INTEGRACIÓN ESTOCÁSTICA

Preliminares

Álgebras y σ-álgebras

Definición 1. *Sea $\mathbb{F}$ un conjunto y $\mathcal{G}$ una familia de subconjuntos de $\mathbb{F}$. Diremos que:*

i. $\mathcal{G}$ es cerrada bajo complementos si $A^c \in \mathcal{G}$ para cualquier $A \in \mathcal{G}$.

ii. $\mathcal{G}$ es cerrada bajo diferencias propias si $B - A \in \mathcal{G}$ para cualquier pareja $A, B \in \mathcal{G}$ tal que $A \subset B$.

iii. $\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) finitas si $\bigcup_{j=1}^{n} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{n} A_j \in \mathcal{G}$) para cualquier colección finita $A_1, A_2, \ldots, A_n$ de elementos de $\mathcal{G}$.

iv. $\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) infinitas numerables si $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{\infty} A_j \in \mathcal{G}$) para cualquier colección infinita numerable $A_1, A_2, A_3, \ldots$ de elementos de $\mathcal{G}$.

v. $\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) monótonas si $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{\infty} A_j \in \mathcal{G}$) para cualquier sucesión creciente (resp. decreciente) $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{G}$.

Definición 2 (álgebra). *Sea $\mathbb{F}$ un conjunto. Se dice que una familia $\mathcal{A}$ de subconjuntos de $\mathbb{F}$ es un álgebra si se satisfacen las siguientes condiciones:*

i. $\mathbb{F} \in \mathcal{A}$.

ii. $\mathcal{A}$ es cerrada bajo complementos.

iii. $\mathcal{A}$ es cerrada bajo uniones finitas.

Definición 3 ($\boldsymbol{\sigma}$-álgebra). *Sea $\mathbb{F}$ un conjunto. Se dice que una familia $\Im$ de subconjuntos de $\mathbb{F}$ es una σ-álgebra si se satisfacen las siguientes condiciones:*

i. $\mathbb{F} \in \Im$.

ii. $\Im$ es cerrada bajo complementos.

iii. $\Im$ es cerrada bajo uniones infinitas numerables.

Definición 4. *Sea $\mathbb{F}$ un conjunto y $\mathcal{G}$ una familia de subconjuntos de $\mathbb{F}$. Llamaremos* ***álgebra generada por*** *$\mathcal{G}$ a la más pequeña familia de subconjuntos de $\mathbb{F}$ que forme un álgebra y que contenga a todos los elementos de $\mathcal{G}$.*

Definición 5 (Intersección de $\boldsymbol{\sigma}$-álgebras). *Dado un conjunto $\mathbb{F}$ y una familia arbitraria de σ-álgebras de subconjuntos de $\mathbb{F}$, se define la intersección de esas σ-álgebras como la familia de conjuntos que pertenecen a todas ellas.*

Definición 6 ($\boldsymbol{\sigma}$ álgebra generada por una familia de conjuntos). *Dada una colección $\mathcal{A}$ de subconjuntos de un conjunto $\mathbb{F}$, se define la σ-álgebra generada por $\mathcal{A}$ como la intersección de todas las σ-álgebras que contienen a todos los conjuntos de $\mathcal{A}$. Denotaremos por $\sigma(\mathcal{A})$ a esta σ-álgebra.*

σ-álgebra de Borel en $\mathbb{R}$

Cuando hablemos de intervalos de números reales, vamos a incluir tanto a los finitos como a los infinitos. Explícitamente, el conjunto de los intervalos de números reales está formado por todos los de los tipos siguientes:

a) Los intervalos con extremos a y b, de cualquier tipo, donde $a, b \in \mathbb{R}$ y $a < b$.

b) Los intervalos de la forma $[a, a]$, $(-\infty, a)$, $(-\infty, a]$, (a, ∞) y $[a, \infty)$, donde a es un número real cualquiera.

c) El intervalo $(-\infty, \infty)$.

De acuerdo con lo anterior, no consideraremos como intervalo al conjunto vacío.

Cuando los dos extremos de un intervalo sean números reales, diremos que el intervalo es finito; en caso contrario, diremos que es infinito.

Definición 7 ($\boldsymbol{\sigma}$-álgebra de Borel en $\mathbb{R}$). *La σ-álgebra de Borel en $\mathbb{R}$, la cual será denotada por $\mathfrak{B}(\mathbb{R})$, es la σ-álgebra de subconjuntos de $\mathbb{R}$ generada por la familia de todos los intervalos de números reales. A los elementos de esa σ-álgebra los llamaremos borelianos de $\mathbb{R}$.*

RESULTADOS

La σ-álgebra de los conjuntos borelianos de $\mathbb{R}$ está generada por cualquiera de las siguientes familias de conjuntos.

a) Los intervalos de la forma $(-\infty, x]$, donde $x \in \mathbb{R}$.

b) Los intervalos de la forma $(-\infty, x)$, donde $x \in \mathbb{R}$.

c) Los intervalos de la forma $(a,b]$, donde $a,b \in \mathbb{R}$.

d) Los intervalos de la forma $[a,b)$, donde $a,b \in \mathbb{R}$.

e) Los intervalos de la forma (a,b), donde $a,b \in \mathbb{R}$.

f) Los intervalos de la forma $[a,b]$, donde $a,b \in \mathbb{R}$.

σ-álgebra de Borel en $\overline{\mathbb{R}}$

Vamos a trabajar con el conjunto de números reales extendidos, el cual consiste del conjunto de números reales y dos elementos especiales, $-\infty$ y ∞, con los cuales operaremos bajo las siguientes convenciones:

Si $c \in \mathbb{R}$, entonces:

$-\infty < c < \infty$,

$c - \infty = -\infty$,

$c + \infty = \infty$,

$c\,(\infty) = \infty$ si $c > 0$,

$c\,(\infty) = -\infty$ si $c < 0$,

$(0)\,(\infty) = (0)\,(-\infty) = 0$,

$\frac{c}{\infty} = \frac{c}{-\infty} = 0$,

$(\infty)\,(\infty) = \infty + \infty = \infty$,

$\infty - \infty$ e $\frac{\infty}{\infty}$ no están definidos.

$\overline{\mathbb{R}}$ denotará al conjunto $\mathbb{R} \cup \{-\infty, \infty\}$.

El conjunto de números reales $\mathbb{R}$ también será denotado por $(-\infty, \infty)$. El conjunto de números reales no negativos será denotado por $[0,\infty)$, o por $\mathbb{R}^+$, y $\overline{\mathbb{R}}^+$ denotará al conjunto $\mathbb{R}^+ \cup \{\infty\}$.

Definición 8 (**$\boldsymbol{\sigma}$-álgebra de Borel en $\overline{\mathbb{R}}$**). *La σ-álgebra de Borel en $\overline{\mathbb{R}}$, la cual será denotada por $\mathfrak{B}\left(\overline{\mathbb{R}}\right)$, es la σ-álgebra de subconjuntos de $\overline{\mathbb{R}}$ generada por la familia de intervalos de la forma $(-\infty, x]$, donde $x \in \overline{\mathbb{R}}$. A los elementos de esa σ-álgebra los llamaremos borelianos de $\overline{\mathbb{R}}$.*

RESULTADOS

1. La σ-álgebra de Borel en $\overline{\mathbb{R}}$ está formada los borelianos de $\mathbb{R}$ y los conjuntos de la forma $B \cup \{\infty\}$, $B \cup \{-\infty\}$ y $B \cup \{-\infty, \infty\}$, donde B es un boreliano de $\mathbb{R}$.

2. La σ-álgebra $\mathfrak{B}\left(\overline{\mathbb{R}}\right)$ de los conjuntos borelianos de $\overline{\mathbb{R}}$ está generada por cualquiera de las siguientes familias de conjuntos.

a) Los intervalos de la forma $[-\infty, x]$, donde $x \in \mathbb{R}$.

b) Los intervalos de la forma $[-\infty, x]$, donde $x \in \overline{\mathbb{R}}$.

c) Los intervalos de la forma $[-\infty, x)$, donde $x \in \mathbb{R}$.

d) Los intervalos de la forma $[-\infty, x)$, donde $x \in \overline{\mathbb{R}}$.

σ-álgebra de Borel en $\mathbb{R}^n$

Definición 9. *Por una celda en $\mathbb{R}^n$ se entenderá un conjunto de la forma $I_1 \times \cdots \times I_n$, donde $I_1, \cdots, I_n$ son intervalos en $\mathbb{R}$.*

Denotaremos por $\mathcal{R}$ a la familia de celdas en $\mathbb{R}^n$.

Definición 10 ($\boldsymbol{\sigma}$-álgebra de Borel en $\mathbb{R}^n$). *La σ-álgebra de Borel en $\mathbb{R}^n$, la cual será denotada por $\mathfrak{B}(\mathbb{R}^n)$, es la σ-álgebra de subconjuntos de $\mathbb{R}^n$ generada por $\mathcal{R}$. A los elementos de esa σ-álgebra los llamaremos borelianos de $\mathbb{R}^n$.*

RESULTADOS

1. La σ-álgebra de Borel en $\mathbb{R}^n$ está generada por cualquiera de las siguientes familias de conjuntos:

$\mathcal{D}_0$: Las celdas de $\mathbb{R}^n$ de la forma:

$(-\infty, x_1] \times (-\infty, x_2] \times \cdots \times (-\infty, x_n]$, donde $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

$\mathcal{D}_1$: Las celdas de $\mathbb{R}^n$ de la forma:

$(-\infty, x_1) \times (-\infty, x_2) \times \cdots \times (-\infty, x_n)$, donde $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

$\mathcal{D}_2$: Las celdas de $\mathbb{R}^n$ de la forma;

$(a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]$, donde $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$.

$\mathcal{D}_3$: Las celdas de $\mathbb{R}^n$ de la forma:

$(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$, donde $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$.

$\mathcal{D}_4$: Las celdas de $\mathbb{R}^n$ de la forma:

$[a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_n, b_n)$, donde $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$.

$\mathcal{D}_5$: Las celdas de $\mathbb{R}^n$ de la forma:

$[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$, donde $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$.

$\mathcal{D}_6$: Los subconjuntos de $\mathbb{R}^n$ de la forma $B_1 \times B_2 \times \ldots \times B_n$, donde $B_1, \ldots, B_n$ son borelianos de $\mathbb{R}$.

$\mathcal{D}_7$: Los subconjuntos abiertos de $\mathbb{R}^n$.

Teoremas de clases monótonas

Definición 11. *Sea $\mathbb{F}$ un conjunto y $\mathcal{G}$ una familia de subconjuntos de $\mathbb{F}$. Diremos que:*

i. $\mathcal{G}$ es cerrada bajo complementos si $A^c \in \mathcal{G}$ para cualquier $A \in \mathcal{G}$.

ii. $\mathcal{G}$ es cerrada bajo diferencias propias si $B - A \in \mathcal{G}$ para cualquier pareja $A, B \in \mathcal{G}$ tal que $A \subset B$.

iii. $\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) finitas si $\bigcup_{j=1}^{n} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{n} A_j \in \mathcal{G}$) para cualquier colección finita $A_1, A_2, \ldots, A_n$ de elementos de $\mathcal{G}$.

iv. $\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) monótonas si $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{\infty} A_j \in \mathcal{G}$) para cualquier sucesión creciente (resp. decreciente) $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{G}$.

