Notas de Probabilidad y Procesos Estocásticos
Latex Export

Carlos Eduardo artínez Aguilar

4 de mayo de 2023

Índice

1. Historia y repaso elementa de probabilidades.

La historia de la probabilidad como rama matemática comienza con Fermat cuando investiga los problemas que un amigo apostador de Pascal le propone. Desde los tiempos de Bernoulli (1713) se conocian resultados como la ley (débil) de los grandes números, sin embargo no es hasta el siglo veinte que se considera una base axiomática basada en la teoría de la medida para la teoria probabilística que se encuentra una demostración rigurosa del la ley de los grandes números y el teorema del límite central. Más aún, existen dos vertientes ideológicas distintas de la teoría de la probabilidad

- Una es la Probabilidad “objertiva”, la cual se basa en la axiomatización dada por la teoría de la medida, la cual fue detallada por Kolmogorov y postula a la probabilidad como el resultado que permite calcular los “outcomes” de una serie de experimentos los cuales son en teoría infinitamente repetibles (von Mises).
- La otra es la Probabilidad "subjetiva", propuesta por Keynes, la cual busca responder preguntas como: “¿Mañana va a lloverá?” o “Existe vida en saturno?” Su influencia mayor está dada por la escuela bayesiana de estadística.

1.1. Medida y probabilidad.

Entenderemos un *espacio de probabilidad* como un espacio de medida $(\Omega, \Sigma, \mathbf{P})$ con medida uno, es decir un conjunto con una *sigma-álgebra* Σ y una medida $\mathbf{P}$, tal que, el espacio total tenga medida $\mathbf{P}(X) = 1$. Se le denomina a los conjuntos medibles de un espacio de probabilidad como *eventos* y a dicha medida le llamamos *medida de probabilidad*, así mismo se le denomina al espacio total X como *espacio de muestras*. Así la medida de probabilidad $\mathbf{P}$ evaluada en un evento E, $\mathbf{P}(E)$ es naturalmente la probabilidad del evento E y a un evento de probabilidad cero se le conoce como *evento nulo*.

Así un espacio de probabilidad $(\Omega, \Sigma, \mathbf{P})$, donde Ω es un conjunto Σ es una sigma-algebra y $\mathbf{P}$ es una medida de probabilidad, es decir Σ cumple que

- $\varnothing \in \Sigma$ y $\Omega \in \Sigma$
- $A \in \Sigma$ implica $\Omega \setminus A \in \Sigma$
- $\{A_n\} \subset \Sigma$ entonces $\bigcup A_n \in \Sigma$

Además $\mathbf{P}$ es una medida de probabilidad, es decir que cumple:

- $\mathbf{P}(A) \geq 0$
- $\mathbf{P}(\Omega) = 1$
- Si $\{A_n\} \subset \Sigma$ es tal que $A_n \cap A_{n+1} = \varnothing$, entonces $\mathbf{P}(\bigcup A_n) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$

Definición 1.1. *Dado un conjunto Ω, decimos que una familia de subconjuntos* F *de* Ω

- *Es un π sistema, si es cerrado bajo intersecciones finitas.*
- *Es un λ sistema o sistema de Dynkin si cumple*
    - $\Omega \in \mathrm{F}$

- $A, B \in \mathrm{F}$ *y* $A \subset B$*, entonces* $A \setminus B \in \mathrm{F}$
- *Si* $\{A_n\}_{n\in\mathbb{N}} \subset \mathrm{F}$ *es una sucesion creciente, entonces* $\bigcup_{n\in\mathbb{N}} A_n \in \mathrm{F}$

Lema 1.1. *Una familia* Σ *de subconjuntos de un conjunto* Ω *es una* σ *álgebra si y sólo si es un* π *sistema y* λ *sistema.*

Definición 1.2. *Sea* F *una familia de subconjuntos de* Ω

- *Definimos a la* σ *álgebra generada por* F *como la* σ *álgebra más chica que contiene a* F*, esto es equivalente a la intersección de todas las* σ *álgebras que contienen a* F *y se denotará por* $\sigma(\mathrm{F})$.
- *De forma similar el* λ *sistema generado por* F *se denota* $\lambda(\mathrm{F})$ *y es el* λ *sistema más chico que contiene a* F*.*

Teorema 1.2 (Teorema de clases monótonas de Sierpinsky)**.** *Sea* F *un* π *sistema de sunconjuntos de* Ω *y* G *un* λ *sistema tal que* $\mathrm{F} \subset \mathrm{G}$*, entonces* $\sigma(\mathrm{F}) \subset \mathrm{G}$ *o equivalentemente* $\lambda(\mathrm{F}) = \sigma(\mathrm{F})$

Demostración. Como una σ álgebra es un π sistema y λ sistema, entonces $\lambda(\mathrm{F}) \subset \sigma(\mathrm{F})$. Por lo tanto si demostramos que $\lambda(\mathrm{F})$ es una σ álgebra, entonces el teorema estaría demostrado, así sólos hay que ver que $\lambda(\mathrm{F})$ es un π sistema. Sea $C \in \mathrm{F}$, entonces definimos

$$\mathrm{G}_1 := \{A \in \lambda(\mathrm{F}) \mid A \cap C \in \lambda(\mathrm{F})\}.$$

Veamos que G_1 es un λ sistema

- Como $\Omega \cap C = C \in \lambda(\mathrm{F})$, entonces $\Omega \in \mathrm{F}$
- Si $A, B \in \mathrm{G}_1$ con $A \subset B$, entonces $A \cap C \subset B \cap C \in \lambda(\mathrm{F})$ y como es λ sistema por definición, entonces $A \cap C \setminus B \cap C = (A \setminus B) \cap C \in \lambda(\mathrm{F})$
- De igual manera, si $\{A_n\} \subset \mathrm{G}_1$ es una sucesión creciente de subconjuntos, entonces $\bigcup A_n \cap C \in \lambda(\mathrm{F})$ y por lo tanto $\bigcup A_n \mathrm{G}_1$.

Por lo tanto D_1 es un λ sistema contenido en $\lambda(\mathrm{F})$ que contiene a F pues éste es un π sistema, así $\lambda(\mathrm{F}) \subset \mathrm{G}_1$, por lo tanto $C \cap A \in \lambda(\mathrm{F})$ para todo $C \in \mathrm{F}$ y $A \in \lambda(\mathrm{F})$. Similarmente, si tomamos $A \in \lambda(\mathrm{F})$ fija, definimos

$$\mathrm{G}_2 = \{B \in \lambda(\mathrm{F}) \mid B \cap A \in \lambda(\mathrm{F})\},$$

se puede demostrar que G_2 es un λ sistema que contiene a F, entonces $\lambda(\mathrm{F}) \subset \mathrm{G}_2$, así $\lambda(\mathrm{F})$ contiene intersecciones finitas y por lo tanto es una σ álgebra. $\square$

Corolario 1.3. *Sea* $(\Omega, \Sigma, \boldsymbol{P})$ *un espacio de probabilidad tal que* $\Sigma = \sigma(\mathrm{F})$ *si* μ_1 *y* μ_2 *son medidas en* Σ *tales que* $\mu_1 = \mu_2$ *en* F*, entonces* $\mu_1 = \mu_2$*.*

Observación. A partir de estos resultados se puede demostrar el teorema de extensión de Caratheodory.
Ejemplo. Si $\mathbf{P}$ y $\mathbf{Q}$ son dos probabilidades en (Ω, Σ), se puede ver que

$$\mathrm{F} = \{A \in \Sigma \,|\, \mathbf{P}(A) = \mathbf{Q}(A)\}$$

es un λ sistema, entonces si un π sistema contenido en F que genera a Σ, entonces $\mathbf{P}$ y $\mathbf{Q}$ son iguales.

Similarmente entendemos por *variable aleatoria* como una función medible $X : \Omega \to \mathbb{R}$ en un espacio de probabilidad que cumple lo siguiente;

$$\mathbb{P}\Big(\{z \in \Omega \,|\, |X(z)| = \infty\}\Big) = 0.$$

Si un evento tiene como complemento un evento nulo, en otra palabras, si el evento tiene probabilidad 1 diremos que el evento es *casi seguro.*
Ejemplo. El ejemplo prototípico son los intervalos finitos de números reales cuya sigma álgebra es el álgebra de los conjuntos *borelianos* o los *medibles de lebesgue* y la medida de Lebesgue reescalada para medir exactamente 1.
Observación. Dada una sucesión $\{x_n\}_{n\in\mathbb{N}}$ denotaremos $x_n \downarrow x$ o $x_n \uparrow x$ si $x_{n+1} \leq x_n$ y $x_{n+1} \geq x_n$ $\lim_{n\to\infty} x_n = x$ respectivamente. Aplicaremos la misma notación para eventos $\{A_n\}_{n\in\mathbb{N}} \subset \Sigma$.
Existen construcciones conjuntistas que son más útiles en el sentido probabilístico, por ejemplo si $\{A_n\}_{n\in\mathbb{N}}$ es una sucesión de eventos aletorios definimos los conjuntos $\{A_n \text{ i.o.}\}$ y $\{A_n \text{ ult.}\}$ dónde i.o significa “infinitely often” y ult “ultimately”, éstos términos se refieren a el comportamiento de las colas de la sucesión de los eventos

$$\{A_n \text{ i.o.}\} = \Big\{\sum_{n\in\mathbb{N}} 1_{A_n} = \infty\Big\} = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k,$$
$$\{A_n \text{ ult.}\} = \Big\{\sum_{n\in\mathbb{N}} 1_{A_n} < \infty\Big\} = \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_k.$$

Notemos que esto se puede expresar en términos de funciones indicadoras como

$$1_{\{A_n \text{ i.o.}\}} = \limsup_{n\to\infty} 1_{A_n} \quad 1_{\{A_n \text{ ult.}\}} = \liminf_{n\to\infty} 1_{A_n}.$$

Ahora por la *desigualdad de Fatou* se cumple lo siguiente con respecto a las probabilidades de estos conjuntos

$$\mathbf{P}(1_{\{A_n \text{ i.o.}\}}) \geq \limsup_{n\to\infty} \mathbf{P}(A_n) \quad \mathbf{P}(\{A_n \text{ ult.}\}) \leq \liminf_{n\to\infty} \mathbf{P}(A_n).$$

Por la subaditividad y continuidad de la probabilidad se obtiene

$$\mathbf{P}(\{A_n \text{ i.o.}\}) = \lim_{n\to\infty} \mathbf{P}(\bigcup_{k\geq n} A_k) \leq \lim_{n\to\infty} \sum_{k\geq n} \mathbf{P}(A_k).$$

Observación. Si $\sum_{n\in\mathbb{N}} \mathbf{P}(A_n)$ entonces las colas tienden a cero y por lo tanto $\mathbf{P}(\{A_n \text{ i.o.}\}) = 0$. Esto corresponde a la parte fácil del lema de *Borel-Cantelli* que veremos en el futuro.

