq E(ZI_A|\mathcal{G}) = \sum_{r=n}^{m} E(ZI_{A_r}|\mathcal{G})$ (c.t.p.), $E(ZI_{A_r}|\mathcal{G}) = \sum_{k=n}^{m-1}(c_k^2 - c_{k+1}^2)E(S_k^2 I_{A_r}|\mathcal{G}) + c_m^2 E(S_m^2 I_{A_r}|\mathcal{G})$ (c.t.p.) y como de la desigualdad de Kolmogorov generalizada, para $r \leq k \leq m$, se tiene

$$E(S_k^2 I_{A_r}|\mathcal{G}) \geq E(S_r^2 I_{A_r}|\mathcal{G}) \geq \frac{\varepsilon^2}{c_r^2} P(A_r|\mathcal{G}) \quad (c.t.p.).$$

Entonces se concluye

$$\begin{aligned}
E(ZI_{A_r}|\mathcal{G}) &\geq \sum_{k=n}^{m-1}(c_k^2 - c_{k+1}^2)\frac{\varepsilon^2}{c_r^2}P(A_r|\mathcal{G}) + c_m^2\frac{\varepsilon^2}{c_r^2}P(A_r|\mathcal{G}) \\
&= \left[c_n^2\frac{\varepsilon^2}{c_r^2}P(A_r|\mathcal{G}) - c_m^2\frac{\varepsilon^2}{c_r^2}P(A_r|\mathcal{G})\right] + c_n^2\frac{\varepsilon^2}{c_r^2}P(A_r|\mathcal{G}) \\
&= c_n^2\frac{\varepsilon^2}{c_r^2}P(A_r|\mathcal{G}) \geq \varepsilon^2 P(A_r|\mathcal{G}) \quad (c.t.p.),
\end{aligned}$$

pues de la hipótesis para $r \leq k \leq m$ se tiene que $c_r \leq c_k \leq c_m$. Por tanto $E(Z|\mathcal{G}) \geq \sum_{r=n}^{m} \varepsilon^2 P(A_r|\mathcal{G}) = \varepsilon^2 P(\uplus_{r=n}^{m} A_r|\mathcal{G}) = \varepsilon^2 P(A|\mathcal{G})$ (c.t.p.), demostrando lo requerido.□

Corolario 3.3.7 *Dadas $\{X_k\}_{k\geq 1}$ una sucesión de variables aleatorias $\mathcal{G}$-independientes con $E(X_k^2|\mathcal{G}) < \infty$; una variable aleatoria $\varepsilon > 0$ (c.t.p.) $\mathcal{G}$-medible, $S_k = \sum_{i=1}^{k} X_i$, $k \geq 1$ entonces se tiene*

$$\varepsilon^2 P(\max_{1\leq k\leq m} |S_k - E(S_k|\mathcal{G}) \geq \varepsilon|\mathcal{G}) \leq \sum_{k=1}^{m} E[X_k - E(X_k|\mathcal{G})|\mathcal{G}]^2 \quad (c.t.p.).$$

Demostración. Es solo tomar en el teorema anterior $n = 1$ y $c_1 = c_2 = \cdots = c_m = 1$.□

Observación 3.3.8 *Al tomar $m \to \infty$ en la desigualdad de Hájek-Rényi condicionada y cambiar $c_k = \frac{1}{k}$, cuando $k \geq 1$, entonces para*

$$C = \left\{\sup_{n\leq k} \frac{|S_k - E(S_k|\mathcal{G})|}{k} \geq \varepsilon\right\},$$

se tiene

$$\varepsilon^2 P\left(C|\mathcal{G}\right) \leq \frac{\sum_{k=1}^{n} E[X_k - E(X_k|\mathcal{G})|\mathcal{G}]^2}{n^2} + \sum_{k=n+1}^{\infty} \frac{E[X_k - E(X_k|\mathcal{G})|\mathcal{G}]^2}{k^2} \quad (c.t.p.).$$