Definición 12. *Sea $\mathbb{F}$ un conjunto y $\mathbb{M}$ una familia de subconjuntos de $\mathbb{F}$. Se dice que $\mathbb{M}$ es una clase monótona si es cerrada bajo uniones e intersecciones monótonas.*

Definición 13. *Dado un conjunto $\mathbb{F}$ y una familia arbitraria de clases monótonas de subconjuntos de $\mathbb{F}$, se define la intersección de esas clases monótonas como la familia de conjuntos que pertenecen a todas ellas.*

Definición 14. *Dada una colección $\mathcal{A}$ de subconjuntos de un conjunto $\mathbb{F}$, se define la clase monótona generada por $\mathcal{A}$ como la intersección de todas las clases monótonas que contienen a todos los elementos de $\mathcal{A}$ y se le denota por $\mathcal{M}(\mathcal{A})$.*

Teorema de clases monótonas para álgebras: Sea $\mathbb{F}$ un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$, entonces $\sigma(A) = \mathcal{M}(\mathcal{A})$.

Definición 15. *Sea $\mathbb{F}$ un conjunto y $\mathcal{P}$ una familia de subconjuntos de $\mathbb{F}$. Se dice que $\mathcal{P}$ es un π-sistema si es cerrada bajo intersecciones finitas.*

Definición 16. *Sea $\mathbb{F}$ un conjunto y $\mathcal{D}$ una familia de subconjuntos de $\mathbb{F}$. Se dice que $\mathcal{D}$ es un d-sistema si $\mathbb{F} \in \mathcal{D}$ y es cerrada bajo diferencias propias y uniones monótonas.*

Definición 17. *Dado un conjunto $\mathbb{F}$ y una familia arbitraria de d-sistemas de subconjuntos de $\mathbb{F}$, se define la intersección de esos d-sistemas como la familia de conjuntos que pertenecen a todas ellas.*

Definición 18. *Dada una colección $\mathcal{G}$ de subconjuntos de un conjunto $\mathbb{F}$, se define el d-sistema generado por $\mathcal{G}$ como la intersección de todos los d-sistemas que contienen a todos los elementos de $\mathcal{G}$ y se le denota por $d(\mathcal{G})$.*

Teorema de clases monótonas para pi sistemas: Sea $\mathbb{F}$ un conjunto y $\mathcal{P}$ un π-sistema de subconjuntos de $\mathbb{F}$, entonces $d(\mathcal{P}) = \sigma(\mathcal{P})$.

Funciones finitamente aditivas y σ-aditivas

Definición 19 (Función finitamente aditiva sobre un álgebra). *Sea $\mathbb{F}$ un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ es finitamente aditiva si dada cualquier familia finita, $A_1, \ldots, A_n$, de elementos de $\mathcal{A}$ tal que $A_i \cap A_j = \emptyset$ para $i \neq j$, entonces $\mu(\bigcup\limits_{k=1}^{n} A_k) = \sum_{k=1}^{n} \mu(A_k)$.*

Definición 20 (Función σ-aditiva sobre un álgebra). *Sea $\mathbb{F}$ un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ es σ-aditiva si es finitamente aditiva y dada cualquier familia infinita numerable, $A_1, A_2, \ldots$, de elementos de $\mathcal{A}$ tal que $A_i \cap A_j = \emptyset$ para $i \neq j$ y $\bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{A}$, entonces $\mu\left(\bigcup\limits_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$.*

Definición 21 (Función σ-aditiva sobre una σ-álgebra). *Sea $\mathbb{F}$ un conjunto y $\Im$ una σ-álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \Im \mapsto \overline{\mathbb{R}}$ es σ-aditiva si es finitamente aditiva y dada cualquier familia infinita numerable, $A_1, A_2, \ldots$, de elementos de $\Im$ tal que $A_i \cap A_j = \emptyset$, entonces $\mu\left(\bigcup\limits_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$.*

RESULTADOS

1. Sea $\mathbb{F}$ un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ una función no negativa finitamente aditiva, entonces:

a) Si $A, B \in \mathcal{A}$ y $A \subset B$, entonces $\mu(A) \leq \mu(B)$.

b) Si $A_1, \ldots, A_n \in \mathcal{A}$, entonces $\mu\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} \mu(A_k)$.

2. Sea $\mathbb{F}$ un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ una función no negativa finitamente aditiva, entonces:

Si $A, B \in \mathcal{A}$, $A \subset B$ y $\mu(A) < \infty$, entonces:

$$\mu(B - A) = \mu(B) - \mu(A)$$

3. Para cualquier pareja $A, B \in \mathcal{A}$ tal que $\mu(A) < \infty$ o $\mu(B) < \infty$, se tiene:

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

La medida de Lebesgue en $\mathbb{R}$

Definición 22. *Diremos que una colección finita o infinita numerable de intervalos abiertos finitos $I_1, I_2, \ldots$ es una cubierta del conjunto A si $A \subset \bigcup_n I_n$.*

Definición 23. *Se define la medida exterior, $m_e(A)$, de un conjunto A, mediante la relación:*

$$m_e(A) = \text{ínf}\left\{\sum_j l(I_j) : I_1, I_2, \ldots \text{ es cubierta de } A\right\}$$

Definición 24. *Se dice que un conjunto E es Lebesgue medible si:*

$$m_e(A) = m_e(A \cap E) + m_e(A \cap E^c)$$

para cualquier conjunto A. Además, en ese caso, se define la medida de E, $m(E)$, como la medida exterior de E.

RESULTADOS

1. La familia de conjuntos Lebesgue medibles forma una σ-álgebra de subconjuntos de $\mathbb{R}$.

2. La función que asigna a cada conjunto Lebesgue medible E su medida, $m(E)$, es una función σ-aditiva.

3. Todo conjunto de medida exterior cero es Lebesgue medible.

4. Todo conjunto boreliano es Lebesgue medible.

5. La σ-álgebra de los conjuntos Lebesgue medibles es la σ-álgebra generada por los intervalos y los conjuntos de medida exterior cero.

Denotaremos por $\mathfrak{L}(\mathbb{R})$ a la σ-álgebra formada por los conjuntos Lebesgue medibles en $\mathbb{R}$ y por λ a la medida m, a la cual llamaremos la medida de Lebesgue sobre $\mathfrak{L}(\mathbb{R})$.

Medidas sobre álgebras y $\sigma-$álgebras

Definición 25 (**Medida**). *Sea $\mathbb{F}$ un conjunto y $\Im$ una σ-álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \Im \mapsto \overline{\mathbb{R}}$ es una medida si μ es σ-aditiva y $\mu(\emptyset) = 0$.*

Definición 26 (**Espacio de medida**). *Llamaremos espacio de medida a una terna $(\mathbb{F}, \Im, \mu)$ donde $\mathbb{F}$ es un conjunto, $\Im$ una σ-álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \Im \mapsto \overline{\mathbb{R}}$ una medida.*

Definición 27. *Diremos que un espacio de medida $(\mathbb{F}, \Im, \mu)$ es completo si $\Im$ contiene a todos los subconjuntos de los conjuntos de medida μ igual a cero.*

Definición 28. *Sea $(\mathbb{F}, \Im, \mu)$ un espacio de medida. Diremos que μ es finita si $\mu(\mathbb{F}) < \infty$. Diremos que μ es σ-finita si existe una colección infinita numerable de conjuntos $E_k \in \Im$ tales que $\mathbb{F} = \bigcup_{k=1}^{\infty} E_k$ y $\mu(E_k) < \infty$ para cualquier k.*

Definición 29. *Sea $\mathbb{F}$ un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ es σ-subaditiva, o que satisface la propiedad de la subaditividad numerable, si dada cualquier colección infinita $A_1, A_2, \ldots$ de elementos de $\mathcal{A}$ tales que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, entonces:*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

RESULTADOS

1. Sea $\mathbb{F}$ un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ una función no negativa y finitamente aditiva, entonces las siguientes propiedades son equivalentes:

a) μ es σ-subaditiva.

b)μ es σ-aditiva.

c) Para cualquier colección infinita $A_1, A_2, \ldots$ de elementos de $\mathcal{A}$ tales que $A_1 \subset A_2 \subset \ldots$ y $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, se tiene:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

2. Sea $\mathbb{F}$ un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \mathbb{R}$ una función no negativa y finitamente aditiva, entonces las siguientes propiedades son equivalentes:

a) μ es σ-subaditiva.

b) μ es σ-aditiva.

c) Para cualquier sucesión creciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$ tales que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, se tiene:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \text{lím}_{n\to\infty}\, \mu(A_n)$$

d) Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$ tales que $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$, se tiene:

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \text{lím}_{n\rightsquigarrow\infty}\, \mu(A_n)$$

e) Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$ tales que $\bigcap_{n=1}^{\infty} A_n = \emptyset$, se tiene:

$$\text{lím}_{n\to\infty}\, \mu(A_n) = 0$$

3. Sea $\mathbb{F}$ un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ una quasi medida. Entonces, para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$, tales que $\mu(A_N) < \infty$ para alguna $N \in \mathbb{N}$ y $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$, se tiene:

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \text{lím}_{n\rightsquigarrow\infty}\, \mu(A_n)$$

4. Sea $(\mathbb{F}, \Im, \mu)$ un espacio de medida. Entonces:

a) μ es σ-subaditiva.

b) Para cualquier sucesión creciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\Im$, se tiene:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \text{lím}_{n\rightsquigarrow\infty}\, \mu(A_n)$$

c) Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\Im$ tales que $\mu(A_N) < \infty$ para alguna $N \in \mathbb{N}$, se tiene:

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \text{lím}_{n\rightsquigarrow\infty}\, \mu(A_n)$$

5. Si un espacio de medida $(\mathbb{F}, \Im_0, \mu)$ no es completo, se puede completar.

6. Dados un espacio de medida $(\mathbb{F}, \Im, \mu)$, un conjunto cualquiera $\mathbb{E}$ y una función $f : \mathbb{F} \mapsto \mathbb{E}$. Definamos $\mathbb{H} = \{B \subset \mathbb{E} : f^{-1}(B) \in \Im\}$, entonces:

a) $\mathbb{H}$ es una $\sigma-$algebra de subconjuntos de $\mathbb{E}$.

b) La función $\mu_f : \mathbb{H} \to \overline{\mathbb{R}}$ definida por $\mu_f(B) = \mu\left[f^{-1}(B)\right]$ es una medida.

Integración con respecto a una medida

Funciones medibles

Definición 30 (Espacio medible). *Llamaremos espacio medible a una pareja* $(\mathbb{E}, \mathcal{E})$ *donde* $\mathbb{E}$ *es un conjunto y* $\mathcal{E}$ *una* σ*-álgebra de subconjuntos de* $\mathbb{E}$*.*

En lo que sigue de esta sección, $(\mathbb{E}, \mathcal{E})$ será un espacio medible.