Definición 1.3. *Dado un espacio de probabilidad* $(\Omega, \Sigma, \boldsymbol{P})$*, un espacio medible* (E, S) *y un conjunto arbitrario* T*, el cual llamaremos a partir de ahora como un* índice arbitrario o abstracto*, es posible definir el espacio medible* $E^T := \{f : T \to E\}$ *con la sigma álgebra definida por los mapeos de evaluación, definidos para cada* $t \in T$ *como* $\pi_t : E^T \to E$*,* $\pi_t(f) := f(t)$*. Así si* $X : \Omega \to E^T$ *es un elemento aleatorio, se define la función* $X(t, \omega) = \pi_t \circ X$*, donde* $X : T \times \Omega \to E$*, a esta función le llamaremos un* proceso aleatorio*. Si* T *es de cardinalidad numerable diremos que* X *es un proceso de tiempo discreto y si* T *es de cardinalidad mayor a la numerable, diremos que* X *es de tiempo continuo.*

Observación. Un proceso $X : T \times \Omega \to E$ es aleatorio (medible) si y solo si para toda $t \in T$, $X_t : \Omega \to E$ definida por $X_t(\omega) := X(t, \omega)$ es medible.

1.1.1. Medidas con signo

Además de medidas de probabilidad, existen las *medidas con signo* es decir, si (Ω, Σ) es un espacio medible, entonces una función $\nu : \Sigma \to \mathbb{R}$ es una medida con signo si cumple

- $\nu(\varnothing) = 0$.
- ν toma a lo más uno de los valores ∞ o $-\infty$.
- Si $\{A_n\} \subset \Sigma$ es una sucesión de conjuntos disjuntos, entonces $\nu(\bigcup A_n) = \sum_{n=1}^{\infty} \nu(A_n)$, donde entendemos que para valores finitos la convergencia es absoluta.

Ejemplo. Un ejemplo que utilizaremos todo el tiempo es la medida con signo definida por una variable aleatoria $X : \Omega \to \mathbb{R}$

$$\nu(A) = \int_A X d\mathbf{P} = \mathbb{E}(X 1_A),$$

en el caso de una variable aleatoria **positiva**, con esperanza finita (o esperanza 1), le llamaremos *distribución de probabilidad*, pero habrá más información sobre estas probabilidades en una sección futura.

Definición 1.4. *Sea* (Ω, Σ, μ) *un espacio de medida y* ν *una medida con signo, decimos que* ν *es absolutamente continua con respecto a* μ *(denotado* $\nu \ll \mu$*) si* $\mu(A) = 0$*, entonces* $\nu(A) = 0$*.*

Una medida con signo ν, definida por una variable aleatoria X es absolutamente continua con respecto a la probabilidad $\mathbf{P}$. Para cualquier función medible $f : \Omega \to \mathbb{R}$ **siempre** podemos descomponerla de la siguiente manera $f = f^+ - f^-$, donde $f^{\pm} = \pm f 1_{E^{\pm}}$ y $E^{\pm} = \{x \in \Omega \mid \pm f > 0\}$. De la misma manera una medida con signo definida por f, ν se descompone en $\nu^{\pm}$.

1.2. Esperanza

Asi mismo entenderemos a la *esperanza* de una variable aleatoria como la integral de dicha funcion medible, es decir

$$\mathbb{E}(f(z)) = \begin{cases} \sum_i p_i f(z) \\ \int_\Omega f(z) d\mu \end{cases} \tag{1}$$

Comunemente se define la probabilidad condicional de un evento A con respecto a otro evento B como la siguiente fórmula

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

sinembargo comunmente se utiliza el concepto más general de *esperanza condicional*

Definición 1.5. *Si* $(\Omega, \Sigma, \boldsymbol{P})$ *es un espacio de probabilidad,* $X : \Omega \to \mathbb{R}$ *es una variable aleatoria inegrable y* $\mathrm{T} \subset \Sigma$ *es una sigma sub álgebra, entonces decimos que una funcion* $Y : \Omega \to \mathbb{R}$ *es una* esperanza *condicional si cumple lo siguiente.*

- Y *es* T *medible*
- $\mathbb{E}(|Y|) < \infty$
- $\mathbb{E}(X1_A) = \mathbb{E}(Y1_A)$ *para todo* $A \in \mathrm{T}$

se denotará a la esperanza condicional de X *con respecto a* T *como* $\mathbb{E}(X|\mathrm{T})$.

Demostración. Podemos demostrar la existencia de la esperanza condicional por medio del teorema de Radon-Nikodym (1930)

Teorema 1.4 (Radon-Nikodym). *Sea* (Ω, Σ) *un espacio medible y* $\mu : \Omega \to \mathbb{R}$ *una medida* (σ) *finita y* ν *una medida con signo* (σ) *finita tal que* ν *sea absolutamente continua con respecto a* μ*, entonces existe* $f : \Omega \to \mathbb{R}$ *funcion medible tal que*

$$\nu(A) = \int_A f d\mu \qquad \forall A \in \Sigma, \tag{2}$$

además la función f *es única casi donde quiera, a esta función se le conoce como *derivada de Radon-Nikodym* y usualmente se denota*

$$f = \frac{d\nu}{d\mu}[\mu].$$

.

Así la existencia de la esperanza condicional proviene de la derivada de Radon-Nikodym de la medida con signo definida por X, para el álgebra F, es decir

$$\nu\pm(A) = \int_A X^{\pm} d\mathbf{P} \qquad \nu = \nu^+ - \nu^-,$$

la cual es absolutamente continua con respecto a la medida de probabilidad $\mathbf{P}$, entonces definimos $Y = \frac{d\nu}{d\mathbf{P}}[\mathbf{P}]$, así simplemente verificamos que en efecto es una esperanza condicional. Por Radon-Nykodym, Y es medible, ahora

- $\mathbb{E}(|Y|) = \int_\Omega |Y| d\mathbf{P} = \int_\Omega |X| d\mathbf{P} < \infty$
- $\mathbb{E}(X1_A) = \int_A X d\mathbf{P} = \nu(A) = \int_A Y d\mathbf{P} = \mathbb{E}(Y1_A)$.

Algunas propiedades de la esperanza condicional son las siguientes:

- Linealidad: para todas $a, b \in \mathbb{R}$ y X, Y **v.a** $\mathbb{E}(aX + bY|\mathrm{F}) = a\mathbb{E}(X|\mathrm{F}) + b\mathbb{E}(Y|\mathrm{F})$.
- Monotonía: si $X \geq 0$, entonces $\mathbb{E}(Y|\mathrm{F}) \geq 0$.
- Si X es F medible, entonces $\mathbb{E}(X|\mathrm{F}) = X$.
- Propiedad de Torre: Si $\mathrm{G} \subset \mathrm{F} \subset \Sigma$, entonces:

$$\mathbb{E}(\mathbb{E}(X|\mathrm{F})|\mathrm{G}) = \mathbb{E}(\mathbb{E}(X|\mathrm{G})|\mathrm{F}) = \mathbb{E}(X|\mathrm{G})$$

Más aún la esperanza condicional hereda las propiedades que son resultados comunes de la teoría de la integración, por ejemplo

- Convergencia monótona: Si $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión de **v.as** no negativas que converge monótonamente $X_N \uparrow X$, entonces $\mathbb{E}(X|\mathrm{F}) = \lim \mathbb{E}(X_n|\mathrm{F})$.
- Convergencia dominada: Si $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión de **v.as** convergente $X_n \to X$, tal que las **v.as** sean integrables, no negativas y se encuentren dominadas por $|X_n| \leq Y$ tal que $\mathbb{E}(Y|\mathrm{F}) < \infty$, entonces $\mathbb{E}(X|\mathrm{F}) = \lim \mathbb{E}(X_n|\mathrm{F})$.
- Desigualdad de Jensen: Si $f : \mathbb{R} \to \mathbb{R}$ es una función convexa (cóncava) y X es una variable aleatoria, entonces $\mathbb{E}(f(X)|\mathrm{F}) \geq \mathbb{E}(X|\mathrm{F})$ (resp $\mathbb{E}(f(X)|\mathrm{F}) \leq \mathbb{E}(X|\mathrm{F})$)

1.2.1. Distribuciones de probabilidad

Consideremos $(\Omega, \Sigma, \mathbf{P})$ un espacio de probabilidad y sea (E, S) un espacio medible y $\xi : \Omega \to E$ una función medible, a esta función le llamremos *elemento aleatorio* así existe una función $\xi^* : S \to \Sigma$, definida como $\xi^*(B) = \xi^{-1}(B)$, entonces es posible definir una medida de probabilidad en $(\xi^*)^{-1}(\Sigma)$, definida por $\mathbf{P} \circ \xi^*(B) = \mathbf{P}(\xi^*(B))$, a la medida $\mathbf{P} \circ \xi^*$, le llamaremos una *distribución* de probabilidad y usualmente no distinguiremos entre una medida de probabilidad y una distribución incluso en la absencia de un elemento aleatorio.
Observación. Cuando $E = \mathbb{R}$ y $S = B_\mathbb{R}$, un elemento aleatorio es una variable aleatoria y cuando $E = \mathbb{R}^d$ y $S = B_{\mathbb{R}^d}$ un elemento aleatorio se le conoce como *vector aleatorio*.