Además si se supone $\sum_{k=1}^{\infty} \frac{E[X_k - E(X_k|\mathcal{G})|\mathcal{G}]^2}{k^2} < \infty$ *(c.t.p.), entonces para una variable aleatoria* $\varepsilon > 0$ *(c.t.p.)* $\mathcal{G}$*-medible se cumple*

$$\lim_{n\to\infty} \varepsilon^2 P(C|\mathcal{G}) = 0 \quad (c.t.p.),$$

luego

$$P(\limsup_{n\to\infty} \frac{|S_n - E(S_n|\mathcal{G})|}{n} \geq \varepsilon \mid \mathcal{G}) = 0 \quad (c.t.p.),$$

lo cual es equivalente

$$P(\liminf_{n\to\infty} \frac{|S_n - E(S_n|\mathcal{G})|}{n} < \varepsilon \mid \mathcal{G}) = 1 \quad (c.t.p.),$$

para toda variable aleatoria $\varepsilon > 0$ *(c.t.p)* $\mathcal{G}$*-medible, de lo que se concluye*

$$P(\lim_{n\to\infty} \frac{|S_n - E(S_n|\mathcal{G})|}{n} = 0 \mid \mathcal{G}) = 1 \quad (c.t.p.).$$

3.4. Ley Fuerte de los grandes números condicionada

Dada $X_1, X_2, \cdots$ una sucesión de variables aleatorias independientes, con media finita y varianza uniformemente acotada. La ley débil de los grandes números 1.3.13 establece que $\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{P} \mu = \frac{E(\sum_{i=1}^{n} X_i)}{n}$. En el año 1930 Kolmogorov mostró que este resultado se podia mejorar demostrando que la convergencia a μ se da no sólo en probabilidad si no también con probabilidad 1, la cual como ya se mostró es un tipo de convergencia más fuerte. A continuación se enuncian las versiones de la ley fuerte de los grandes números de Cantelli y de Kolmogorov, las cuales están demostradas en la Sección 3 del Capítulo IV de Shiryaev [16].

Teorema 3.4.1 (de Cantelli) *Sean* $X_1, X_2, \cdots$ *variables aleatorias independientes con cuarto momento finito (*$E|X_i|^4 < \infty$ *para todo* $i \in \mathbb{N}$*) y sea*

$$E|X_i - E(X_i)|^4 \leq C,$$

con C *constante. Entonces*

$$P\left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = \mu\right) = 1, \tag{3.2}$$

donde $\mu = \frac{E(\sum_{i=1}^{n} X_i)}{n}$.

Teorema 3.4.2 (de Kolmogorov) *Sea $X_1, X_2, \cdots$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con $E|X_1| < \infty$. Entonces*

$$P\left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = E(X_1)\right) = 1.$$

siguiendo esta idea se van a enunciar y demostrar las versiones condicionadas de los anteriores resultados.

Teorema 3.4.3 (De Cantelli condicinada) *Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias $\mathcal{G}$-independientes, donde para algún $r \geq 1$ se cumple*

$$\sum_{n=1}^{\infty} \frac{E(|X_n - E(X_n|\mathcal{G})|^{2r} \mid \mathcal{G})}{n^{r+1}} < \infty \quad (c.t.p.).$$

Entonces

$$P(\lim_{n\to\infty} \frac{S_n - E(S_n|\mathcal{G})}{n} = 0 \mid \mathcal{G}) = 1 \quad (c.t.p.).$$

Demostración. Sin perdida de generalidad se asumira $E(X_k|\mathcal{G}) = 0$ (c.t.p.), $k \geq 1$, de lo cual $E(S_n|\mathcal{G}) = 0$ (c.t.p.). Si se toma $D_k = \{\omega : |S_n| > n\varepsilon, n \in [2^k, 2^{k+1}]\}$, para $\varepsilon > 0$ (c.t.p.) una variable aleatoria $\mathcal{G}$-medible, entonces de la definición del conjunto D_k, se tiene que $|S_n| > \varepsilon 2^k$ para $2^k \leq n \leq 2^{k+1}$ y por la desigualdad generalizada de Kolmogorov para $k \geq 1$

$$(\varepsilon 2^k)^{2r} P(D_k|\mathcal{G}) \leq E(|S_n|^{2r}|\mathcal{G}) \leq E(|S_{2^{k+1}}|^{2r}|\mathcal{G}) \leq (2^{k+1})^r \sum_{n=1}^{2^{k+1}} E(|X_n|^{2r}|\mathcal{G}) \quad (c.t.p.).$$