Definición 31. *Diremos que una función* $f : \mathbb{E} \mapsto \overline{\mathbb{R}}$ *es medible si* $\{y \in \mathbb{E} : f(y) \leq x\} \in \mathcal{E}$ *para cualquier* $x \in \mathbb{R}$*.*

Definición 32 (Función medible). *De manera general, si* $(\mathbb{F}, \mathcal{F})$ *es un espacio medible, diremos que una función* $f : \mathbb{E} \mapsto \mathbb{F}$ *es medible si* $f^{-1}(B) \in \mathcal{E}$ *para cualquier conjunto* $B \in \mathcal{F}$*.*

RESULTADOS

Las funciones medibles $f : (\mathbb{E}, \mathcal{E}) \to \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ tienen las siguientes propiedades:

1. Si $(f_n)_{n\in\mathbb{N}}$ es una sucesión de funciones medibles, entonces:

a) Para cualquier $n \in \mathbb{N}$, las funciones $\min\{f_1, \ldots, f_n\}$ y $\max\{f_1, \ldots, f_n\}$ son medibles.

b) Las funciones $\inf\{g_n : n \in \mathbb{N}\}$ y $\sup\{g_n : n \in \mathbb{N}\}$ son medibles.

c) Las funciones $\liminf_{n\to\infty} f_n$ y $\limsup_{n\to\infty} f_n$ son medibles.

d) Si $\lim_{n\to\infty} f_n(x)$ existe para cualquier $x \in \mathbb{E}$ (donde ∞ y $-\infty$ también se incluyen como posibles límites), entonces la función $f : \mathbb{E} \to \overline{\mathbb{R}}$ definida por $f(x) = \lim_{n\to\infty} f_n(x)$ es medible.

2. Si f es una función medible, entonces $f^+ = \max\{f, 0\}$ y $f^- = \max\{-f, 0\}$ son medibles.

3. Si f y g son funciones medibles y $c \in \mathbb{R}$. Entonces las funciones $f + c$, cf y fg son medibles.

4. Si f y g son funciones medibles y $h : \mathbb{E} \to \overline{\mathbb{R}}$ es una función tal que $h(x) = f(x) + g(x)$ en todos los puntos $x \in \mathbb{E}$ para los cuales $f(x) + g(x)$ esté definida y h es constante en el conjunto de puntos $x \in \mathbb{E}$ para los cuales $f(x) + g(x)$ no esté definida, entonces h es medible.

5. Si $(\mathbb{E}, \mathcal{E}, \mu)$ es un espacio de medida completo, f una función medible y $g : \mathbb{E} \to \overline{\mathbb{R}}$ una función tal que $g = f$ excepto a lo más en un conjunto de medida cero, entonces g es medible.

6. Si $h : \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right) \mapsto \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ es una función boreliana y f es una función medible, entonces $h \circ f$ es medible.

7. Si f es una función medible que toma únicamente valores en $\mathbb{R}$ y $h:\mathbb{R}\to\mathbb{R}$ es una función continua, entonces $h\circ f$ es medible.

8. Si f es una función medible, entonces las funciones $g=\frac{1}{f}I_{\{x\in\mathbb{E}:f(x)\neq 0\}}$ y $h=|f|^{\alpha}$, donde $\alpha\geq 0$, son medibles.

Funciones medibles simples

Definición 33 (Función simple). *Diremos que una función medible $\varphi:\mathbb{E}\to\overline{\mathbb{R}}$ es simple si tiene la forma $\varphi=\sum_{k=1}^{m}b_kI_{E_k}$, donde $b_1,\dots,b_m$ son números reales y $E_1,\dots,E_m$ son elementos de $\mathcal{E}$.*

Teorema 1. *Sea $f:\mathbb{E}\to\overline{\mathbb{R}}$ una función medible no negativa, entonces existe una sucesión no decreciente de funciones medibles simples no negativas $\varphi_n:\mathbb{E}\to\overline{\mathbb{R}}$ tales que* $\lim_{n\to\infty}\varphi_n(y)=f(y)$ *para cualquier $y\in\mathbb{E}$.*

Demostración

Para cada $n\in\mathbb{N}$, definamos:

$$\varphi_n(y)=\begin{cases}\sum_{m=1}^{n2^n}\frac{m-1}{2^n}I_{\left\{z\in\mathbb{E}:\frac{m-1}{2^n}\leq f(z)<\frac{m}{2^n}\right\}}(y) & \text{si } f(y)<n\\ n & \text{si } f(y)\geq n\end{cases}$$

Obsérvese que aunque, para cada $y\in\mathbb{E}$ tal que $f(y)<n$, $\varphi_n(y)$ es una suma, únicamente uno de los términos es distinto de cero.

También podríamos escribir la definición de φ_n de la siguiente manera:

$$\varphi_n(y)=\begin{cases}\frac{m-1}{2^n} & \text{si } \frac{m-1}{2^n}\leq f(y)<\frac{m}{2^n} \text{ y } m\in\{1,2,\dots,n2^n\}\\ n & \text{si } f(y)\geq n\end{cases}$$

Si $y\in\mathbb{E}$ es tal que $f(y)=\infty$, entonces $\varphi_n(y)=n$ para cualquier $n\in\mathbb{N}$, así que:

$$\lim_{n\to\infty}\varphi_n(y)=\infty=f(y)$$

Si $n\in\mathbb{N}$ y $y\in\mathbb{E}$ es tal que $f(y)<n$, sea m el único número natural tal que $\frac{m-1}{2^n}\leq f(y)<\frac{m}{2^n}$. Entonces, como $\varphi_n(y)=\frac{m-1}{2^n}$, se tiene:

$f(y)-\frac{1}{2^n}<\varphi_n(y)\leq f(y)$, así que $\lim_{n\to\infty}\varphi_n(y)=f(y)$.

Ahora bien, como $\frac{2(m-1)}{2^{n+1}} \leq f(y) < \frac{2m}{2^{n+1}}$, se tiene que, o bien $\frac{2m-2}{2^{n+1}} \leq f(y) < \frac{2m-1}{2^{n+1}}$ o bien $\frac{2m-1}{2^{n+1}} \leq f(y) < \frac{2m}{2^{n+1}}$.

En el primer caso, se tiene:

$$\varphi_{n+1}(y) = \frac{2m-2}{2^{n+1}} = \frac{m-1}{2^n} = \varphi_n(y)$$

En el segundo, se tiene:

$$\varphi_{n+1}(y) = \frac{2m-1}{2^{n+1}} > \frac{2m-2}{2^{n+1}} = \varphi_n(y)$$

Así que, en cualquier caso, $\varphi_n(y) \leq \varphi_{n+1}(y)$.

Así que, φ_n es una sucesión no decreciente de funciones simples no negativas tales que $\lim_{n\to\infty} \varphi_n(y) = f(y)$ para cualquier $y \in \mathbb{E}$.

■

Si $\varphi = \sum_{k=1}^{m} b_k I_{G_k}$ es una función simple entonces el conjunto de los valores que toma es finito. Sea $\{a_1, \ldots, a_n\}$ el conjunto formado por todos los distintos posibles valores no nulos de φ y, para $k \in \{1, \ldots, n\}$, sea $E_k = \{y \in \mathbb{E} : \varphi(y) = a_k\}$, entonces los conjuntos $E_1, \ldots, E_n$ son ajenos por parejas y $\varphi = \sum_{k=1}^{n} a_k I_{E_k}$. Esta última sumatoria será llamada la **representación canónica** de φ.

Integración de funciones medibles no negativas

Definición 34. *Si $\varphi : (\mathbb{E}, \mathcal{E}, \mu) \to \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ es una función medible simple no negativa con representación canónica $\varphi = \sum_{k=1}^{n} a_k I_{E_k}$, se define la integral de φ, $\int_{\mathbb{E}} \varphi d\mu$, de la siguiente manera:*

$$\int_{\mathbb{E}} \varphi d\mu = \sum_{k=1}^{n} a_k m\left(E_k\right)$$

Definición 35. *Si $f : (\mathbb{E}, \mathcal{E}, \mu) \to \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ es una función medible no negativa, se define la integral de f, con respecto a la medida μ, $\int_{\mathbb{E}} f d\mu$, de la siguiente manera:*

$$\int_{\mathbb{E}} f d\mu = \sup\left\{ \int_{\mathbb{E}} \varphi d\mu : \varphi \text{ es simple y } 0 \leq \varphi \leq f \right\}$$

Obsérvese que el conjunto $\left\{ \int_{\mathbb{E}} \varphi d\mu : \varphi \text{ es simple y } 0 \leq \varphi \leq f \right\}$ no necesariamente está acotado, de manera que estamos considerando que, en ese caso, el supremo del conjunto lo definimos como ∞.

Definición 36. *Si f es una función medible no negativa y $E \in \mathcal{E}$, se define:*

$$\int_E f d\mu = \int_{\mathbb{E}} (I_E \cdot f)\, d\mu$$

RESULTADOS

1. Si f y g son dos funciones medibles no negativas, entonces:

a) Si $f \leq g$, entonces $\int_{\mathbb{E}} f d\mu \leq \int_{\mathbb{E}} g d\mu$.

b) Si $f \leq g$ sobre un conjunto $E \in \mathcal{E}$, entonces $\int_E f d\mu \leq \int_E g d\mu$.

2. Sea f una función medible no negativa y $E \in \mathcal{E}$ tal que $\mu(E) = 0$; entonces:

$$\int_E f d\mu = 0$$

3. **Teorema de la convergencia monótona:** Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión no decreciente de funciones medibles no negativas, entonces:

$$\int_{\mathbb{E}} \lim_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_{\mathbb{E}} f_n d\mu$$

4. Si f y g son dos funciones medibles no negativas, entonces:

a) $\int_{\mathbb{E}} [af + bg]\, d\mu = a \int_{\mathbb{E}} f d\mu + b \int_{\mathbb{E}} g d\mu$ para cualesquiera números reales a y b no negativos.

b) $\int_{\mathbb{E}} f d\mu = 0$ si y sólo si $\mu\{y \in \mathbb{E} : f(y) > 0\} = 0$.

5. Sea f una función medible no negativa. Entonces, la función $m : \mathcal{E} \to \overline{\mathbb{R}}$, definida por $m(E) = \int_E f d\mu$, es una medida.

6. **Lema de Fatou:** Sea $\{f_n\}_{n\in N}$ una sucesión de funciones medibles no negativas. Entonces:

$$\int_{\mathbb{E}} \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_{\mathbb{E}} f_n d\mu$$

Funciones medibles integrables

Definición 37. *Se dice que una función medible f es integrable sobre un conjunto $E \in \mathcal{E}$ si $\int_E |f|\, d\mu < \infty$.*

Definición 38. *Si f es una función medible e integrable sobre un conjunto $E \in \mathcal{E}$, se define su integral sobre E, $\int_E f d\mu$, de la siguiente manera:*

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu$$

RESULTADOS

1. Si f es una función medible no negativa tal que $\int_{\mathbb{E}} f d\mu < \infty$, entonces $\mu\{y \in \mathbb{E} : f(y) = \infty\} = 0$.

2. Una función medible f es integrable sobre un conjunto $E \in \mathcal{E}$ si y sólo si f^+ y f^- son integrables sobre E.