Definición 1.6. *Función de una distribución Dado un espacio de probabilidad $(\Omega, \Sigma, \boldsymbol{P})$ y un elemento medible $\xi : \Omega \to E$, definimos la función de distribución como la función $F_\xi : \mathbb{R} \to [0, 1]$*

$$F_\xi(\alpha) = \boldsymbol{P} \circ \xi^*(-\infty, \alpha] = \boldsymbol{P}(\xi \leq \alpha)$$

Algunas propidades: Supongamos que F es la función de distribución asociada a una variable aleatoria X, entonces.

1. $F : \mathbb{R} \to [0,1]$ es no decreciente, es decir $F(x) \leq F(y)$ siempre que $x \leq y$.
2. $\lim_{x\to\infty} F(x) = 1$ y $\lim_{x\to-\infty} F(x) = 0$.
3. F es continua por la derecha.

Demostración. La demostración de estos hechos es evidente, el único que no es inmediatemente obvio es el punto número tres, para esto notamos que si tenemos la susesión $\{x + 1/n\}$, entonces por convergencia monónona tenemos que

$$\lim_{n\to\infty} \mathbf{P}\Big(X \leq x + \frac{1}{n}\Big) = \mathbf{P}\Big(X \leq x\Big)$$

Existencia: Ahora supongamos que F es una función que cumple las propiedades 1.,2. y 3., entonces similar a la existencia de Lebesgue, es posible construir una medida $\nu : \mathbb{B}_{\mathbb{R}} \to \mathbb{R}^+$ tal que $\nu((-\infty, x]) = F(x)$. A la medida ν se le conoce como la medida de Lebesgue-Stieltjes de F.

Ejemplo. (La representación de Skorokhod de una variable aleatoria con una función de distribución prescrita en el intervalo [0,1])
Supongamos que $F : \mathbb{R} \to [0,1]$ cumple con las propiedades 1.,2. y 3., como antes, entonces construimos una variable aleatoria en $([0,1], \mathbb{B}[0,1], \lambda)$, donde λ es la medida de Lebesgue. Definimos

$$X^+(\omega) := \inf\{\, z \mid F(z) > \omega\} = \sup\{\, y \mid F(y) \leq \omega\}$$
$$X^-(\omega) := \inf\{\, z \mid F(z) \geq \omega\} = \sup\{\, y \mid F(y) < \omega\}.$$

Observemos que por definición de X^-, si $\omega \leq F(c)$, entonces $X^-(\omega) \leq c$. Ahora, si $z > X^-(\omega)$, entonces $F(z) \geq \omega$, entonces por continuidad a la derecha de F, $F(X^-(\omega)) \geq \omega$ y por lo tanto

$$X^-(\omega) \leq c \implies \omega \leq F(X^-(\omega)) \leq F(c)$$

Por lo tanto tenemos que $\{\,\omega \mid X^-(\omega) \leq c\} = \{\,\omega \mid \omega \leq F(c)\}$ lo que implica que

$$\mathbf{P}(X \leq c) = \mathbf{P}(\omega \leq F(c)) = \int_{-\infty}^{c} F(x)\, dx = F(c)$$

□

Observación. X^+ tiene la mimsa función de distribución ya que

$\mathbf{P}(\mathrm{X}^- = X^+) = 1.$

1.3. Independencia

Nos enfocaremos en indepencenia de σ álgebras y describimos a partir de este la noción usual de independencia.

Definición 1.7. *Independencia Sea $(\Omega, \Sigma, \boldsymbol{P})$ un espacio de probabilidad, decimos que una familia de sub σ álgebras $\{\mathrm{F}_i\}_{i\in I}$ son* independientes *si para cualquier $\{i_1, \ldots, i_n\} \subset I$ subconjunto fininto, se tiene que*

$$\boldsymbol{P}(A_{i_1} \cap \cdots \cap A_{i_n}) = \prod_{j=1}^{n} \boldsymbol{P}(A_{i_j}) \quad A_{i_j} \in \mathrm{F}_{i_j}. \tag{3}$$

Similarmente a los eventos que cumplan con la ecuación 3, les llamaremos eventos independientes.

Definición 1.8. *Variables aleatorias independientes. Decimos que las variables aleatorias $\{X_i\}_{i\in I}$ en un espacio de probabilidad $(\Omega, \Sigma, \boldsymbol{P})$ son* independientes *si*

$$\{\sigma(X_i)\}_{i\in I}$$

son independientes. Se denotará la relación de independencia entre X e Y como $X \perp\!\!\!\perp Y$ y no distinguiremos entre variables aleatorias y σ álgebras, e incluso denotaremos $X \perp\!\!\!\perp \mathrm{F}$ cuando una variable aleatoria sea indepeniente de una sub σ álgebra.

En términos de de *eventos independientes* como sabemos la ecuación que se cumple es

$$\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_n}) = \prod_{j=1}^{n} \mathbf{P}(A_{i_j}),$$

queremos generalizar este tipo de ecuaciones pero en el sentido de π sistemas en lugar de en nuestro sentido original de σ álgebras ya que estas son más complicadas de usar. Por lo que el siguiente lema es muy útil para este tipo de cuestiones

Lema 1.5. *Sean* G *y* H *sub sigma álgebras de Σ y sean I y J π sistemas tales que*

$$\sigma(I) = \mathrm{G}, \quad \sigma(J) = \mathrm{H}.$$

Entonces G *y* H *son independientes si y solo si*

$$\boldsymbol{P}(A \cap B) = \boldsymbol{P}(A)\boldsymbol{P}(B) \quad A \in J,\, B \in J$$

Demostración. Es claro que si las σ álgebra son independientes entonces los π sistemas son independientes. Por lo que sólo es necesario demostrar que si los π sistemas son independientes entonces las σ álgebras lo son, entonces dadas $A \in I$ y $B \in J$ fijos, definimos las siguientes medidas finitas

$$C \mapsto \mathbf{P}(A \cap C) \quad C \mapsto \mathbf{P}(B)\mathbf{P}(C).$$

Como ambas medidascoinciden en I y J, por el lema de clases monótonas coincide en $\sigma(I) = \mathrm{G}$ y $\sigma(J) = \mathrm{H}$ y por lo tanto el lema es válido. $\square$

Corolario 1.6. *Ahora si X e Y son variables aleatorias en un espacio de probabilidad $(\Omega, \Sigma, \boldsymbol{P})$ tales que para todo $\{x, y\} \subset \mathbb{R}$ se cumple*

$$\boldsymbol{P}(\{X \leq x\} \cap \{Y \leq y\}) := \boldsymbol{P}(X \leq x\,,\, Y \leq y) = \boldsymbol{P}(X \leq x)\boldsymbol{P}(Y \leq y)$$

entonces X e Y son variables aleatorias independientes

Demostración. Esto es claro de la proposición anterior debido a que la familia de conjuntos $\{X^{-1}(-\infty, x] \,|\, x \in \mathbb{R}\}$ es un π sistema para $\sigma(X)$ para toda **v.a** X, así del lema anterior se sigue el resultado. □

Observación. En términos de la esperanza conicional, si $X \perp\!\!\!\perp \mathrm{F}$, entonces $\mathbb{E}(X|\mathrm{F}) = \mathbb{E}(X)$, además $Y = \mathbb{E}(X)$ es F medible pues

$$\forall A \in \mathrm{F} \quad \mathbb{E}(X 1_A) = \mathbb{E}(X)\mathbf{P}(A) = Y\mathbb{E}(1_A) = \mathbb{E}(Y 1_A)$$

Proposición 1.7. *Sea $(\Omega, \Sigma, \boldsymbol{P})$ espacio de probabilidad, $\mathrm{F} \subset \Sigma$ sub σ álgebra y X, Y v.a. tal que X sea F medible, entonces para toda $\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ Borel medible, se cumple*

$$\mathbb{E}(\varphi(X, Y)|\mathrm{F}) = \int_{y \in \mathbb{R}} \varphi(X, y)\boldsymbol{P}_Y(dy) = \mathbb{E}(\varphi(x, Y))|_{x \in \mathbb{R}}$$

Lema 1.8 (Borel-Cantelli)**.** *Sea $\{E_n\}$ una sucesión de eventos indpendientes en un espacio de probabilidad $(\Omega, \Sigma, \boldsymbol{P})$, entonces si*

$$\sum_{n=1}^{\infty} \boldsymbol{P}(E_n) = \infty \quad \textit{implica} \quad \boldsymbol{P}(\{E_n,\ i.o\ \}) = \boldsymbol{P}(\ \limsup E_n\ ) \tag{4}$$

Demostración. Primero, sabemos que se cumple

$$\Omega \setminus (\limsup E_n) = \liminf\{\Omega \setminus E_n\} = \bigcup_{m \in \mathbb{N}} \bigcap_{n \geq m} (\Omega \setminus E_n).$$

Ahora si $p_n = \mathbf{P}(E_n)$, entonces por independencia se cumple que

$$\mathbf{P}\Big(\bigcap_{n \geq m} (\Omega \setminus E_n)\Big) = \prod_{n \geq m} (1 - p_n).$$

Por monotonía de ambos lados de la ecuación, al tomar el límite cuando $r \uparrow \infty$ y la condición $\{r \geq n \geq m\}$, obtenemos gracias a la independencia lo siguiente. Para cada $x \geq 0$, $1 - x \leq \exp(-x)$, así como la serie de los p_n es divergente por hipótesis, entonces

$$\prod (1 - p_n) \leq \mathbb{E}\Big(-\sum_{n \geq m} p_n\Big) = 0.$$

Por lo que $\mathbf{P}(\{\Omega \setminus \limsup E_n\}) = 0$ □

Ejemplo.