Luego se cumple (c.t.p.)

$$\sum_{k=0}^{\infty} P(D_k|\mathcal{G}) \leq \frac{1}{\varepsilon^{2r}} \sum_{k=0}^{\infty} \frac{2^{(k+1)r}}{2^{2kr}} \sum_{n=1}^{2^{k+1}} E(|X_n|^{2^r}|\mathcal{G}) = \frac{1}{\varepsilon^{2r}} \sum_{n=1}^{\infty} E(|X_n|^{2^r}|\mathcal{G}) \sum_{k:2^{k+1}\geq n} \frac{2^{(k+1)r}}{2^{2kr}},$$

como

$$\begin{aligned}
\sum_{k:2^{k+1}\geq n} \frac{2^{(k+1)r}}{2^{2kr}} &\leq \sum_{k=0}^{\infty} \frac{2^{(k+1)r}}{2^{2kr}} = \sum_{k=0}^{\infty} \frac{1}{(2^r)^{k-1}} = 2^r + \sum_{k=1}^{\infty} \left(\frac{1}{2^r}\right)^{k-1} = \frac{2^{2r}}{2^r - 1} \\
&\leq \frac{2^{2r}}{2^r - 1} \cdot \frac{(2^{k_n+1})^{r+1}}{n^{r+1}} = c_r \frac{1}{n^{r+1}},
\end{aligned}$$

donde k_n es el entero más pequeño k tal que $2^{k+1} \geq n$. Por lo tanto de lo anterior y de lo supuesto en la hipótesis

$$\sum_{k=0}^{\infty} P(D_k|\mathcal{G}) \leq \frac{c_r}{\varepsilon^{2r}} \sum_{n=1}^{\infty} \frac{E(|X_n|^{2^r}|\mathcal{G})}{n^{r+1}} < \infty \quad (c.t.p.),$$

es decir, $P(\{\omega : \sum_{k=0}^{\infty} P(D_k|\mathcal{G}) < \infty\}) = 1$ y al aplicar el punto (2) de 3.2.4 se tiene que $P(\limsup_{k\to\infty} D_k) = 0$, lo cual es igual a $P(\lim_{n\to\infty}\{\omega : |S_n| > n\varepsilon\}) = 0$. Entonces de la definición de esperanza condicionada lo anterior equivale a $E(P(\lim_{n\to\infty}\{\omega : |S_n| > n\varepsilon\}|\mathcal{G})) = 0$, luego $P(\lim_{n\to\infty}\{\omega : |S_n| > n\varepsilon\}|\mathcal{G}) = 0$ (c.t.p.), de lo cual $P(\lim_{n\to\infty} \frac{|S_n|}{n} = 0 \mid \mathcal{G}) = 1$ (c.t.p.).$\square$

Lema 3.4.4 (Toeplitz) *Sean $\{a_n\}_{n\in\mathbb{N}}$ una sucesión de números no negativos, $b_n = \sum_{i=1}^{n} a_i$, $b_n > 0$ para $n \geq 1$, con $b_n \uparrow \infty$ cuando $n \to \infty$ y $\{X_n\}_{n\in\mathbb{N}}$ una sucesión convergente a X. Entonces*

$$\lim_{n\to\infty} \frac{1}{b_n}\sum_{j=1}^{n} a_j X_j = X,$$

si en particular $a_n = 1$, para $n \in \mathbb{N}$, entonces

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = X.$$

Demostración. Si $X_n \to X$, entonces para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para $n > n_0$, se tiene $|X_n - X| < \frac{\varepsilon}{2}$, como $b_n \uparrow \infty$ entonces existe $n_1 \in \mathbb{N}$ tal que $n_1 > n_0$ y