3. Si f y g dos funciones medibles e integrables sobre un conjunto $E \in \mathcal{E}$, entonces:

a) Para cualquier número real c, la función cf es integrable sobre E y $\int_E cf d\mu = c \int_E f d\mu$.

b) Si $f \leq g$ sobre E, entonces $\int_E f d\mu \leq \int_E g d\mu$.

c) $\left|\int_E f d\mu\right| \leq \int_E |f| \, d\mu$

4. Sean $f_1, \ldots, f_n$ n funciones medibles e integrables sobre un conjunto medible E, $a_1, \ldots, a_n$ números reales y $h : \mathbb{F} \to \overline{\mathbb{R}}$ una función medible tal que $h(x) = \sum_{k=1}^{n} a_k f_k(x)$ en todos los puntos $x \in E$ para los cuales $\sum_{k=1}^{n} a_k f_k(x)$ esté definida, entonces h es integrable sobre E y:

$$\int_E h d\mu = \sum_{k=1}^{n} a_k \int_E f_k d\mu$$

5. Sea g una función medibles e integrable tal que $\int_E g d\mu \geq 0$ para cualquier $E \in \mathcal{E}$, entonces:

$$\mu\{x \in \mathbb{E} : g(x) < 0\} = 0$$

6. Sean f y g dos funciones medibles e integrables.

a) Si $\int_E f d\mu \leq \int_E g d\mu$ para cualquier $E \in \mathcal{E}$, entonces:

$$\mu\{x \in \mathbb{F} : f(x) > g(x)\} = 0$$

b) Si $\int_E f d\mu = \int_E g d\mu$ para cualquier $E \in \mathcal{E}$, entonces:

$$\mu\{x \in \mathbb{F} : f(x) \neq g(x)\} = 0$$

7. Sea g una función no negativa, integrable sobre un conjunto $E \in \mathcal{E}$, $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tales que $|f_n| \leq g$ y $\lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero y f una función medible tal que $f = \lim_{n\to\infty} f_n$, donde este límite existe, entonces:

$$\lim_{n\to\infty} \int_E |f_n - f| \, d\mu = 0$$

8. **Teorema de la convergencia dominada:** Sea g una función no negativa, integrable sobre un conjunto medible E, $\{f_n\}_{n\in N}$ una sucesión de funciones medibles tales que $|f_n| \leq g$

y $\lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero y f una función medible tal que $f = \lim_{n\to\infty} f_n$, donde este límite existe entonces, f es integrable sobre E y:

$$\int_E f d\mu = \lim_{n\to\infty} \int_E f_n d\mu$$

Integrabilidad uniforme

En esta sección asumimos que la medida μ es finita.

RESULTADOS

1. Si f es una función medible no negativa, entonces:

$$\int_{\mathbb{F}} f d\mu = \int_0^{\infty} \mu\left(\{y \in \mathbb{F} : f(y) > x\}\right) dx$$

2. Si f es una función integrable, entonces la serie:

$$\sum_{k=1}^{\infty} \mu\left(\{y \in \mathbb{F} : |f(y)| \geq k\}\right)$$

converge.

3. Si f es una función integrable, entonces

$$\lim_{\alpha\to\infty} \mu\left[|f| > \alpha\right] = 0$$

4. Una función medible f es integrable si y sólo si:

$$\lim_{\alpha\to\infty} \int_{[|f|>\alpha]} |f| \, d\mu = 0$$

5. Una función medible f es integrable si y sólo si dada $\varepsilon > 0$ existe $\delta > 0$ tal que $\int_A |f| \, d\mu < \varepsilon$ para cualquier conjunto $A \in \mathcal{F}$ tal que $\mu(A) < \delta$.

Definición 39 (**Integrabilidad uniforme**)**.** *Se dice que una familia $\mathcal{H}$ de funciones medibles es uniformemente integrable si:*

$$\lim_{\alpha\to\infty} \sup\left\{\int_{[|f|>\alpha]} |f| \, d\mu : f \in \mathcal{H}\right\} = 0$$

RESULTADOS

1. Una familia $\mathcal{H}$ de funciones medibles es uniformemente integrable si y sólo si el conjunto $\left\{\int_{\mathbb{F}} |f| \, d\mu : f \in \mathcal{H}\right\}$ está acotado y, dada cualquier $\varepsilon > 0$, existe $\delta > 0$ tal que $\int_A |f| \, d\mu < \varepsilon$ para cualesquiera $f \in \mathcal{H}$ y $A \in \mathcal{F}$ tal que $\mu(A) \leq \delta$.

2. Sea f una función medible e integrable, Γ un conjunto cualquiera y $\{f_\gamma : \gamma \in \Gamma\}$ una familia de funciones medibles tales que $|f_\gamma| \leq f$ para cualquier $\gamma \in \Gamma$, entonces la familia $\{f_\gamma : \gamma \in \Gamma\}$ es uniformemente integrable.

3. Sea $\{f_n\}_{n \in N}$ una familia uniformemente integrable de funciones medibles tal que $f = \lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero, entonces f es integrable y:

$$\lim_{n\to\infty} \int_{\mathbb{F}} |f_n - f|\, d\mu = 0$$

4. **Teorema de la convergencia uniformemente integrable:** Sea $\{f_n\}_{n \in N}$ una familia uniformemente integrable de funciones medibles tales que $\lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero, entonces $f = \lim_{n\to\infty} f_n$ es integrable y:

$$\int_{\mathbb{F}} \lim_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_{\mathbb{F}} f_n d\mu$$

5. Sean $(f_n)_{n \in N}$ una sucesión de funciones integrables y f una función medible tales que $\lim_{n\to\infty} \int_{\mathbb{F}} |f_n - f|\, d\mu = 0$. Entonces la familia $\{f_n : n \in \mathbb{N}\}$ es uniformemente integrable.

6. Sea $(f_n)_{n \in N}$ una sucesión de funciones integrables no negativas tal que $f = \lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero y $\lim_{n\to\infty} \int_{\mathbb{F}} f_n d\mu = \int_{\mathbb{F}} f d\mu < \infty$. Entonces $\lim_{n\to\infty} \int_{\mathbb{F}} |f_n - f|\, d\mu = 0$.

7. Sea $\{f_n\}_{n \in N}$ una sucesión de funciones integrables no negativas tal que $f = \lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero y $\lim_{n\to\infty} \int_{\mathbb{F}} f_n d\mu = \int_{\mathbb{F}} f d\mu < \infty$. Entonces la familia $\{f_n : n \in \mathbb{N}\}$ es uniformemente integrable.

Combinando lo anterior, se tiene el siguiente resultado:

8. Sea $(f_n)_{n \in N}$ una sucesión de funciones integrables tales que $f = \lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero, entonces las siguientes tres condiciones son equivalentes:

a) La familia $\{f_n : n \in \mathbb{N}\}$ es uniformemente integrable.

b) $\lim_{n\to\infty} \int_{\mathbb{F}} |f_n - f|\, d\mu = 0$

c) $\lim_{n\to\infty} \int_{\mathbb{F}} |f_n|\, d\mu = \int_{\mathbb{F}} |f|\, d\mu < \infty$

9. Sea $\{f_n\}_{n \in N}$ una sucesión de funciones integrables no negativas tales que $f = \lim_{n\to\infty} f_n$ existe excepto a lo más en un conjunto de medida cero, entonces las siguientes tres condiciones son equivalentes:

a) La familia $\{f_n : n \in \mathbb{N}\}$ es uniformemente integrable.

b) $\lim_{n\to\infty} \int_{\mathbb{F}} |f_n - f|\, d\mu = 0$

c) $\lim_{n\to\infty}\int_{\mathbb{F}} f_n d\mu = \int_{\mathbb{F}} f d\mu < \infty$

10. Sea $\mathcal{H}$ una familia de funciones medibles uniformemente integrable. Entonces la familia:

$\mathcal{G} = \Big\{ f : \mathbb{F} \to \overline{\mathbb{R}}^+$: Existe una sucesión $(f_n)_{n\in\mathbb{N}}$ de funciones en $\mathcal{H}$ tales que

$\lim_{n\to\infty} f_n = f$ excepto a lo más en un conjunto de medida cero$\}$

es uniformemente integrable.

Funciones de variación acotada

Definición 40. *Dada una función $g : [a,b] \to \mathbb{R}$ y una partición $P = \{x_0, x_1, \ldots, x_n\}$ del intervalo $[a,b]$, definamos $V_g(P) = \sum_{k=1}^{n} |g(x_k) - g(x_{k-1})|$.*

Diremos que g es de variación acotada en $[a,b]$ si:

$V_g[a,b] = \sup\{V_g(P) : P$ *es una partición de* $[a,b]\} < \infty$

RESULTADOS

1. El conjunto de funciones de variación acotada, definidas en un mismo intervalo $[a,b]$, forma un espacio vectorial.

2. Si $g : [a,b] \to \mathbb{R}$ es una función monótona, entonces es de variación acotada.

3. Si $g_1 : [a,b] \to \mathbb{R}$ y $g_2 : [a,b] \to \mathbb{R}$ son funciones no decrecientes, entonces $g_1 - g_2$ es de variación acotada.

4. Si $g_1 : \mathbb{R} \to \mathbb{R}$ y $g_2 : \mathbb{R} \to \mathbb{R}$ son funciones no decrecientes, entonces $g_1 - g_2$ es de variación acotada sobre cualquier intervalo compacto.

5. Si $g : \mathbb{R} \to \mathbb{R}$ es una función de variación acotada sobre cualquier intervalo compacto, entonces se puede expresar como la diferencia de dos funciones no decrecientes.

6. Una función $g : \mathbb{R} \to \mathbb{R}$ es de variación acotada sobre cualquier intervalo compacto si y sólo si se puede expresar como la diferencia de dos funciones no decrecientes.

7. Una función de variación acotada sobre cualquier intervalo compacto tiene a lo más un conjunto numerable de discontinuidades.

8. Si $g : \mathbb{R} \to \mathbb{R}$ es una función de variación acotada sobre cualquier intervalo compacto, entonces se puede expresar como la diferencia de dos funciones no decrecientes, $f_1 : \mathbb{R} \to \mathbb{R}$ y $f_2 : \mathbb{R} \to \mathbb{R}$, las cuales no tienen discontinuidades en común del mismo lado.

9. Si $g : \mathbb{R} \to \mathbb{R}$ es una función continua por la derecha y de variación acotada sobre cualquier intervalo compacto, entonces se puede expresar como la diferencia de dos funciones no decrecientes continuas por la derecha.

10. Si $g : \mathbb{R} \to \mathbb{R}$ es una función continua por la izquierda y de variación acotada sobre cualquier intervalo compacto, entonces se puede expresar como la diferencia de dos funciones no decrecientes continuas por la izquierda.

11. Sea $f : \mathbb{R} \to \overline{\mathbb{R}}$ una función medible. Si f es integrable en el intervalo $[a, b]$, entonces la función $F : [a, b] \to \mathbb{R}$ definida por:

$$F(x) = \int_a^x f(y)\, dy$$

es de variación acotada.