1.4. TODO Tipos de convergencia,Teoremas límite, Ley fuerte de los grandes números y teorema del límite central.

2. Procesos estocáticos: Temario del curso

- Existencia de procesos y Martingalas
- Cadenas de Markov (discretas y continuas)
- Procesos de Poisson y renovacion
- Movimiento Browniano

2.1. Procesos estocáticos

Estudian eventos que evolucionan el tiempo (discreto o continuo), es decir tenemos una varible aleatoria que representa un estado en el sistema al tiempo t, digamos $X(t)$. Todo va a ser modelado en el mismo espacio de probabilidad $(\Omega, \Sigma, \mathbf{P})$

Definición 2.1. *Un procesos estocáticos es una colección de variables aleatorias* $\{X_t\}_{t \in T}$*, donde* T *le llamamos* espacio de parámetros*.*

$$X_t : (\Omega, \Sigma, \boldsymbol{P}) \to (X, S)$$

Estas funciones toman valores en un espacio medible, llamado espacio de estados*.*

Podemos pensar que nuestro proceso es una funcion de dos variables

$$X : T \times (\Omega, \Sigma, \boldsymbol{P}) \to (X, S)$$

Se clasifican los procesos estocácticos dependiendo de si el espacio de parámetros es **discreto** o **continuo**, similarmente se clasifican si el espacio de estados es **discreto** o **continuo**.

2.1.1. TODO Relaciones de dependencia

- Procesos de varibles aletorias independientes

$$s \neq t \qquad X_t \perp\!\!\!\perp X_s \qquad \{X_n\}_{n \geq 0}.$$

- Procesos con crecimientos independientes

$$\forall 0 \leq s \leq t \leq r \leq u \qquad X_t - X_s \perp\!\!\!\perp X_u - X_r.$$

  además si $\{X_n\}_{n \geq 0}$ son variables aleatorias independientes, entonces

$$Y_n = \sum_{j=1}^{n} X_j$$

  tiene crecimientos independientes.

- Procesos que cumplen la propiedad de Markov, intuitivamente esto quiere decir que estos procesos son independientes el pasado y el futuro dado el presente. Es decir si $\Sigma_t = \sigma(X_u\, u \leq t)$ es la sigma-álgebra generada por X_u (la cual pensamos como el álgebra del pasado) y $\overline{\Sigma_t} = \sigma(X_u\, u \geq t)$ es la sigma-algebra del futuro, entonces $\{X_t,\, t \geq 0\}$ cumple la propiedad de Markov si para todo $A \in \Sigma_t$ y $B \in \overline{\Sigma_t}$, emtomces

$$\mathbf{P}(AB|X_t) = \mathbf{P}(A|X_t)\mathbf{P}(B|X_t).$$

El futuro independiente del pasado si para todo $A \in \Sigma$, se cumple que

$$\mathbb{E}(X_{s+t} \in A \mid \Sigma_t) = \mathbb{E}(X_{s+t} \in A \mid X_t)$$

Usando funciones medibles:

$$\mathbb{E}(f(X_{s+t}) \mid \Sigma_t) = \mathbb{E}(f(X_{s+t}) \mid X_t)\, \forall f \text{ medible.}$$

  - Martingala: Es decir $\{X_t,\, t \geq 0\}$ es integrable si $\mathbb{E}(X_{t+s}|X_t)$ y el pasado no dice nada del futuro.
  - Trayectorias continuas o cadlag $X_t = B_t$ es movimiento browniano.
  - Incrementos Gaussianos: Para todo

Definición 2.2. *Una familia de funciones* $\mathfrak{H} = \{f : \Omega \to \mathbb{R}\}$ *es una clase monotona si*

- $1 \in \mathfrak{H}$
- $\forall a, b \in \mathbb{R}\, f, g \in \mathfrak{H}$, *entonces* $af + bg \in \mathfrak{H}$
- *contiene sus límites no decrecientes de funciones positivas y acotadas*

Teorema 2.1 (De clases monótonas funcional)**.** *Si una clase de funciones* $\mathfrak{H}$ *es una clase monótona y* F *es un* π *sistema que genera a* Σ*, tal que* $1_A \in \mathfrak{H}$ *para todo* $A \in \mathrm{F}$*, entonces* $\mathfrak{H}$ *contiene a todas las funciones medibles y acotadas.*

Demostración. Sea $\mathrm{G} = \{A \in \mathrm{F} \mid 1_A\}$, asi $\mathrm{F} \subset \mathrm{G}$, veamos que G es un λ sistema

- $1_\Omega \in \mathfrak{H}$ claramente, así $\Omega \in \mathrm{G}$
- Sea $A \subset B$ conjuntos en G, como $1_{B\setminus A} = 1_B - 1_A \in \mathfrak{H}$ por lienalidad así $B \setminus A \in \mathfrak{H}$
- Sea $\{A_n\}$ una sucesion creciente de conjuntos en $\mathfrak{H}$ y $A = \bigcup A_n$ como 1_A es acotada y límite creciente de 1_{A_n} así $A \in \mathfrak{H}$

Por el teorema de clases monónonas, por hipótesis $\mathfrak{H}$ contiene las funciones características de todo conjunto medible y por linealidad todas las funciones simples y por monotonía todas las funciones medibles acotadas.

Observación. En el teorema anterior se puede precindir de la hipótesis de acotamiento.

Corolario 2.2. *Tarea Sea* $X : (\Omega, \Sigma, \boldsymbol{P}) \to (E, \xi)$ *medible y* $\mathrm{T} \subset \Sigma$ *una sub sigma álgebra genrada por una familia numerable de funciones* $\{Y_n\}$*, donde* $Y_n : (\Omega, \Sigma, \boldsymbol{P}) \to (Z, \tau)$*. Sea* $f : E \to \mathbb{R}$ *medible, entonces una variable aleatoria* U *que es* T *medible y acotada es* $U = \mathbb{E}(f|\mathrm{T})$ *si y sólo si para todo* $n \in \mathbb{N}$ *y* $f_1, \ldots, f_n : (Z, \mathrm{T}) \to \mathbb{R}$ *medibles y acotadas, se cumple*

$$\mathbb{E}(f(X) \prod f_i(Y_i)) = \mathbb{E}(U \prod f_i(Y_i))$$

Observación. $X : (\Omega, \Sigma, \mathbf{P}) \to (E, S)$ una variable aleatoria y $\mathrm{T} \subset \Sigma$ sub álgebra, entonces la esperanza condicional de una función medible $f : E \to \mathbb{R}$ siempre está bien definida por medio de

$$\mathbb{E}(f(X)|\mathrm{T})$$

Definición 2.3. *Dada una* ***v.a*** $X : \Omega \to E$ *y* $mathrmT \subset \Sigma$ *es una sub álgebra, definimos la ley* X *condicionada a* T *como*

$$\boldsymbol{P}_{X|\mathrm{T}}(B) = \mathbb{E}(1_B|\mathrm{T}).$$

Observación. Para cada $B \in S$ fija, la función $\mathbb{E}(1_B|\mathrm{T})(w)$ es una **v.a** T medible, pero no necesariamente una probabilidad.

Definición 2.4. *Sean* (E, S) *y* (F, T) *espacios medibles, decimos que una función* $Q : E \times T \to \mathbb{R}$ *es una medida de transisción si cumple*

- $\forall\, B \in T$ *fija* $Q(w, B)$ *es* S *medible*
- $\forall\, w \in E$ *fija* $Q(w, B)$ *es una medida en* (F, T)

Si $Q(w, F) = 1$ *para toda* w*, decimos que* Q *es una probabilidad de transición.*

Definición 2.5. *Decimos que* $\boldsymbol{P}_{X|\mathrm{T}}(\cdot)$ *si existe una probabilidad de transición* Q *tal que*

$$\mathbb{E}(1_B(X)|\mathrm{T})(w) = Q(w, B)$$

Recordamos que un espacio medible es separable si la σ álgebra se puede definir por un número numerable de conjuntos, el ejemplo prototípico son los reales.

Teorema 2.3 (Teorema de Jirina)**.** *Sea* $(\Omega, \Sigma, \boldsymbol{P})$ *un espacio de probabilidad y* $\mathrm{T} \subset \Sigma$ *una sub álgebra, sea* $X : \Omega \to E$ *una *v.a* tal que* (E, S) *sea separable, entonces existe una probabilidad de transición* $Q : \Omega \to E$ *tal que* $Q(w, B)$ *es una esperanza condicional*

Demostración. Lo demostramos para $E = \mathbb{R}$. Sea $r \in \mathbb{Q}$ y definimos $C_r = \mathbb{E}(1_{\{x \le r\}}|\mathrm{T})$, así claramente C_r es una **v.a** T medible, entonces por monotonia $C_r \le C_q$ si $r < q$, asi si definimos

$$\Omega_{r,q} = \{\omega \in \Omega \,|\, C_r \le C_q\},$$

asi $\Omega_{r,q} \in \mathrm{T}$ y $\mathbf{P}(\Omega_{r,q}) = 1$. Entonces $\Omega_0 = \bigcap_{r<q\in\mathbb{Q}} \Omega_{r,q} \in \mathrm{T}$, además $\mathbf{P}(\Omega_0) = 1$. Ahora para $w \in \Omega_0$, la función $C_\cdot : \mathbb{Q} \to \mathbb{R}$ es no decreciente y no negativa, por lo que es posible extenderla a todo $\mathbb{R}$. Notamos que $C_\cdot$ está acotada por 1 y es continua a la izquierda y los límites por la derecha existen, por lo tanto $w \in \Omega_0$ y por lo tanto $C_\cdot$ es una distribución en $\mathbb{R}$. Ahora definimos

$$F(w,x) := 1_{\Omega_0} C_x + 1_{\Omega_0}(w) 1_{x\geq 0}.$$

Para cada $w \in \Omega$ fija, $F(w,x)$ es una distribución, entonces existe una medida de probabilidad $Q(w,\cdot)$ en $\mathbb{R}$ tal que $F(w,x) = Q(w,(-\infty,x])$. Así para $x \in \mathbb{R}$ fija $F(w,x)$ es una **v.a** T medible. Enconces como $\mathrm{C} = \{(-\infty,x] \,|\, x \in re\}$ es un π sistema, así por el treorema de clases monóntanas, entonces $Q(x,B)$ es una probabilida de transición, sólo falta ver que es una esperanza condicional. Si $B = (-\infty, r]$ con $r \in \mathbb{Q}$.