$$\frac{2}{\varepsilon}[\sum_{j=1}^{n_0} a_j|X_j - X|] < b_{n_1},$$

luego para $n > n_1 > n_0$ se tiene

$$\begin{aligned}
|\frac{1}{b_n}\sum_{j=1}^{n} a_j X_j - X| &\leq \frac{1}{b_n}\sum_{j=1}^{n} a_j|X_j - X| \\
&= \frac{1}{b_n}\sum_{j=1}^{n_0} a_j|X_j - X| + \frac{1}{b_n}\sum_{j=n_0+1}^{n} a_j|X_j - X| \\
&\leq \frac{1}{b_{n_1}}\sum_{j=1}^{n_0} a_j|X_j - X| + \frac{1}{b_n}\sum_{j=n_0+1}^{n} a_j|X_j - X| \\
&< \frac{\varepsilon}{2} + \frac{1}{b_n}\sum_{j=n_0+1}^{n} a_j \cdot \frac{\varepsilon}{2} < \varepsilon.\square
\end{aligned}$$

El siguiente resultado es un caso especial del teorema anterior, en este caso la sucesión de variables es idénticamente distribuida dada la σ-álgebra $\mathcal{G}$ y es conocido como la versión condicional de la ley de los grandes números para distribuciones condicionadas idénticamente distribuidas.

Teorema 3.4.5 (De Kolmogorov condicionada) *Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias $\mathcal{G}$-independientes e idénticamente distribuidas dada $\mathcal{G}$, entonces*

$$\lim_{n\to\infty} \frac{S_n}{n} = Y \quad (c.t.p.)$$

si y sólo si $E(X_1|\mathcal{G}) = Y$ (c.t.p.).

Demostración. ($\Leftarrow$) Sea $Y_n = X_n I_{\{|X_n|\leq n\}}$, donde las variables aleatorias $\{X_n\}_{n\in\mathbb{N}}$ son idénticamente distribuidas dada $\mathcal{G}$, entonces

$$\begin{aligned}
\sum_{n=1}^{\infty} P(X_n \neq Y_n|\mathcal{G}) &= \sum_{n=1}^{\infty} P(|X_n| > n|\mathcal{G}) = \sum_{n=1}^{\infty} P(|X_1| > n|\mathcal{G}) \\
&= \sum_{n=1}^{\infty}\sum_{k\geq n} P(k < |X_1| \leq k+1|\mathcal{G}) \\
&= \sum_{k=1}^{\infty} kP(k < |X_1| \leq k+1|\mathcal{G}) \\
&= \sum_{k=0}^{\infty} E(kI_{k<|X_1|\leq k+1}|\mathcal{G}) \\
&< \sum_{k=0}^{\infty} E(|X_1|I_{k<|X_1|\leq k+1}|\mathcal{G}) = E(X_1|\mathcal{G}) \\
&\leq \sum_{k=0}^{\infty} E((k+1)I_{k<|X_1|\leq k+1}|\mathcal{G}) \\
&= \sum_{n=1}^{\infty} P(|X_1| > n|\mathcal{G}) + \sum_{k=0}^{\infty} P(k < |X_1| \leq k+1|\mathcal{G}) \\
&= \sum_{n=1}^{\infty} P(X_n \neq Y_n|\mathcal{G}) + 1 \quad (c.t.p.),
\end{aligned}$$

es decir, $\sum_{n=1}^{\infty} P(X_n \neq Y_n|\mathcal{G}) < \infty$ (c.t.p.). Entonces por (2) del lema 3.2.4

$$P(\limsup_{n\to\infty}\{\omega : X_n \neq Y_n\}) = 0,$$

que es igual a $P(\limsup_{n\to\infty} |X_n| > n) = 0$, es decir $P(\lim_{n\to\infty}(\inf\{\omega : |X_n| \leq n\})) = 1$, por tanto $P(\bigcup_{n=1}^{\infty}\{\omega : X_n = Y_n\}) = 1$. Entonces demostrar que $\lim_{n\to\infty} \frac{S_n}{n} = E(X_1|\mathcal{G})$ (c.t.p.) equivale a que $\lim_{n\to\infty} \frac{Y_1+\cdots+Y_n}{n} = E(X_1|\mathcal{G})$ (c.t.p.). Como