La integral de Riemann-Stieltjes

Definición 41. *Sean $f : [a, b] \to \mathbb{R}$ y $g : [a, b] \to \mathbb{R}$ dos funciones acotadas y $P = \{x_0, x_1, \ldots, x_n\}$ una particion del intervalo $[a, b]$. Una suma de Riemann-Stieltjes $S(P, f, g)$ de f con respecto a g, correspondiente a la particion P, es una suma de la forma $S(P, f, g) = \sum_{k=1}^{n} f(\xi_k)\, [g(x_k) - g(x_{k-1})]$, donde $\xi_k \in [x_{k-1}, x_k]$ para $k \in \{1, 2, \ldots, n\}$.*

Definición 42. *Se dice que f es integrable con respecto a g en el intervalo $[a, b]$ si existe un numero real I tal que para cualquier $\varepsilon > 0$ existe una particion P_ε del intervalo $[a, b]$ tal que $|S(P, f, g) - I| < \varepsilon$ para cualquier particion P que sea un refinamiento de P_ε y cualquier suma de Riemann-Stieltjes $S(P, f, g)$ de f con respecto a g, correspondiente a la particion P. Al numero real I de esta definición se le llama la integral de f con respecto a g y se le denota por $\int_a^b f dg$.*

Criterio de Cauchy

Definición 43. *Se dice que la pareja de funciones acotadas $f : [a, b] \to \mathbb{R}$ y $g : [a, b] \to \mathbb{R}$ satisface el criterio de Cauchy si para cada $\varepsilon > 0$ existe una partición P_ε del intervalo $[a, b]$ tal que si P y P' son dos refinamientos de P_ε y $S(P, f, g)$, $S(P', f, g)$ son sumas de Riemann-Stieltjes de f con respecto a g, entonces:*

$$|S(P, f, g) - S(P', f, g)| < \varepsilon$$

RESULTADOS

1. Una función f es integrable con respecto a g en el intervalo $[a, b]$ si y sólo si la pareja f, g satisface el criterio de Cauchy.

2. Si g es de variación acotada, entonces toda función continua es integrable con respecto a g.

3. Si $f_1 : [a,b] \to \mathbb{R}$ y $f_2 : [a,b] \to \mathbb{R}$ son integrables con respecto a $g : [a,b] \to \mathbb{R}$ y $\alpha_1, \alpha_2 \in \mathbb{R}$, entonces $\alpha_1 f_1 + \alpha_2 f_2$ es integrable con respecto a g y:

$$\int_a^b (\alpha_1 f_1 + \alpha_2 f_2)\, dg = \alpha_1 \int_a^b f_1 dg + \alpha_2 \int_a^b f_2 dg$$

4. Si $f : [a,b] \to \mathbb{R}$ es integrable con respecto a $g_1 : [a,b] \to \mathbb{R}$ y con respecto a $g_2 : [a,b] \to \mathbb{R}$ y $\alpha_1, \alpha_2 \in \mathbb{R}$, entonces f es integrable con respecto a $\alpha_1 g_1 + \alpha_2 g_2$ y:

$$\int_a^b f d\,(\alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 \int_a^b f dg_1 + \alpha_2 \int_a^b f dg_2$$

5. Si $f : [a,b] \to \mathbb{R}$ es integrable con respecto a una función no decreciente $g : [a,b] \to \mathbb{R}$, entonces $|f|$ es integrable con respecto a g y:

$$\left| \int_a^b f dg \right| \leq \int_a^b |f|\, dg$$

6. Sean $f : [a,b] \to \mathbb{R}$ una función continua y $g : [a,b] \to \mathbb{R}$ una función de variación acotada, entonces la función $F : [a,b] \to \mathbb{R}$ definida por:

$$F\,(t) = \int_a^t f dg$$

es de variación acotada.

7. Sea $g : [a,b] \to \mathbb{R}$ una función acotada que no es de variación acotada en $[a,b]$. Entonces existe una función continua $f : [a,b] \to \mathbb{R}$ la cual no es Riemann-Stieltjes integrable con respecto a g.

8. Si toda función continua $f : [a,b] \to \mathbb{R}$ es integrable con respecto a la función acotada $g : [a,b] \to \mathbb{R}$, entonces g es de variación acotada.

9. Sean $f : [a,b] \to \mathbb{R}$ y $g : [a,b] \to \mathbb{R}$ dos funciones acotadas y supongamos que f es integrable con respecto a g, entonces g es integrable con respecto a f y, además, se tiene:

$$\int_a^b g df = g(b)f(b) - g(a)f(a) - \int_a^b f dg$$

10. Sea $g : [a,b] \to \mathbb{R}$ una función continua de variación acotada y $F : \mathbb{R} \to \mathbb{R}$ una función de clase C^1 , entonces $F \circ g$ es de variación acotada y:

$$F(g\,(t)) = F(g\,(a)) + \int_a^t F'(g\,(s))dg\,(s)$$

para cualquier $t \in [a,b]$.

Convergencia casi en todas partes

En esta sección, $(\mathbb{E}, \mathcal{E}, \mu)$ será un espacio de medida completo.

Definición 44 (**Convergencia casi en todas partes**). *Diremos que una sucesión de funciones medibles* $\{f_n\}_{n\in\mathbb{N}}$ *converge casi en todas partes a una función medible* f *si* $\lim_{n\to\infty} f_n = f$ *excepto a lo más en un conjunto de medida cero. Si éste es el caso, se escribirá* $f_n \xrightarrow{c.t.p.} f$.

Las siguientes propiedades se siguen inmediatamente de las correspondientes propiedades para las sucesiones de números reales:

i. Si una sucesión $(f_n)_{n\in\mathbb{N}}$ converge casi en todas partes a f, entonces cualquier subsucesión de $(f_n)_{n\in\mathbb{N}}$ también converge casi en todas partes a f.

ii. Si $(f_n)_{n\in\mathbb{N}}$ es una sucesión de funciones medibles tal que $f_n \xrightarrow{c.t.p.} f$ y $f_n \xrightarrow{c.t.p.} g$, entonces $f = g$ casi en todas partes.

iii. Si c es una constante y $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{c.t.p.} f$, entonces $cf_n \xrightarrow{c.t.p.} cf$.

iv. Si $(f_n)_{n\in\mathbb{N}}$ y $(g_n)_{n\in\mathbb{N}}$ son dos sucesiones de funciones medibles tales que $f_n \xrightarrow{c.t.p.} f$ y $g_n \xrightarrow{c.t.p.} g$, entonces $f_n + g_n \xrightarrow{c.t.p.} f + g$ y $f_n g_n \xrightarrow{c.t.p.} fg$.

Los resultados que siguen, con relación a la convergencia casi en todas partes, se entienden mejor si se tienen en mente los siguientes conceptos:

Dada una sucesión $A_1, A_2, \ldots$ de subconjuntos de un conjunto $\mathbb{F}$, se define el límite inferior (lím ínf) y el límite superior (lím sup) de esa sucesión de la siguiente manera:

$$\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

Obsérvese que $\liminf A_n$ está formado por todos los elementos $x \in \mathbb{F}$ para los cuales existe $N \in \mathbb{N}$ tal que $x \in A_n$ para cualquier $n \geq N$, mientras que $\limsup A_n$ está formado por todos los elementos $x \in \mathbb{F}$ que pertenecen a una infinidad de conjuntos de la sucesión. Así que se tiene siempre $\liminf A_n \subset \limsup A_n$.

RESULTADOS

1. Supongamos que μ es finita y sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles. Entonces $f_n \xrightarrow{c.t.p.} 0$ si y sólo si para cualquier $\varepsilon > 0$, se tiene:

$$\mu\left(\lim\sup\{y \in \mathbb{E} : |f_n(y)| > \varepsilon\}\right) = 0$$

2. Supongamos que μ es finita y sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles. Entonces $f_n \xrightarrow{c.t.p.} f$ si y sólo si para cualquier $\varepsilon > 0$, se tiene:

$$\mu\left(\lim\sup\{y \in \mathbb{F} : |f_n(y) - f(y)| > \varepsilon\}\right) = 0$$

3. **Lema de Borel-Cantelli:** Sea $E_1, E_2, \ldots$ una sucesión de conjuntos medibles tales que $\sum_{n=1}^{\infty} \mu(E_n) < \infty$, entonces:

$$\mu\left(\lim\sup E_n\right) = 0$$

4. Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tales que:

$$\textstyle\sum_{n=1}^{\infty} \mu\left[|f_n| > \varepsilon\right] < \infty \text{ para cualquier } \varepsilon > 0$$

Entonces $f_n \xrightarrow{c.t.p.} 0$.

5. Sea f una función medible y $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tales que:

$$\textstyle\sum_{n=1}^{\infty} \mu\left[|f_n - f| > \varepsilon\right] < \infty \text{ para cualquier } \varepsilon > 0$$

Entonces $f_n \xrightarrow{c.t.p.} f$.

Convergencia en medida

Definición 45 (Convergencia en medida). *Diremos que una sucesión de funciones medibles $(f_n)_{n\in\mathbb{N}}$ converge en medida si existe una función medible f tal que:*

$$\lim_{n\to\to\infty} \mu\left(\{y \in \mathbb{E} : |f_n(y) - f(y)| > \varepsilon\}\right) = 0$$

para cualquier $\varepsilon > 0$. Si éste es el caso, se escribirá $f_n \xrightarrow{\mu} f$.

Definición 46. *Diremos que una sucesión de funciones medibles $(f_n)_{n\in\mathbb{N}}$ es de Cauchy en medida si, para cualquier $\varepsilon > 0$ y cualquier $\delta > 0$ existe $N \in \mathbb{N}$ tal que:*

$$\mu\left(\{y \in \mathbb{F} : |f_n(y) - f_m(y)| > \varepsilon\}\right) < \delta$$

para cualquier par de números naturales n y m mayores o iguales a N.

RESULTADOS

1. Sea $\{f_n\}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{\mu} f$ y $f_n \xrightarrow{\mu} g$, entonces $f = g$ casi en todas partes.

2. Sea c una constante y $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{\mu} f$, entonces $cf_n \xrightarrow{\mu} cf$.

3. Sean $(f_n)_{n\in\mathbb{N}}$ y $(g_n)_{n\in\mathbb{N}}$ dos sucesiones de funciones medibles tales que $f_n \xrightarrow{\mu} f$ y $g_n \xrightarrow{\mu} g$, entonces $f_n + g_n \xrightarrow{\mu} f + g$.

4. Sean f una función medible, $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{\mu} f$ y $g : \mathbb{R} \to \mathbb{R}$ una función uniformemente continua, entonces $g \circ f_n \xrightarrow{\mu} g \circ f$.

5. Sean f una función medible, $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tales que $f_n \xrightarrow{\mu} f$ y $g : \mathbb{R} \to \mathbb{R}$ una función continua nula fuera de un intervalo $[a, b]$, entonces $g \circ f_n \xrightarrow{\mu} g \circ f$.

6. Supongamos que μ es finita y sean f una función medible, $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{\mu} f$ y $g : \mathbb{R} \to \mathbb{R}$ una función continua, entonces $g \circ f_n \xrightarrow{\mu} g \circ f$.

7. Supongamos que μ es finita y sean f y g dos funciones medibles, y $(f_n)_{n\in\mathbb{N}}$ y $(g_n)_{n\in\mathbb{N}}$ dos sucesiones de funciones medibles tales que $f_n \xrightarrow{\mu} f$ y $g_n \xrightarrow{\mu} g$, entonces $f_n g_n \xrightarrow{\mu} fg$.

8. Supongamos que μ es finita y sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{c.t.p.} 0$, entonces $f_n \xrightarrow{\mu} 0$.

9. Supongamos que μ es finita y sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{c.t.p.} f$, entonces $f_n \xrightarrow{\mu} f$.

10. Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles que converge en medida, entonces $(f_n)_{n\in\mathbb{N}}$ es de Cauchy en medida.

11. Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles y supongamos que $(f_n)_{n\in\mathbb{N}}$ es de Cauchy en medida, entonces $(f_n)_{n\in\mathbb{N}}$ contiene una subsucesión $(f_{n_k})_{k\in\mathbb{N}}$ que converge casi en todas partes a una función medible f y tal que, dada $\delta > 0$, existe un conjunto medible A tal que $\mu(A) < \delta$ y la sucesión $(f_{n_k})_{k\in\mathbb{N}}$ converge uniformemente a f sobre A^c.

12. Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles que converge en medida, entonces $(f_n)_{n\in\mathbb{N}}$ contiene una subsucesión $(f_{n_k})_{k\in\mathbb{N}}$ que converge casi en todas partes a una función medible f y tal que, dada $\delta > 0$, existe un conjunto medible A tal que $\mu(A) < \delta$ y la sucesión $(f_{n_k})_{k\in\mathbb{N}}$ converge uniformemente a f sobre A^c.

13. Supongamos que μ es finita y sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles, de Cauchy en medida, entonces $(f_n)_{n\in\mathbb{N}}$ converge en medida..

14. Si μ es finita y $(f_n)_{n\in\mathbb{N}}$ es una sucesión de funciones medibles que converge casi en todas a la función medible f, entonces, dada $\delta > 0$, existe un conjunto medible A tal que $\mu(A) < \delta$ y la sucesión $(f_n)_{n\in\mathbb{N}}$ converge uniformemente a f sobre A^c.

15. Sean g una función no negativa e integrable, f una función medible y $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tales que $|f_n| \leq g$ para cualquier $n \in \mathbb{N}$ y $f_n \xrightarrow{\mu} f$, entonces $\lim_{n\to\infty} \int_{\mathbb{E}} |f_n - f| \, d\mu = 0$.

Espacios L^p

Asumimos que se tiene definido un espacio de medida $(\mathbb{F}, \mathcal{F}, \mu)$.

Para $p \in (0,\infty)$, denotaremos por $\mathcal{L}^p$ al conjunto de funciones medibles f tales que $\int_{\mathbb{F}} |f|^p \, d\mu < \infty$. También, denotaremos por $\mathcal{L}^\infty$ al conjunto de funciones medibles y acotadas excepto a lo más en un conjunto de medida cero.

Obsérvese que si $f \in \mathcal{L}^p$, entonces f es finita casi en todas partes.

Para $p \in (0,\infty]$, el conjunto de clases de equivalencia en las cuales queda partido L^p, mediante la relación de equivalencia definida por la igualdad casi en todas partes, será denotado por L^p.

Cada elemento de L^p es un conjunto de funciones con la propiedad de que cualquier par de ellas son iguales casi en todas partes.

Si $f : \mathbb{F} \to \overline{\mathbb{R}}$, la notación $f \in L^p$ significará que la clase de equivalencia de la cual forma parte f, pertenece a L^p, de manera que cualquier propiedad que se demuestre para f será en realidad una propiedad de la clase de equivalencia de la cual forma parte.

Si $f \in L^\infty$, diremos que $M \in \mathbb{R}$ es cota esencial de f si $|f| \leq M$ casi en todas partes. Además, definimos el supremo esencial de f, $\sup es(f)$, de la siguiente manera:

$$\sup es(f) = \inf \{M \in \mathbb{R} : M \text{ es cota esencial de } f\}$$

Obsérvese que si $f \in L^\infty$, entonces $|f| \leq \sup es(f)$ casi en todas partes. Además, no hay ningún número real M menor que $\sup es(f)$ y tal que $|f| \leq M$ casi en todas partes. Es decir, si $|f| \leq M$ casi en todas partes, entonces $\sup es(f) \leq M$.

Si $p \in [1,\infty)$ y $f \in L^p$, definimos:

$$||f||_p = \left(\int_{\mathbb{F}} |f|^p \, d\mu\right)^{\frac{1}{p}}$$

Si $p \in (0,1)$ y $f \in L^p$, definimos:

$$||f||_p = \int_{\mathbb{F}} |f|^p \, d\mu$$

Si $f \in L^\infty$, definimos:

$$||f||_\infty = \sup es\,(|f|)$$

RESULTADOS

1. Para cualquier $p \in (0,\infty]$, L^p es un espacio vectorial sobre $\mathbb{R}$.

2. **Desigualdad de Hölder:** Sean $p, q \in [1,\infty]$ tales que $\frac{1}{p} + \frac{1}{q} = 1$, $f \in L^p$ y $g \in L^q$, entonces $fg \in L^1$ y:

$$||fg||_1 \leq ||f||_p \, ||g||_q$$

3. **Desigualdad de Minkowski:** Sea $p \in [1,\infty]$ y $f, g \in L^p$, entonces:

$$||f+g||_p \leq ||f||_p + ||g||_p$$

4. Para cualquier $p \in (0,\infty]$, la función $||\bullet||_p$, definida sobre L^p, es una norma.

5. Supongamos que la medida μ es finita, entonces, para cualquier $r \in (0,\infty]$, si $f \in L^r$, entonces $f \in L^p$ para cualquier $p \in (0,r)$.

Definición 47. *Para $p \in (0,\infty]$, se dice que una sucesión $(f_n)_{n\in\mathbb{N}}$ de funciones medibles converge en L^p a la función medible f si $f, f_n \in L^p$ para cualquier $n \in \mathbb{N}$ y:*

$$\text{lím}_{n\to\infty} ||f_n - f||_p = 0$$

Si éste es el caso, se escribirá $f_n \xrightarrow{L^p} f$.

RESULTADOS

1. La convergencia en L^p tiene las propiedades comunes a la convergencia en cualquier espacio vectorial normado. En particular, si X es un espacio vectorial normado, sobre $\mathbb{R}$, con norma $\|\bullet\|$, se tiene lo siguiente:

i. Si $(x_n)_{n\in\mathbb{N}}$ es una sucesión que converge x, entonces cualquier subsucesión de $(x_n)_{n\in\mathbb{N}}$ también converge a x.

ii. Si $(x_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de X tales que $x_n \to x$ y $x_n \to y$, entonces $x = y$.

iii. Si $c \in \mathbb{R}$ y $(x_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de X tales que $x_n \to x$, entonces $cx_n \to cx$.

iv. Si $(x_n)_{n\in\mathbb{N}}$ y $(y_n)_{n\in\mathbb{N}}$ son dos sucesiones de elementos de X tales que $x_n \to x$ y $y_n \to y$, entonces $x_n + y_n \to x + y$.

v. Si $(x_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de X tales que la serie $\sum_{n=1}^{\infty} \|x_n\|$ converge, entonces la sucesión $(y_n)_{n\in\mathbb{N}}$ definida por $y_n = \sum_{k=1}^{n} x_k$ es de Cauchy.

vi. Si para cualquier sucesión $(x_n)_{n\in\mathbb{N}}$ de elementos de X tales que la serie $\sum_{n=1}^{\infty} \|x_n\|$ converge, la sucesión $(y_n)_{n\in\mathbb{N}}$, definida por $y_n = \sum_{k=1}^{n} y_k$, converge. Entonces X es completo.

2. Para cualquier $p \in (0, \infty]$, L^p, con la norma $||\bullet||_p$ es un espacio normado completo.

3. Sea $p \in (0, \infty]$ y $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $f_n \xrightarrow{L^p} f$, entonces $f_n \xrightarrow{\mu} f$.

3. Sean $p \in (0, \infty)$, $\{f_n\}_{n\in\mathbb{N}}$ una sucesión en L^p y f una función medible. Si μ es una medida finita, las siguientes propiedades son equivalentes:

a) $f_n \xrightarrow{\mu} f$ y la familia $\{|f_n|^p : n \in \mathbb{N}\}$ es uniformemente integrable.

b) $f \in L^p$ y $f_n \xrightarrow{L^p} f$.

c) $f \in L^p$, $f_n \xrightarrow{\mu} f$ y $\lim_{n\to\infty} \int_{\mathbb{F}} |f_n|^p \, d\mu = \int_{\mathbb{F}} |f|^p \, d\mu$.

4. Supongamos que la medida μ es finita, entonces, para cualquier $r \in (0, \infty]$, si $f_n \xrightarrow{L^r} f$, entonces $f_n \xrightarrow{L^p} f$ para cualquier $p \in (0, r]$.

Densidad de las funciones simples en L^p

Lema 1. *Sea $p \in (0, \infty)$. Una función simple no negativa φ pertenece a L^p si sólo si φ es nula fuera de un conjunto de medida finita.*

Teorema 2. *Sea $p \in (0, \infty)$. El conjunto de las funciones simples nulas fuera de un conjunto de medida finita es denso en L^p.*

Demostración

Sea $f \in L^p$ y $(\varphi_n)_{n\in\mathbb{N}}$ y $(\psi_n)_{n\in\mathbb{N}}$ dos sucesiones no decrecientes de funciones simples no negativas tales que $\lim_{n\to\infty} \varphi_n(x) = f^+(x)$ y $\lim_{n\rightsquigarrow\infty} \psi_n(x) = f^-(x)$ para cualquier $x \in \mathbb{F}$. Sea $A = \{y \in: |f(y)| < \infty\}$. Entonces, para cualquier $n \in \mathbb{N}$, $\varphi_n I_A$ y $\psi_n I_A$ siguen siendo simples.

Para cualquier $n \in \mathbb{N}$ se tiene $\varphi_n I_A \leq f^+$ y $\psi_n I_A \leq f^-$, así que $\varphi_n I_A \in L^p$ y $\psi_n I_A \in L^p$. Por lo tanto, $\varphi_n I_A$ y $\psi_n I_A$ son nulas fuera de un conjunto de medida finita.

Además, como $\mu(A^c) = 0$, $\lim_{n\to\infty} \varphi_n I_A = f^+$ y $\lim_{n\to\infty} \psi_n I_A = f^-$ casi en todas partes.

Para cada $n \in \mathbb{N}$ definamos $s_n = \varphi_n I_A - \psi_n I_A$. Entonces:

$\lim_{n\to\infty} s_n = f^+ - f^- = f$ casi en todas partes. Así que:

$\lim_{n\to\infty} |f - s_n|^p = 0$ casi en todas partes.

Además:

$$|s_n| = |\varphi_n I_A - \psi_n I_A| \leq \varphi_n + \psi_n = |f|$$

Así que:

$$|f - s_n|^p \leq 2^p \left(|f|^p + |s_n|^p\right) \leq 2^{p+1} |f|^p$$

Por lo tanto, aplicando el teorema de la convergencia dominada, se tiene:

$$\lim_{n\to\infty} \int_{\mathbb{F}} |f - s_n|^p \, d\mu = 0$$

Así que la sucesión $(s_n)_{n\in\mathbb{N}}$ converge a f en L^p.

■

Espacios de Probabilidad

Definición 48 (**Espacio de probabilidad**). *Llamaremos espacio de probabilidad a una terna $(\Omega, \Im, P)$, donde Ω es un conjunto, $\Im$ una σ-álgebra de subconjuntos de Ω y P una medida sobre $\Im$ tal que $P(\Omega) = 1$, a la cual llamaremos medida de probabilidad. A Ω lo llamaremos el espacio muestral, a los elementos de $\Im$ eventos y a la medida P de un evento A la probabilidad de A.*

Considerando que cualquier espacio de medida se puede completar, asumiremos que cualquier medida de probabilidad con la que trabajemos es completa.

En lo que sigue asumimos que tenemos definido un espacio de probabilidad completo $(\Omega, \Im, P)$.

Variables Aleatorias

Definición 49. *Llamaremos variable aleatoria real a cualquier función medible de $(\Omega, \Im)$ en $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. Una variable aleatoria con valores en el conjunto de números reales extendido será una función medible de $(\Omega, \Im)$ en $\left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$. Un vector aleatorio real será cualquier función medible de $(\Omega, \Im)$ en $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$.*

Obviamente, una variable aleatoria real puede considerarse también como una función de $(\Omega, \Im)$ en $\left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$, y esta función es medible.