2.2. Construccion de medidas de probabilidad a partir de dos probabilidades.

- Una construccion es a travez de la medida producto.
- Otra es con la probabilidad de transición.

Definición 2.6. *Sean (E,S) y (F,T) espacios medibles y sea μ una probabilidad en (E,S) y $Q : E \times T \to \mathbb{R}$ una probabilidad de transicion de E a F. Definimos una probabilidad en $(E \times F, S \otimes T)$ como*

$$\pi(A \times B) = \int_E \mu(dx) 1_A(x) Q(x,B) = \int_E \mu(dx) 1_A(x) \int_F Q(x,dy) 1_B(y)$$

El teorema de clases monónotonas funcional $f, : E \times F \to \mathbb{R}$ medible, se tiene que

$$\int f d\pi = \int_{E\times F} \pi(dx,dy) f(x,y) = \int_E \mu(dx) \int_F Q(x,dy) f(x,y).$$

Dicho de otra manera tenemos,

$$\mathbf{P}_1(dx)\mathbf{P}_2(dy) = \pi(dx.dy) = \mu Q(x,dy),$$

lo cual se puede interpretar como que π es la drisribucion conjunta de X y Y, μ la distribución de X y $Q(x,dy)$ la probabilidad condicional de Y dado X. Existe la versión inversa de este problema y se conoce como el problema de desintegración.

Teorema 2.4 (Teorema de desintegración)**.** *Sea π una probabilidad en $(E \times F, S \otimes T)$, supongamos que T es separable, entonces existe una medida de probabilidad μ en (E,S) y una probabilidad de transición Q entre (E,S) a (F,T) tal que*

$$\pi(dx,dy) = \mu(dx) Q(x,dy)$$

Aplicamos inducción para encontrar un espacio con n variables aleatorias y la extensión numerable de esto fue realizada por *Clonescu-Tulcea* y la extensión no numerable fue hecha por *Kolmogorov* la cual llevó a la definición de la propiedad de Kolmogorov. Sea T n conjunt de parámetros (el cual usualmente lo pensaremos como tiempo) y sea (E_t, Σ_t) un espacio medible para todo $t \in T$.

Definición 2.7. *Definimos el espacio producto como $\prod_{t\in T} E_t$ como un producto de Tychonoff, es decir la sigma álgebra se define por medio de los conjuntos llamados **rectángulos***

$$\times_{t\in T} A_t = \{x \in E_t \,|\, x_t \in A_t \forall t \in T\},$$

donde $A_t \in \Sigma$ y además requerimos que $A_t = E_t$ para a lo más una cantidad finita de $t \in T$. Así denotamos por $\bigotimes_{t\in T} \Sigma_t$ a la sigma álgebra generada por los rectángulos.

Observamos que $\bigotimes_{t\in T} \Sigma_t$ es la sigma álgebra más chica generada por las proyecciones $\pi_s : \prod_{t\in T} E_t \to E_s$. Ahora sólo tenemos que definir una probabilidad en este espacio. Primero suponemos que $T = \mathbb{N}$, así tomamos μ una probabilidad en (E_0, Σ_0) y sea κ_1 una probabilidad de transición de (E_0, Σ_0) a (E_1, Σ_1), luego inductivamente definimos κ_{n+1} como la probabilidad de transición de $(F^0, \Sigma^0) = (E_0 \times \cdots \times E_n, \Sigma_0 \otimes \cdots \otimes \Sigma_n)$ a (E_{n+1}, Σ_{n+1}), así $\mu(A)$ es la probabilidad de que ocurra A y dado $(x_0, \ldots, x_n)$, $\kappa_{n+1}((x_0, \ldots, x_n); A)$ modela la probabilidad de que el $n+1$ resultado esté en A dado que ya sucedió $(x_0, \ldots, x_n)$. Sean Y_n los primeros n resultados, entonces Y_n es un elemento de (F^0, Σ^0), el cual tiene medida de probabilidad

$$\pi_n(dx_0, \ldots, dx_n) = \mu(dx_0)\kappa_1(x_0; dx_1)\kappa_2(x_0, x_1; dx_2) \ldots \kappa_n(x_0, \ldots, x_n; dx_n).$$

Por lo que queremos un espacio de probabilidad $(\Omega, \Sigma, \mathbf{P})$ donde todas las (Y_n) sean elementos. Si definimos

$$(\Omega, sig) = \prod_{n\in\mathbb{N}} (E_n, \Sigma_n),$$

entonces definimos $Y_n : (\Omega, \Sigma) \to (F^0, \Sigma^0)$ como la proyeccion $Y_n(\omega) = (\omega_1, \ldots, \omega_n)$, así definimos a la sigma álgebra generada por Y_n como F_n, notamos que si $A \in \mathrm{F}_n$, entonces exite un $B \in \Sigma^0$ tal que $A = B \times E_{n+1} \times \ldots$. Ahora queremos una $\mathbf{P}$ tal que $\mathbf{P}(A) = \pi_n(B)$, al conjunto $\{\pi_n(B) \,|\, n \in \mathbb{N}\, B \in \mathrm{F}\}$ se les llama *disribuciones finito dimenionales.*

Teorema 2.5. *Existe una única medida de probabilidad $\boldsymbol{P}$ es (Ω, Σ) tal que $\boldsymbol{P}(A) = \pi_n(B)$ para todo $B \in \mathrm{F}_n$*

Demostración. Definimos $\Sigma^0 = \bigcup_{n\in\mathbb{N}} \mathrm{F}_n$, entonces probamos lo siguiente:

1. Σ^0 es un álgebra
2. Σ^0 es un π sistema

3. $\sigma(\Sigma^0) = \Sigma = \sigma(\textbf{rectangulos})$

#TERMINAR: VER FOTOS EN EL ONEPLUS y ver Kallenberg
Ejemplo. Se puede afirmar la existencia de una infinidad de variables aleatorias independientes. Para cada $n \in \mathbb{N}$ definimos en $(\Omega_n, \Sigma_n, \mathbf{P}_n)$ el espacio de probabilidad por medio de

$$\kappa_{n+1}(x_0, \ldots, x_n; A) = \mathbf{P}_{n+1}(A).$$

Así existe un único espacio de probabilidad $(\Omega, \Sigma, \mathbf{P})$ espacio de probabilidad como previamente lo definimos y $X_n(\omega) = \omega_n$ con $\omega \in \Omega$
Ejemplo. Cadenas de Markov sea (E, F) espacio medible, así definimos $(E_n, \mathrm{F}_n) = (E, \mathrm{F})$ para toda n. Sea μ una probabilidad en (E, F) y tomamos una probabilidad de transición $Q(x, dy)$, entonces $\kappa_{n+1}(x_0, \ldots, x_n; A) = Q(x_n; A)$. Si queremos una cadena con espacio de estados finita nos dan $\mathbf{P}_{x,y} = \mathbf{P}(X_1 | x_0 = x,$ entonces

$$Q(x, A) = \sum_{y \in A} \mathbf{P}_{x,y}.$$

2.3. Construccion de procesos para tiempo continuo

Sea (E, F) un espacio medible separable y sea $J \subset T$ finito, entonces podemos construir un espacio de probabilidad $(E^J, \mathrm{F}^J, \pi^J)$. Ahora queremos construir un espacio de probabilidad $(\Omega, \Sigma, \mathbf{P})$ tal que $(\Omega, sig) = (\prod_{t \in T} E, \bigotimes_{t \in T} \mathrm{F})$, para esto necesitamos la condición de consistencia #ver fotos

Definición 2.8. *Decimos que una familia $\{\pi_J \mid J \subset T \text{ finito}\}$ es consistente si para todo $K \subset J \subset T$ com J finito $\pi_K(A) = \pi_J(R_{J,K})^{-1}(A))$*

Teorema 2.6. *Extensión de Kolmogorov Supongamos que (E, F) es un espacio medible separable con $\{\pi_J \mid J \subset K \text{ finito}\}$ consistente, entonces existe una única medida de probabilidad $\boldsymbol{P}$ en (Ω, Σ) tal que $\boldsymbol{P}(X_J \in A) = \pi_J(A)$ para todo $A \in \mathrm{F}^J$ con $J \subset T$ finito.*