$$\begin{aligned}
\sum_{n=1}^{\infty}\frac{E(|Y_n-E(Y_n|\mathcal{G})|^2\mid\mathcal{G})}{n^2} &\leq \sum_{n=1}^{\infty}\frac{E(|Y_n|^2\mid\mathcal{G})}{n^2}=\sum_{n=1}^{\infty}\frac{E(|X_n|^2 I_{\{|X_n|\leq n\}}\mid\mathcal{G})}{n^2}\\
&= \sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{k=1}^{n}E(|X_1|^2 I_{\{k-1\leq|X_1|\leq k\}}|\mathcal{G})\\
&= \sum_{k=1}^{\infty}E(|X_1|^2 I_{\{k-1\leq|X_1|\leq k\}}|\mathcal{G})\sum_{n=k}^{\infty}\frac{1}{n^2}\\
&\leq \sum_{k=1}^{\infty}E(|X_1|^2 I_{\{k-1\leq|X_1|\leq k\}}|\mathcal{G})\frac{2}{k}\\
&\leq 2\sum_{k=1}^{\infty}E(|X_1| I_{\{k-1\leq|X_1|\leq k\}}|\mathcal{G})\\
&= 2E(|X_1|\mid\mathcal{G})<\infty \quad (c.t.p.),
\end{aligned}$$

entonces por el teorema 3.4.3

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}[Y_k-E(Y_k|\mathcal{G})]}{n}=0 \quad (c.t.p.),$$

es decir

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}Y_k}{n}=\lim_{n\to\infty}\frac{\sum_{k=1}^{n}E(Y_k|\mathcal{G})}{n} \quad (c.t.p.).$$

Por otro lado como las $\{X_n\}_{n\in\mathbb{N}}$ son variables aleatorias idénticamente distribuidas dada $\mathcal{G}$, entonces

$$\lim_{n\to\infty}E(Y_n|\mathcal{G})=\lim_{n\to\infty}E(X_n I_{\{|X_n|\leq n\}}\mid\mathcal{G})=\lim_{n\to\infty}E(X_1 I_{\{|X_1|\leq n\}}\mid\mathcal{G})=E(X_1|\mathcal{G}) \quad (c.t.p.),$$

luego del lema de Toeplitz se tiene

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}E(Y_k|\mathcal{G})}{n}=E(X_1|\mathcal{G}).$$

Por tanto se concluye que

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}Y_k}{n}=E(X_1|\mathcal{G}) \quad (c.t.p.),$$

demostrando lo deseado.

($\Rightarrow$) Ahora si $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias $\mathcal{G}$-independientes e idénticamente distribuidas dada $\mathcal{G}$ tal que

$$\lim_{n\to\infty}\frac{S_n}{n}=Y \quad (c.t.p.),$$

entonces

$$\lim_{n\to\infty} \frac{X_n}{n} = \lim_{n\to\infty} \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1} = 0 \quad (c.t.p.),$$

de lo cual $P\left(\limsup\{\left|\frac{X_n}{n}\right| > 1\}\right) = 0$, luego del punto (2) de 3.2.4

$$\sum_{n=1}^{\infty} P\left(\left|\frac{X_1}{n}\right| > 1 \mid \mathcal{G}\right) < \infty \quad (c.t.p.)$$

y como en la primera implicación se obtuvo

$$\sum_{n=1}^{\infty} P(|X_1| > n \mid \mathcal{G}) < E(X_1|\mathcal{G}) \leq \sum_{n=1}^{\infty} P(|X_1| > n \mid \mathcal{G}) + 1 \quad (c.t.p.),$$

luego $E(X_1|\mathcal{G}) < \infty$ (c.t.p.), entonces por lo hecho en la primera implicación se obtiene que $E(X_1|\mathcal{G}) = Y$ (c.t.p.).$\square$

Conclusiones

Se establecieron algunas propiedades generales de la relación de independencia condicionada que, junto a las propiedades especiales enunciadas por Van Putten y Van Schuppen [17], conforman una serie de resultados los cuales son de utilidad en la solución de algunos problemas que involucren independencia condicionada.