Dado un conjunto finito de variables aleatorias, $X_1, X_2, \ldots, X_n$, con valores en $\overline{\mathbb{R}}^n$, la función $(X_1, X_2, \ldots, X_n) : \Omega \to \overline{\mathbb{R}}^n$ definida por:

$$(X_1, X_2, \ldots, X_n)(\omega) = (X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$$

es medible. A una función así definida la llamaremos vector aleatorio con valores en $\overline{\mathbb{R}}^n$. También usaremos la notación $\overline{X}$ para un vector aleatorio $(X_1, X_2, \ldots, X_n)$.

A menos que se indique otra cosa, una variable aleatoria (resp. vector aleatorio con n componentes) será considerada con valores en $\overline{\mathbb{R}}$ (resp. con valores en $\overline{\mathbb{R}}^n$).

En lo que se refiere a la convergencia de una sucesión de variables aleatorias, hay algunos cambios en la terminología. En el contexto de la Teoría de la Probabilidad, la convergencia casi en todas partes será denominada **convergencia casi segura** y a la convergencia en medida la llamaremos **convergencia en probabilidad**.

Dada una variable aleatoria $X : \Omega \to \overline{\mathbb{R}}$ y $B \in \mathfrak{B}\left(\overline{\mathbb{R}}\right)$, la notación $[X \in B]$ será una manera abreviada de representar al conjunto $\{\omega \in \Omega : X(\omega) \in B\}$. Si $a, b \in \overline{\mathbb{R}}$, las notaciones $[a < X < b]$, $[a \leq X \leq b]$, $[a \leq X < b]$, $[a < X \leq b]$, $[X < b]$, $[X \leq b]$, $[X > a]$ y $[X \geq a]$ se entenderán en un sentido similar. Para el caso de un vector aleatorio utilizaremos una notación análoga.

Si $(X_1, X_2, \ldots, X_n)$ es un vector aleatorio con valores en $\overline{\mathbb{R}}^n$ y $B_1, B_2, \ldots, B_n$ son conjuntos borelianos en $\overline{\mathbb{R}}$, denotaremos por $[X_1 \in B, X_2 \in B_2, \ldots, X_n \in B_n]$ al conjunto $\bigcap_{k=1}^{n} [X_k \in B_n]$.

Definición 50. *Sea X una variable aleatoria con valores en $\overline{\mathbb{R}}$. La proyección de P bajo X será denotada por μ_X y la llamaremos la distribución de la variable aleatoria X. Si $(X_1, X_2, \ldots, X_n)$ es un vector aleatorio, la proyección de P bajo $(X_1, X_2, \ldots, X_n)$ será denotada por $\mu_{X_1, X_2, \ldots, X_n}$ y la llamaremos la distribución del vector aleatorio $(X_1, X_2, \ldots, X_n)$.*

RESULTADOS

1. Sea X una variable aleatoria con valores en $\overline{\mathbb{R}}$. La función $\mu_X : \mathfrak{B}\left(\overline{\mathbb{R}}\right) \mapsto \mathbb{R}$ definida por:

$$\mu_X(B) = P[X \in B]$$

es una medida de probabilidad.

2. Si X es una variable aleatoria con valores en $\overline{\mathbb{R}}$, μ_X su distribución y $f : \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right), \mu_X\right) \mapsto \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ una función medible, no negativa o integrable, se tiene:

$$\int_{\nleq} f(X)\, dP = \int_{\overline{\mathbb{R}}} f d\mu_X$$

3. Si X es una variable aleatoria real y la consideramos como una función de $(\Omega, \Im)$ en $\left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$, entonces X es una variable aleatoria con valores en el conjunto de números reales extendido y su distribución μ_X, aunque está definida sobre $\mathfrak{B}\left(\overline{\mathbb{R}}\right)$, está concentrada en $\mathbb{R}$ ya que $\mu_X(\{-\infty, \infty\}) = 0$.

4. Si $(X_1, X_2, \ldots, X_n)$ es una vector aleatorio con valores en $\overline{\mathbb{R}}^n$, $\mu_{X_1,X_2,\ldots,X_n}$ su distribución y $f : \left(\overline{\mathbb{R}}^n, \mathfrak{B}\left(\overline{\mathbb{R}}^n\right), \mu_{X_1,X_2,\ldots,X_n}\right) \mapsto \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ una función medible, no negativa o integrable, se tiene:

$$\int_{\nleq} f(X_1, X_2, \ldots, X_n)\, dP = \int_{\overline{\mathbb{R}}} f d\mu_{X_1,X_2,\ldots,X_n}$$

Definición 51 (σ álgebra generada por una familia de funciones). *Sea* $(\mathbb{E}, \mathfrak{E})$ *un espacio medible y* $\mathbb{F}$ *un conjunto cualquiera. Dada una colección de funciones:*

$$\mathcal{H} = \{f_\gamma : \mathbb{F} \to (\mathbb{E}, \mathfrak{E}) : \gamma \in \Gamma\}$$

donde Γ es un conjunto de índices cualquiera, se define la σ-álgebra generada por $\mathcal{H}$ como la más pequeña σ-álgebra de subconjuntos de $\mathbb{F}$ tal que toda función $f \in \mathcal{H}$ es medible. Denotaremos a esta σ-álgebra por por $\sigma(\mathcal{H})$ o por $\sigma(\{f_\gamma : \gamma \in \Gamma\})$.

Obsérvese que si $f : \mathbb{F} \to (\mathbb{E}, \mathfrak{E})$ es cualquier función, la familia de conjuntos $\{f^{-1}(B) : B \in \mathfrak{E}\}$ es una σ-álgebra de subconjuntos de $\mathbb{F}$. Sin embargo, si $\{f_\gamma : \mathbb{F} \to (\mathbb{E}, \mathfrak{E}) : \gamma \in \Gamma\}$ es una colección de funciones, la familia de conjuntos $\left\{f_\gamma^{-1}(B) : \gamma \in \Gamma \text{ y } B \in \mathfrak{E}\right\}$ no es, en general, una σ-álgebra, pero la σ-álgebra generada por esa familia de conjuntos es la σ-álgebra generada por la familia de funciones $\{f_\gamma : \mathbb{F} \to (\mathbb{E}, \mathfrak{E}) : \gamma \in \Gamma\}$.

Por otra parte, si $f_1, f_2, \ldots, f_n$ son funciones de $\mathbb{F}$ en $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ y definimos $f : \mathbb{F} \to \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}^n\right)\right)$ mediante la relación:

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$$

sabemos que si $\Im$ es una σ-álgebra de subconjuntos de $\mathbb{F}$, entonces f es $\Im$-medible si y sólo si f_k es $\Im$-medible para cualquier $k \in \{1, 2, \ldots, n\}$. Por lo tanto:

$$\left\{f^{-1}(B) : B \in \mathfrak{B}\left(\overline{\mathbb{R}}^n\right)\right\} = \sigma(f) = \sigma(\{f_1, f_2, \ldots, f_n\})$$

Proposición 1. *Sean* $Y_1, \ldots, Y_n$ *n variables aleatorias con valores en* $\overline{\mathbb{R}}$ *y* $Z : \Omega \mapsto \overline{\mathbb{R}}$ *una función* $\sigma(Y_1, \ldots, Y_n)$*-medible. Entonces, existe una función boreliana* $h : \overline{\mathbb{R}}^n \mapsto \overline{\mathbb{R}}$ *tal que* $Z = h(Y_1, \ldots, Y_n)$.

Demostración

Si $Z = I_E$, donde $E \in \sigma(Y_1, \ldots, Y_n)$, entonces existe un boreliano $B \subset \overline{\mathbb{R}}^n$ tal que $Z = I_B(Y_1, \ldots, Y_n)$. Por lo tanto, se tiene el resultado para el caso de una función simple Z $\sigma(Y_1, \ldots, Y_n)$-medible.

Si Z es una variable aleatoria no negativa, sea $(Z_n)_{n\in\mathbb{N}}$ una sucesión no decreciente de funciones simples no negativas tales que $Z = \text{lím}_{n\rightsquigarrow\infty} Z_n$ y, para cada $n \in \mathbb{N}$, sea $h_n : \overline{\mathbb{R}}^n \mapsto \overline{\mathbb{R}}$ una función boreliana no negativa tal que $Z_n = h_n(Y_1, \ldots, Y_n)$.

Sea $D = \left\{x \in \overline{\mathbb{R}}^n : \text{lím}_{n\rightsquigarrow\infty} h_n(x) \text{ existe}\right\}$. Entonces D es un conjunto boreliano de $\overline{\mathbb{R}}^n$ y contiene a la imagen de Ω bajo la función $(Y_1, \ldots, Y_n)$. Definamos la función $h : \overline{\mathbb{R}}^n \mapsto \overline{\mathbb{R}}$ de la siguiente manera:

$$h(x) = \begin{cases} \text{lím}_{n\rightsquigarrow\infty} h_n(x) & \text{si } x \in D \\ 0 & \text{en otro caso} \end{cases}$$

h es entonces una función boreliana y $Z = h(Y_1, \ldots, Y_n)$.

∎

Corolario 1. *Sean $Y_1, \ldots, Y_n$ n variables aleatorias con valores en $\mathbb{R}$ y $Z : \Omega \mapsto \mathbb{R}$ una función $\sigma(Y_1, \ldots, Y_n)$-medible. Entonces, existe una función boreliana $h : \mathbb{R}^n \mapsto \mathbb{R}$ tal que $Z = h(Y_1, \ldots, Y_n)$.*

Independencia de variables aleatorias

Definición 52. *Diremos que las variables aleatorias de una familia no vacía cualquiera $\{X_\gamma\}$, finita o infinita, son independientes, si dada cualquier subcolección finita de ellas, $X_{\gamma_1}, \ldots, X_{\gamma_n}$ y cualquier colección de subconjuntos borelianos de $\overline{\mathbb{R}}$, $A_1, \ldots, A_n$, donde $n \in \mathbb{N}$, se tiene:*

$$P\left[X_{\gamma_1} \in A_1, \ldots, X_n \in X_{\gamma_n}\right] = P\left[X_{\gamma_1} \in A_1\right] \cdots P\left[X_{\gamma_n} \in A_n\right]$$

RESULTADOS

1. n variables aleatorias $X_1, \ldots, X_n$, son independientes si y sólo si para cualquier colección de subconjuntos borelianos de $\overline{\mathbb{R}}$, $A_1, \ldots, A_n$, se tiene:

$$P[X_1 \in A_1, \ldots, X_n \in A_n] = P[X_1 \in A_1] \cdots P[X_n \in A_n].$$

2. Las variables aleatorias de una familia infinita numerable, $\{X_1, X_2, \ldots\}$, son independientes si y sólo si para cualquier $n \in \mathbb{N}$ y cualquier colección de subconjuntos borelianos en $\overline{\mathbb{R}}$, $A_1, \ldots, A_n$, se tiene:

$$P[X_1 \in A_1, \ldots, X_n \in A_n] = P[X_1 \in A_1] \cdots P[X_n \in A_n].$$

3. Sean $X_1, \ldots, X_n$ n variables aleatorias independientes y $f_1, \ldots, f_n$ n funciones borelianas de $\overline{\mathbb{R}}$ en $\overline{\mathbb{R}}$. Entonces las variables aleatorias $f_1(X_1), \ldots, f_n(X_n)$ son independientes.