Demostración. Definimos $I_f = \{J \subset T \mid J \text{ finito}\}$ y $I_c = \{J \subset T \mid J \text{ numerable infinito}\}$, tomamos $J \in I_c$, así $J = \{t_0, \ldots, t_n, \ldots\}$ y sea $J = \{t_0, \ldots, t_n\}$, inductivamente el teorema de desintegración implica que existe una medida $\mu_0(A) = \pi_{J_0}(A)$ $A \in \mathrm{F}$ y una sucesión de probabilidades de transición $\{\kappa_n\}_{n \in \mathbb{N}}$ de $(E^{J_{n-1}}, \mathrm{F}^{J_{n-1}})$ a (E, F), digamos

$$\pi_{J_n}(dx_0, \ldots, dx_n) = \mu_0(dx_0)\kappa_1(x_0; dx_1) \ldots \kappa(x_0, \ldots, x_{n-1}; dx_n),$$

entonces el teorema de Ionescu-Tulcea, existe $\mathbf{P}_J$ en (E^J, F^J) tal que $\mathbf{P}_J(R^{-1}_{J,J_n}(A)) = \pi_{J_n}(A)$ para toda $n \in \mathbb{N}$ y $A \in \mathrm{F}^{J_n}$. Así tenemos $\mathbf{P}_J$ para toda $J \in I_c$, veamos que es consistente, si $L \subset J$ con $J, L \in I_c$, veamos que $\mathbf{P}_L(R^{-1}_{L,J}(A)) = \pi_L(A)$ para toda $A \in \mathrm{F}^L$, así sea $H \in \mathrm{F}^L$, entonces existe $n \in \mathbb{N}$ y $B \in L_n = \{s_0, \ldots, s_n\}$ tal que $H = R^{-1}_{L_n,L}(B)$, por lo tanto $\mathbf{P}_L(R^{-1}_{L,L_n}(B)) = \pi_L(H) = \pi_{L_n}(B)$, ahora como $L \subset J$, exite $m \in \mathbb{N}$ talque $L_n \subset J_m$ y por consistencia tenemos

$$\pi_{L_n}(R^{-1}_{J_m,L_n}(B)) = \pi_{L_n}(B) = \mathbf{P}_J(R^{-1}_{J,J_m} R^{-1}_{J_m,L_n}(B)) = \tag{5}$$

$$\mathbf{P}_J(R^{-1}_{J,L_n}(B)) = \mathbf{P}_J(R^{-1}_{J,L} R^{-1}_{L,L_n}(B)) = \mathbf{P}_J(R^{-1}_{J,L}(H)). \tag{6}$$

como la familia de rectangulos genera a F^L, entonces la ecuaciones anteriores son válidas para todo $H \in \mathrm{F}^L$. Ahora en el caso general por propieadades básicas de σ álgebras, si $H \in \Sigma$, entonces existe $J \in I_c$ tal que para toda $t \in J$ $A_t = E$, es decir sus colas son totales y $H = R^{-1}_{T,J}(B)$ con $B \in \mathrm{F}^J$. Así definimos $\mathbf{P}(H) = \mathbf{P}_J(B)$ y por consitencia esto es suficiente pues la familia de rectángulos es un π sistema que genera a Σ.
Ejemplo. Familias de Markov a tiempo continuo Sea $T = [0, \infty)$ y μ una probabilidad en (E, F), para cada $s < t \in \mathbb{R}^+$ tomamos una probabilidad de transición $\mathbf{P}_{s,t}$ de (E, F) en sí mismo y llamamos μ_0 a la probabilidad inicial y diremos que $\mathbf{P}_{s,t}(x; A)$ es la probabilidad condicional a que $X_t \in A$ dado $X_s = x$. Tomamos $J = (0 < t_1 < t_2 \ldots < t_n)$ y definimos

$$\pi_J(dx_0, \ldots, dx_n) = \mu_0(dx_0)\mathbf{P}_{0,t_1}(x_0; dx_1) \ldots \mathbf{P}_{t_{n-1},t_n}(x_{n-1;dx_n}).$$

La consistencia en este caso equivale a las *ecuaciones de Chapman-Kolmogorov*

$$\mathbf{P}_{s,t}(x; B) = \int_E \mathbf{P}_{s,u}(x, dy)\mathbf{P}_{u,t}(y; B)$$

$B \in \mathrm{F}$, $s < u < t$. Si definimos $\mathrm{F}_t = \sigma(X_u \mid u \leq t)$ como

$$X_u : (\Omega, sig) \to (E, \mathrm{F}) \quad X_u(\omega) = \omega_0$$

entonces $\mathbf{P}(X_0 \in A) = \mu(A)$ y $\mathbf{P}(X_t \in A \mid \mathrm{F}_s) = \mathbf{P}_{s,t}(X_s, A)$ lo cual interpretamos como la probabilidad de si en el tiempo pasado s, ¿Cuál es la probabilidad de al tiempo t estar en A? Le llamamos a $A(X_t, x \geq 0)$, *procesos de Markov con probabilidades de transición.* Si $\mathbf{P}_{s,t} = \mathbf{P}_{t-s}$, es decir que sólo depende de la distancia, le llamamos *proceso de Markov homogeneo* y en este caso las ecuaciones de Chapman-Kolmogorov se reducen a

$$\mathbf{P}_t\mathbf{P}_s = \mathbf{P}_s\mathbf{P}_t = \mathbf{P}_{s+t} \quad \{\mathbf{P}_t\}_{t\in\mathbb{R}^+} \text{ semigrupo de transición.}$$

Usualmente esto se hace con **matrices de Markov**

3. Martingales

Esta teoría intuitivamente se puede pensar como la teoría de "juegos justos", es decir son juegos donde no puedes ganar dinero apostando.

Definición 3.1. *Martingalas Sea* $\{X_n \mid n \in \mathbb{N}\}$ *una sucesión de* ***v.a****, decimos que ésta es* martingala, submartingala o subramartingala *si*

$$\mathbb{E}(X_n|\{X_0, \ldots, X_m\}) = X_m \quad \mathbb{E}(X_n|\{X_0, \ldots, X_m\}) \geq X_m \quad \mathbb{E}(X_n|\{X_0, \ldots, X_m\}) \leq X_m$$

respectivamente.

Las siguientes son propiedades últiles de las martingalas

- Si N es un tiempo finito, entonces $\mathbb{E}(X_N) = \mathbb{E}(X_0)$
- Si $\{X_n\}$ es submartingala tal que $\sup_{n\in\mathbb{N}} \mathbb{E}(X_n^{pm}) < \infty$, emtonces converge **c.d.**

Definición 3.2. *Filtración Sea $(\Omega, \Sigma, \boldsymbol{P})$ un espacio de probabilidad, una filtración es una familia creciente de sub σ álgebras, es decir*

$$\{\mathrm{F}_n \subset \Sigma\}_{n\in\mathbb{N}} \quad \mathrm{F}_m \subset \mathrm{F}_n \, n < m.$$

Definición 3.3. *Sea $(\Omega, \Sigma, \boldsymbol{P})$ un espacio de probabilidad, $\{\mathrm{F}_n \subset \Sigma\}_{n\in\mathbb{N}}$ una filtración y $\{X_n\}_{n\in\mathbb{N}}$ una familia de **v.as**, decimos que esta familia es*

- *Adaptada si cada X_n es F_n medible*
- *Subadaptada si cada X_n es F_{n-1} medible*

Observación. Sea $(\Omega, \Sigma, \mathbf{P})$ un espacio de probabilidad y $\{X_n\}_{n\in\mathbb{N}}$ una familia de **v.as**, podemos definir una filtración adaptada llamada *filtración natural* por medio de

$$\mathrm{F}_n = \sigma(X_0, \ldots, X_n)$$

Definición 3.4. *Sea $(\Omega, \Sigma, \boldsymbol{P})$ un espacio de probabilidad, $\{\mathrm{F}_n \subset \Sigma\}_{n\in\mathbb{N}}$ una filtración. Decimos que una familia de **v.as** $\{X_n\}_{n\in\mathbb{N}}$, es martingala con respecto a $\{\mathrm{F}_n\}_{n\in\mathbb{N}}$ si*

- *Si $\{X_n\}$ es adaptada a $\{\mathrm{F}_n\}$*
- $\mathbb{E}(|X|_n) < \infty$
- $\mathbb{E}(X_n|\mathrm{F_m}) = X_m$ *para cada $m \leq n$ (adaptar para sub y supra martingala)*

Observación. :

1. $\{X_n\}$ es martingala si y sólo si es sub y supra martingala.
2. Si $\{X_n\}$ es submartingala, entonces $\{-X_n\}$ es supra martingala.
3. Como una martingala es F_n adaptada, entonces $\mathbb{E}(X_n 1_A) = \mathbb{E}(X_m 1_A)$ para toda $A \in \mathrm{F}$.
4. Si $\{X_n\}$ es submartingala, entonces $\mathbb{E}(\mathbb{E}(X_n|\mathrm{F}_m)) = \mathbb{E}(X_n)$ con $m < n$ de forma creciente, el resultado es similar para supramartingala.
5. Demostrar que $\{X_n\}$ es martingala, basta ver que es adapatada e integrable con $\mathbb{E}(X_{n+1}|\mathrm{F}_n) = X_n$ para toda n por recursión sobre m y la propiedad de la torre, el resultado para sub y supramartingala es análogo.

Teorema 3.1. *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una* $\{\mathrm{F}_n\}$ *maringala y sea* $\varphi : \mathbb{R} \to \mathbb{R}$ *convexa creciente tal que* $\mathbb{E}(\varphi(X_n)) < \infty$*, entonces* $\{\varphi(X_n)\}_{n\in\mathbb{N}}$ *es submartingala*

Demostración. Como φ es convexa creaciente, es continua y por lo tanto Borel medible así $\varphi(X_n)$ es F_n medible, entonces por la desigualdad de Jensen

$$\mathbb{E}(\varphi(X_{n+1})|\mathrm{F}_n) \geq \varphi(\mathbb{E}(X_{n+1}|\mathrm{F}_n)) = \varphi(X_n)$$

ya que $\{X_n\}$ es martingala, así si φ es creciente, tenemos el resultado.

Corolario 3.2. *Si* $\{X_n\}$ *es martingala y* $\mathbb{E}(|X_n|^p) < \infty$*, entonces*

- $\{X_n^p\}$ *es submartingala*
- $\{|X_n|\}$ *es submartingala*
- $\{X_n^+\}$ *es submartingala*

Ejemplo.

1. Si X es una **v.a** integrable y $\{\mathrm{F}_n\}$ es una fitración, entonces $X_n = \mathbb{E}(X|\mathrm{F}_n)$ es maringala.