Se establecieron algunos resultados de familias recíprocas y markovianas, mostrando la relación entre estas familias y algunas propiedades de estas mismas.

Se establecieron algunas versiones condicionadas de resultados clásicos de la teoría de probabilidad. En concreto se presento una versión condicionada del lema de Borel-Cantelli, la cual es más general que la correspondiente a la formulada por Yan [19] y que la establecida por Majerek et al. [9].

Bibliografía

[1] Ash R. B. *Real Analysis and Probability.* Academic Press, Inc., New York, 1972.

[2] Bauer H. *Probability Theory and Elements of Measure Theory.* Holt, Rinehart and Winston, Inc., 1972.

[3] Billingsley P. *Probability and Measure.* John Wiley and Sons, Inc., New York, 1979.

[4] Brzeźniak Z. and Zastawniak T. *Basic Stochastic Processes*. Springer-Verlarg, 1999.

[5] Feng C.; Li L. and Shen J. *On Borel-Cantelli Lema and its generalizations*, C. R. Acad. Sci. Paris (2009), Ser. I347, 1313-1316.

[6] Gill B. J. *Markovian extensions and reductions of a family of σ-algebras.* Internat. J. Math. and Math. Sci. (1984), Vol. 7, No. 3, 523-528.

[7] Jamison B. *Reciprocal processes: the stacionary Gaussian case.* The Annals of Mathematical Statistics (1970), Vol. 41, No. 5, 1624-1630.

[8] Loève M. *Probability Theory II.* Four Edition. Springer-Verlag, New York, 1978.

[9] Majerek, D.; Nowak, W. and Zieba, W. *Conditional strong law of large number*. International Journal of Pure and Applied Mathematics (2005), Vol. 20, No. 2, 143-157.

[10] Marmolejo L. Miguel A.; Mendoza Evelin; Muñoz Andrés F. *Versión condicionada de una generalización del lema de Borel-Cantelli.* Revista de Ciencias, Universidad del Valle (2011), Vol. 15, 83-92.

[11] Marmolejo L. Miguel A.; Muñoz Andrés F. *Algunas propiedades de la independencia condicionada.* Aceptado para publicación en la Revista integración, Universidad Industrial de Santander (2013).

[12] Nogales A. G. and Oyola J. A. *Some Remarks on Sufficiency, Invariance and conditional independence.* The Annals of Statistics (1996), 24 No. 2, 906-909.

[13] Nogales A. G.; Oyola J. A. ; Pérez P. *On conditional independence and the relationship between sufficience and invariance under the Bayesian point of view.* Statistics and Probability Letters (2000), 46, 75-84.

[14] Petrov, V. A generalization of the Borel-Cantelli lemma. Statistics and Probability Letters (2004), 67, 233-239.

[15] Prakasa Rao B. L. S. *Conditional independence, conditional mixing and conditional association.* Annals of the Institute of Statistical Mathematics (2009), Vol. 61, No. 2, 441-460.

[16] Shiryaev A. N. *Probability.* Second Edition Springer, New York, 1996.

[17] Van Putten C. and Van Schuppen J. H. *Invariance properties of the conditional independence relation.* The Annals of Probability (1985), Vol. 13, No. 3, 934-945.

[18] Williams D. *Probability with Martingales.* Cambridge University Press, Cambridge, 1997.

[19] Yan J. *A simple proof of two generalized Borel-Cantelli Lemmas.* In memoriam Paul-André Meyer: Seminar on Probability Theory XXXIX, in: Lecture Notes in Mathematics (2006), Vol. 1874, 77-79, Springer-Verlag, Berlin.

Índice alfabético