2. **Caminata aleatoria simétrica**: Sean $\{\epsilon_i\}$ **v.as** independientes e integrables de media nula $\mathbb{E}(\epsilon_i) = 0$, entonces

$$X_n = \sum_{i=0}^{n} \epsilon_i \quad \text{es martingala con respecto a } \mathrm{F}_n = \sigma(\epsilon_0, \dots, \epsilon_n)$$

Claramente X_n es $mathrmF_n$ medible además sucede que

$$\mathbb{E}(|X_n|) \leq \sum_{i=1}^{n} \mathbb{E}(|\epsilon_i|) < \infty.$$

Ahora si $m < n$, sucede que

$$\mathbb{E}(X_n|\mathrm{F}_m) = \mathbb{E}(\sum_{i=1}^{m} \epsilon_i + \sum_{i=m}^{n} \epsilon_i|\mathrm{F}_m) = X_m + \sum_{i=m} \mathbb{E}(\epsilon_i) = X_m$$

Si $\{\epsilon_i\}$ son independientes, integrables y $\mathbb{E}(\epsilon_i) < 0$ o $\mathbb{E}(\epsilon_i) > 0$,entonces

$$X_n = \sum_{i=0}^{n} \epsilon_i \quad \text{es martingala o submartingala respectivamente.}$$

Más general si $\{\epsilon_i\}$ son variables aleatorias independientes e integrables entonces

$$X_n = \sum_{i=1}^{n} \epsilon_i - \mathbb{E}(\epsilon_i) \quad \text{es martingala.}$$

3. Sean $\{\epsilon_i\}_{i\in\mathbb{N}}$ **v.a** independientes e intregrables con la misma distribución (resumido usualmente como **i.i.d**). Sea $\mu_i = \mathbb{E}(\epsilon_i)$ y $\sigma_{\epsilon_i}^2 = \mathbb{E}((\epsilon_i - \mu_i)^2)$ la esperanza y *la varianza* de la variable ϵ_i respectivamente, laos cuales resumimos en $\mu = \sum_{i=0}^{n} \mu_i$ y $\sigma^2 = \sum_{i=0}^{n} \sigma_i^2$, entonces

$$Y_n = (\sum_{i=0}^{n} \epsilon_i - n\mu_i)^2 - n\sigma^2 \quad \text{es martingala.}$$

Con respecto a $\mathrm{F}_n = \sigma(\epsilon_0, \dots, \epsilon_n)$, esto es pues claramente Y_n es F_n medilbe además de que

$$\mathbb{E}(|Y_n|) \leq \sum_{i=0}^{n} \mathbb{E}((\epsilon_i - n\mu)^2) + n\sigma^2 = (2n+1)\sigma^2 < \infty.$$

Ahora si $m \leq n$, calculamos

$$\mathbb{E}(Y_n|\mathrm{F}_m) = \mathbb{E}((S_m + \sum_{i=m+1}^{n} \epsilon_i - n\mu)^2 - n\sigma^2|\mathrm{F}_m),$$

donde $S_m = \sum_{i=0}^{m} \epsilon_i$, así si expandimos el binomio al cuadrado $((S_m - m\mu) + (\sum_{i=m+1}^{n} \epsilon_i - (n-m)\mu))^2$, entonces

$$\begin{aligned}
&\mathbb{E}(Y_n|\mathrm{F}_m) \\
&= \mathbb{E}((S_m - m\mu)^2 + 2(S_m - m\mu)(\sum_{i=m+1}^{n} \epsilon_i - (n-m)\mu) + (\sum_{i=m+1}^{n} \epsilon_i - (n-m)\mu)^2 - n\sigma|\mathrm{F}_m) \\
&= (S_m - m\mu)^2 + 2(S_m - m\mu)\mathbb{E}(\sum_{i=m+1}^{n} \epsilon_i - (n-m)\mu)) + \mathbb{E}((\sum_{i=m+1}^{n} \epsilon_i - (n-m)\mu)^2) - n\sigma^2 \\
&= Y_n,
\end{aligned}$$

ya que $S_m - m\mu$ es F_m medible y $\sum_{i=m+1}^{n} \epsilon_i - (n-m)\mu$ es independiente a F_m y por lo tanto $\{Y_n\}_{n\in\mathbb{N}}$ es martingala.

4. **Urna de Polya**:

##TAREA:

Ejemplo. Sea $\{X_n\}$ una maringala o submartingala y $\{\mathrm{F}_n\}$ una filtración con T un tiempo de paro asociado. Veamos que

$$\mathrm{F}_T = \{A \in \Sigma \mid A \cap \{T \leq n\} \in \mathrm{F}_n \ \forall n \in \mathbb{N}\}.$$

es una sub σ álgebra llamada la σ álgebra detenida. Más aún se cumple que

- T es F_T medible
- Si T es finito, entonces X_T es F medible.

- Si $S \leq T$ es un tiempo de paro acotado, entonces $\mathrm{F}_S \subset \mathrm{F}_T$ y X_S, X_T son $L_1(\Omega, \Sigma, \mathbf{P})$ y
$$\mathbb{E}(X_T|\mathrm{F_s}) = X_S$$
.

- Más aún si $\{X_n\}$ es un proceso $\{\mathrm{F}_n\}$ adaptado con $X_n \in L_1(\Omega, \Sigma, \mathbf{P})$ para todo $n \in \mathbb{N}$ y para todo $S \leq T$ tiempos de paro acotados se cumple $\mathbb{E}(X_S) = \mathbb{E}(X_T)$, entonces X es martingala.

3.1. Desigualdades Maximales

##TERMINAR(ver fotos) Dado un proceso $\{X_n\}$ podemos definir el proceso maximal
$$X_n^* = \max_{0 \leq m \leq n} X_m. \tag{7}$$

Teorema 3.3. *Desigualdad Maximal de Doob Sea $\{X_n\}$ submartingala no negativa, entonces para toda $\lambda \in \mathbb{R}^+$*
$$\lambda \boldsymbol{P}(X_n^* \geq \lambda) \leq \mathbb{E}(X_n 1_{X_n^* \geq \lambda}) \leq \mathbb{E}(X_n).$$

Demostración. Sea $n \in \mathbb{N}$ fija, entonces definimos
$$T := \min_{0 \leq k \leq n} \{X_k \geq \lambda\},$$
y definimos S el mínimo entre T y n, entonces S es un tiempo de paro acotado aplicado. Como X_n es submartingala, entonces $\mathbb{E}(X_S) \leq \mathbb{E}(X_n)$. Notamos que $\{S < T\}$ implica $T = \infty$ y $S = n$, así para toda $0 \leq k \leq n$, $X_k < \lambda$ y $X_k^* < \lambda$. Ahora si $\{S = T\} = \{S = n\} \cap \{X_n^* < \lambda\}$, entonces
$$\mathbb{E}(X_n) \geq \mathbb{E}(X_S) = \mathbb{E}(X_S 1_{S<T}) + \mathbb{E}(X_S 1_{S=T}) \geq \mathbb{E}(X_n 1_{X_n^* < \lambda}) + \lambda \mathbf{P}(X_n^* \geq \lambda).$$

Teorema 3.4. *Desigualdad L_p de Doob*

Demostración. Recordamos que
$$\mathbb{E}(Y^p) = \int p y^{p-1} \mathbf{P}(Y \geq y) dy.$$

Así si $Y = X_n^*$, por Fubini y la desigualdad maximal de Doob obtenemos
$$\mathbb{E}((X_n^*)^p) = \int p y^{p-1} \mathbf{P}(X_n^* \geq y) dy \leq . \int p y^{p-1} y^{-1} \mathbb{E}(X_n 1_{X_n^* \geq y}) dy \int = p y^{p-1} y^{-1} \mathbb{E}(X_n 1_{X_n^* \geq y}) dy.$$

Con Hölder y despejando sale.

3.1.1. Convergencia de submartingalas.

#TERMINAR(ver fotos)

Definición 3.5. *Sea* $\{\mathrm{F}_n\}_{n\in\mathbb{N}}$ *una sucesión decreciente de* σ *álgebras decimos que sucesión de variables aleatorias* $\{X_n\}$*, es* martingala reversa *si se cumple lo siguiente*

- X_n *es* F_n *medible*
- $X_n \in L_1(\Omega, \Sigma, \boldsymbol{P})$
- $\mathbb{E}(X_n|\mathrm{F}_{n+1}) = X_{n+1}$

Observación. $X_n = \mathbb{E}(X_0|\mathrm{F}_n)$

Teorema 3.5. *Convergencia de martingalas reversas Sea* X_n *una* $\{\mathrm{F}_n\}$ *martingala reversa, entonces* X_n *converge casi siempre y en* L_1 *a* $X = \mathbb{E}(X_0| \cap \mathrm{F}_n)$

Demostración. Denotamos por $D_{[a,b]}(X_0, X_1, \dots)$ al número de cruces hacia abajo en el intervalo $[a,b]$ por las variables $\{X_n\}$, más aún $D^n_{[a,b]}(X_0, \dots, X_n)$ es el número de cruces hacia abajo al tiempo n. Sea $m \in \mathbb{N}$ y $Y_n^{(m)} = -X_{\text{máx}\{m-n,0\}}$ y $\mathrm{H}_n^{(m)} = \mathrm{F}_{\text{máx}\{m-n,0\}}$. Notemos que $\{\mathrm{H}_n^{(m)}\}$ es filtración, con esto, vemos que $\{Y_n^{(m)}\}$ es martingala. Sea $n \in \mathbb{N}$

$$\begin{aligned}\mathbb{E}(Y_{n+1}^{(m)}|\mathrm{H}_n) =& \mathbb{E}(-X_{m-n+1}^{(m)}|\mathrm{F}_{m-n}) \\ & \mathbb{E}(X_0|\mathrm{F}_0) = -X_{m-n} \\ = -X_0.\end{aligned}$$

notemos que $D^n_{[a,b]}(X_0, \dots, X_n) = D^n_{[a,b]}(Y_0, \dots, Y_n)$, así por la desigualdad de Doob para $\{Y_n^{(m)}\}$, tenemos

$$\mathbb{E}(D^n_{[a,b]}(X_0, \dots, X_n)) = \mathbb{E}(D^n_{[a,b]}(Y_0, \dots, Y_n)) \leq \frac{1}{(b-a)}(\mathbb{E}(|Y_n^{(m)}|) + |b|] = \frac{1}{(b-a)}[\mathbb{E}(|X_0|) + |b|) < \infty$$

así D es finita casi dondequiera e integrable, por lo tanto existe $X = \text{lím}_{n\to\infty} X_n$, además $\{X_n = \mathbb{E}(X_0|\mathrm{F}_n)\}$ es uniformemente integrable y la convergencia de X es en L_1. Falta ver que $X = \mathbb{E}(X_0| \cap \mathrm{F}_n)$. Como $\mathrm{F}_n \subset \mathrm{F}_m$ si $m≤ n$y sabemos que X_n es F_n medible, entonces es F_m medible y por lo tanto para toda $m \leq n$ $X=lím$_{\mathrm{n}\to\infty} X_{n+m}$, así X es F_m medible para toda m y por lo tanto es $\cap \mathrm{F}_n$ medible. ☐

Observación. Sea $A \in \mathrm{F}_m$, por lo tanto $\mathbb{E}(X_0 1_A)\mathbb{E}(X_n 1_A) \to \mathbb{E}(X 1_A)$ y por lo tanto X es esperanza condicional de X_0 con respecto a $\cap \mathrm{F}_n$.

3.2. Extensión a tiempos continuos

#Ver fotos

Definición 3.6. *Una filtración a tiempo continuo es una familia de sub* σ *álgebras* $\{\mathrm{F}_t\}_{t\in\mathbb{R}^+}$ *tales que* $\mathrm{F}_s \subset \mathrm{F}_t$ *para* $s \leq t$

Definición 3.7. *Una familia de sub* σ *álgebras* $\{\mathrm{F}_t\}_{t\in\mathbb{R}^+}$ *es continua a la derecha si* $\mathrm{F}_t = \cap_{s\geq t}\mathrm{F}_s$.

3.3. Cadenas de Markov

Definición 3.8. *Sea* $X = \{X_n\}_{n\in\mathbb{N}}$ *un proceso, definimos* $\mathrm{F}_n = \sigma(X_0,\dots,X_n)$ *y* $\mathrm{G}_n = \sigma(X_n, X_{n+1}\dots)$*, decimos que* X *cumple con la propiedad si para toda variable* Y *que sea* G_n *medible y acotada, se cumple*

$$\mathbb{E}(Y|\mathrm{F}_n) = \mathbb{E}(Y|X_n)$$

Proposición 3.6. *Sea* $X = \{X_n\}_{n\in\mathbb{N}}$ *un proceso con espacio de estado numerable, entonces* X *cumple con la propiedad de Markov si y sólo si para toda* $n \in \mathbb{N}$ *y* $\{i_0,\dots,i_n\} \subset E$ *tales que* $\boldsymbol{P}(i_0,\dots,i_n) > 0$ *se cumple*

$$\boldsymbol{P}(i_{n+1}|i_0,\dots,i_n) = \boldsymbol{P}(i_{n+1}|i_n)$$

Demostración. **Clases monótonas**

Definición 3.9. *Para cada* $n \in \mathbb{N}$ *definimos* $P_{i,j} = \boldsymbol{P}(X_n = j|X_{n-1}i)$

Lema 3.7. *Sea* $\{X_n\}$ *una cadena de Markov. Si* $\{i_0,\dots,i_n\} \subset E$

$$\boldsymbol{P}(i_0,\dots,i_n) = \boldsymbol{P}(i_0)P^{(1)}_{i_0,i_1}P^{(1)}_{i_0,i_1}\dots P^{(1)}_{i_0,i_1}$$

Demostración.

$\boldsymbol{P}(i_0,\dots,i_n) =$ *haz los pasos intermedios usando que es Markov para separar las probabilidades.* $\boldsymbol{P}(i_0)P^{(1)}_{i_0,i_1}P^{(1)}_{i_0,i_1}$

#TERMINAR(creo que faltó una clase, checa el libro)

3.4. Propiedad de Markov y propiedad de Markov fuerte

Observación. Denotamos por $Markov(\lambda p)$ a una cadena de Markov con dirtibución inicial λ y matriz de transición P.

Teorema 3.8. *Sea* $\{X_n\}$ *una cadena de Markov* $Markov(\lambda p)$*, entonces condicionalmente a* X_m*, el proceso* $\{X_{n+m}\}$ *es* $Markov(\delta_i p)$ *e indepandiente de* $\{X_0,\dots X_m\}$*, es decir*

$$\mathbb{E}(f(X_{n+m})|\mathrm{F}_m) = \mathbb{E}_{X_m}(f(X_n)).$$

Demostración. Sea $A \in \mathrm{F}_m = \sigma(X_0,\dots,X_m)$ tenemos que demonstrar que

$$\mathbb{E}(f(X_{n+m})1_A|X_m) = \mathbb{E}(f(X_{n+m})X_m)\mathbf{P}(A).$$

Entonces por el teorema de clases monótonas

$$\mathbf{P}(X_0,\dots,X_m,X_{m+1},\dots,X_{n+m}) = \delta_i(j_0)\mathbf{P}(X_0,\dots,X_m)P_{j_0j_1}P_{j_1j_2}\dots P_{j_nj_{n+1}}.$$

así

$$\mathbf{P}(A, f(X_{n+m})|X_m) = \frac{1}{\mathbf{P}(X_m)\mathbf{P}()}\mathbf{P}(X_m,A)\mathbf{P}(A,X_m,X_{m+1},\dots,X_{n+m})\mathbf{P}(X_m,A)$$

así por la propiedad de Markov tenemos que

$$\mathbf{P}(X_m,A)\mathbf{P}(A,X_m,X_{m+1},\dots,X_{n+m})\mathbf{P}(X_m,A) = \delta_i(j_m)P_{j_nj_{n+1}}P_{j_{n+1}j_{n+2}\dots}P_{j_{n+m-1}j_{n+m}}$$

Teorema 3.9. *Propiedad de Markov fuerte Sea* $\{X_n\}$ *una cadena* $Markov(\lambda p)$ *y sea* T *un tiempo de paro. Entonces condicionalmente a* $T < \infty$ *y* $X_T = i$*, el proceso* $\{X_{n+T}\}$ *es* $Markov(\delta p)$ *y es independiente de* $\{X_0, \ldots, X_T\}$

Demostración. Sea $A \in \mathrm{F}$, entonces $A \cap \{T = m\} \in \mathrm{F}_m$. Por la propiedad de Markov para m tenemos

$$\mathbf{P}(X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n) \cap A \cap \{T = m\} \cap \{X_T = i\} = \mathbf{P}_i(X_0 = j_0, \ldots, X_n = j_n)\mathbf{P}(A \cap \{T = m\} \cap \{X_T = i$$

si sumamos sobre las m, obenemos

$$\mathbf{P}(\{X_T = j_0, \ldots, X_{T+n}\} \cap A \cap \{X_T\} \cap \{T < \infty\}) = \mathbf{P}_i(X_T = j_0, \ldots, X_{T+n} = j_n)\mathbf{P}(A \cap \{X_T\} \cap \{T < \infty\})$$

ahora si dividimos entre $\mathbf{P}(\{T < \infty\} \cap \{X_T = i\})$, tenemos

$$\mathbf{P}(\{X_T = j_0, \ldots, X_{T+n}\}, A, \{T < \infty\}, \{X_T = i\}) = \mathbf{P}_i(X_0 = j_0, \ldots, X_n = j_n)\mathbf{P}(A|, \{T < \infty\}, \{X_T = m\})$$

Así si nuestro proceso es en el espacio producto que definimos previamente, podemos definir una función de traslación $\Theta_m : \Omega \to \Omega$, donde $\omega \in \Omega$, entonces $\omega = (\omega_n)_{n\in\mathbb{N}}$, así $\Theta(\omega_n) = (\omega_{n+m})$ y por lo tanto tenemos que

$$\mathbb{E}(f \circ \Theta_m | \mathrm{F}) = \mathbb{E}_{X_m}(f)$$

asi para T, tenemos

$$\mathbb{E}(f \circ \Theta_T | \mathrm{F}) = \mathbb{E}_{X_T}(f)$$

si tomamos $f = 1_{X_0 = j_0, \ldots, X_{m+n} = j_n}$ tenemos el resultado. □

Ejemplo. (Principio de Reflexión) Sea S_n una caminata aleatoria simétrica $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$, donde $S_0 = 0$, $S_n = X_1 + \cdots + X_n$. Entonces para todo natural a, se tiene que

$$\mathbf{P}(\sup_{m \le n}\{S_m \ge a\}) \le \mathbf{P}(S_n \ge a)$$

Sea

Referencias

[1] Lawler G,F,. (2006) *Introduction to Stochastic Processes*, Chapman and Hall.

[2] Resnik, S (1992) *Adventures in Stochastic Processes* Birkhäuser.

[3] Kallenber, O,. (1997) *Foundations of Modern Probability*, Springer-Verlag.

[4] William, D. (1991) *Probability with Martingales* Cambridge University Press.

[5] Durrett, *Probability Theory and Examples*

[6] Kingman, J.F.C *Poisson Processes*, Oxford Studies in Probability.

[7] Revuz D., Yor M., *Continuous Martingales and Brownian Motions*, Max-Planck-Institut für Matematik, Germany.

[8] Hoffman, Jorgensen, *Probability with a view towards Statistics*, IMPA Monographs, Springer International Publishing, Switzerland.

[9] Billingsley P., *Probability and Measure*, Springer-Verlag.

[10] Chow, Y.S., Teicher (1980) *Probability Theory: Independence, Interchangability and Martingales*, Springer-Verlag.