4. Sean $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{n+m}$ $n+m$ variables aleatorias independientes y $f : \overline{\mathbb{R}}^n \mapsto \overline{\mathbb{R}}$ y $g : \overline{\mathbb{R}}^m \mapsto \overline{\mathbb{R}}$ dos funciones borelianas. Entonces, las variables aleatorias $f(X_1, \ldots, X_n)$ y $g(X_{n+1}, \ldots, X_{n+m})$ son independientes.

Funciones de distribución

Definición 53 (**Función de distribución**). *Si X es una variable aleatoria real, la función $F_X : \mathbb{R} \mapsto \mathbb{R}$, definida por $F_X(x) = P[X \leq x]$, es llamada la función de distribución de X.*

Definición 54 (**Función de distribución conjunta**). *Sean $X_1, \ldots, X_n$ n variables aleatorias reales. La función $F_{X_1,\ldots,X_n} : \mathbb{R}^n \mapsto \mathbb{R}$, definida por:*

$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P[X_1 \leq x_1, \ldots, X_n \leq x_n]$

será llamada la función de distribución conjunta de $X_1, \ldots, X_n$.

RESULTADOS

1. Sea X una variable aleatoria real y F_X su función de distribución, entonces:

i. F_X es una función no decreciente y continua por la derecha.

ii. $\lim_{x \rightsquigarrow \infty} F_X(x) = 1$

iii. $\lim_{x \to -\infty} F_X(x) = 0$

iv. $F_X(x-) = P[X < x]$ para cualquier $x \in \mathbb{R}$.

2. Sean $X_1, \ldots, X_n$ n variables aleatorias y sea $F_{X_1,\ldots,X_n}$ su función de distribución conjunta, entonces, para cada:

$(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$, se tiene:

i. La función $x \mapsto F_{X_1,\ldots,X_n}(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$, definida sobre $\mathbb{R}$, es no decreciente y continua por la derecha.

ii. $\lim_{x \rightsquigarrow \infty} F_{X_1,\ldots,X_n}(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$

$= F_{X_1,,\ldots,X_{j-1},X_{j+1},\ldots,X_n}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$.

iii. $\lim_{x \to -\infty} F_{X_1,\ldots,X_n}(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n) = 0$.

Una familia de variables aleatorias $X_1, \ldots, X_n$ puede verse como la función de Ω en $\mathbb{R}^n$ que asigna a cada $\omega \in \Omega$ el vector $(X_1(\omega), \ldots, X_n(\omega))$; de esta forma, podemos decir que las variables aleatorias forman un **vector aleatorio** $(X_1, \ldots, X_n)$.

Recordemos que un rectángulo en $\mathbb{R}^n$ es un conjunto de la forma $I_1 \times \cdots \times I_n$, en donde $I_1, \cdots, I_n$ son intervalos en $\mathbb{R}$.

Si $R = I_1 \times \cdots \times I_n$ es un rectángulo en $\mathbb{R}^n$ y $a_1, b_1, \ldots, a_n, b_n$ son los extremos de $I_1, \cdots, I_n$, respectivamente, los intervalos I_k serán llamados los lados del rectángulo y los puntos del conjunto $V_{(a_1,b_1,\ldots,a_n,b_n)} = \{(x_1, \ldots, x_n) : x_k \in \{a_k, b_k\} \text{ para toda } k\}$ serán llamados los vértices del rectángulo.

El rectángulo $(a_1, b_1] \times \cdots \times (a_n, b_n]$ será denotado por $R_{(a_1,b_1,\ldots,a_n,b_n)}$ y $S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}$ denotará al conjunto:

$\{(x_1, \cdots, x_n) : x_i = a_i \text{ para } k \text{ índices } i \text{ y } x_i = b_i \text{ para el resto de índices}\}$.

Obviamente, $V_{(a_1,b_1,\ldots,a_n,b_n)} = \bigcup_{k=0}^n S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}$.

3. Sean $X_1, \ldots, X_n$ n variables aleatorias y $(a_1, b_1] \times \cdots \times (a_n, b_n]$ un rectángulo de $\mathbb{R}^n$.

Entonces, para cualquier evento A se tiene:

$P([a_1 < X_1 \leq b_1, \cdots, a_n < X_n \leq b_n] \cap A)$

$= \sum_{k=0}^n (-1)^k \sum_{\left\{(x_1,\cdots,x_n) \in S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}\right\}} P([X_1 \leq x_1, \cdots, X_n \leq x_n] \cap A)$.

4. Sean $X_1, \ldots, X_n$ n variables aleatorias y $(a_1, b_1] \times \cdots \times (a_n, b_n]$ un rectángulo de $\mathbb{R}^n$. Entonces:

$P([a_1 < X_1 \leq b_1, \cdots, a_n < X_n \leq b_n])$

$= \sum_{k=0}^n (-1)^k \sum_{\left\{(x_1,\cdots,x_n) \in S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}\right\}} F_{X_1,\ldots,X_n}(x_1, \cdots, x_n)$.

5. Sean $X_1, \ldots, X_n$ n variables aleatorias y $F_{X_1,\ldots,X_n}$ su función de distribución conjunta, entonces:

$$\sum_{k=0}^n (-1)^k \sum_{\left\{(x_1,\cdots,x_n) \in S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}\right\}} F_{X_1,\ldots,X_n}(x_1, \cdots, x_n) \geq 0$$

para cualquier rectángulo $(a_1, b_1] \times \cdots \times (a_n, b_n]$.

6. Sean $X_1, \ldots, X_n$ n variables aleatorias y $F_{X_1,\ldots,X_n}$ su función de distribución conjunta, entonces:

$$\lim_{m \to 0} F_{X_1,\ldots,X_n}\left(x_1 + \delta_1^{(m)}, \cdots, x_n + \delta_n^{(m)}\right) = F_{X_1,\ldots,X_n}(x_1, \cdots, x_n)$$

para cualquier vector $(x_1, \cdots, x_n) \in \mathbb{R}^n$ y cualquier sucesión $\left(\left(\delta_1^{(m)}, \cdots, \delta_n^{(m)}\right)\right)_{m\in\mathbb{N}}$ que converja al vector $\overline{0} \in \mathbb{R}^n$ y tal que $\delta_1^{(m)}, \cdots, \delta_n^{(m)}$ sean números reales positivos.

7. n variables aleatorias reales, $X_1, \ldots, X_n$, son independientes si y sólo si:

$$F_{X_1,\ldots,X_n}(x_1, \cdots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n)$$

para cualquier vector $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Funciones de distribución como medidas

Definición 55. *Diremos que una función $F : \mathbb{R} \mapsto \mathbb{R}$ es una función de distribución finita en 1 variable si satisface las siguientes propiedades:*

i. F es una función no decreciente y continua por la derecha.

ii. $\lim_{x\rightsquigarrow\infty} F_X(x) < \infty$

iii. $\lim_{x\rightsquigarrow-\infty} F_X(x) = 0$

Definición 56. *Para $n \in \{2, 3, \ldots\}$, diremos que una función $F : \mathbb{R}^n \mapsto \mathbb{R}$ es una función de distribución finita en n variables si satisface las siguientes propiedades:*

i. $\sum_{k=0}^{n} (-1)^k \sum_{\left\{(x_1,\cdots,x_n)\in S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}\right\}} F(x_1, \cdots, x_n) \geq 0$

para cualquier rectángulo $(a_1, b_1] \times \cdots \times (a_n, b_n]$.

ii. $\lim_{m\to 0} F\left(x_1 + \delta_1^{(m)}, \cdots, x_n + \delta_n^{(m)}\right) = F(x_1, \cdots, x_n)$

para cualquier vector $(x_1, \cdots, x_n) \in \mathbb{R}^n$ y cualquier sucesión $\left(\left(\delta_1^{(m)}, \cdots, \delta_n^{(m)}\right)\right)_{m\in\mathbb{N}}$ que converja al vector $\overline{0} \in \mathbb{R}^n$ y tal que $\delta_1^{(m)}, \cdots, \delta_n^{(m)}$ sean números reales positivos.

iii. $\lim_{x\to-\infty} F(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n) = 0$

para cualquier $(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$.

iv. Para cada $(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$, el límite

$\lim_{x\to\infty} F(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$

existe y la función $G : \mathbb{R}^{n-1} \mapsto \mathbb{R}$ definida por:

$G(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = \lim_{x\to\infty} F(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$

es una función de distribución finita en $n - 1$ variables.

Cuando $\lim_{(x_1,\cdots,x_n)\to(\infty,\ldots,\infty)} F(x_1,\cdots,x_n)=1$, diremos simplemente que F es una función de distribución en n variables.

Si $a,b\in\overline{\mathbb{R}}$ y $a\leq b$, entonces definimos $(a,b|$ de la siguiente manera:

$$(a,b|=\begin{cases}(a,b] & \text{si } b\in\mathbb{R}\\ (a,b) & \text{si } b=\infty\end{cases}$$

Si $F:\mathbb{R}^n\mapsto\mathbb{R}$ es una función de distribución finita y

$(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\in\mathbb{R}^{n-1}$, definimos:

$F(x_1,\ldots,x_{j-1},\infty,x_{j+1},\ldots,x_n)=\lim_{x\to\infty} F(x_1,\ldots,x_{j-1},x,x_{j+1},\ldots,x_n)$.

$F(x_1,\ldots,x_{j-1},-\infty,x_{j+1},\ldots,x_n)=0$.

Con estas convenciones, se tiene que:

$\sum_{k=0}^{n}(-1)^k\sum_{\left\{(x_1,\cdots,x_n)\in S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}\right\}} F_{X_1,\ldots,X_n}(x_1,\cdots,x_n)\geq 0$

para cualquier rectángulo $(a_1,b_1|\times\cdots\times(a_n,b_n|$.

Definición 57. *Si $F:\mathbb{R}^n\mapsto\mathbb{R}$ es una función de distribución finita y $R=(a_1,b_1|\times\cdots\times(a_n,b_n|$ es un rectángulo en $\mathbb{R}^n$, definimos $\mu_F(R)$ de la siguiente manera:*

$$\mu_F(R)=\sum_{k=0}^{n}(-1)^k\sum_{\left\{(x_1,\cdots,x_n)\in S^{(k)}_{(a_1,b_1,\ldots,a_n,b_n)}\right\}} F(x_1,\cdots,x_n)$$

RESULTADOS

1. Sea $F:\mathbb{R}^n\mapsto\mathbb{R}$ una función de distribución finita y $R=(a_1,b_1|\times\cdots\times(a_n,b_n|$ un rectángulo en $\mathbb{R}^n$. Para cada intervalo $(a_i,b_i|$ consideremos una partición:

$$P_i=\left\{a_i=c_0^{(i)}<c_1^{(i)}<\cdots<c_{m_i}^{(i)}=b_i\right\}$$

Entonces:

$$\mu_F(R)=\sum_{j_i\in\{1,\ldots,m_i\}}\mu_F\left(R_{\left(c^{(1)}_{j_1-1},c^{(1)}_{j_1},\ldots,c^{(n)}_{j_n-1},c^{(n)}_{j_n}\right)}\right)$$

2. Sea $F:\mathbb{R}^n\mapsto\mathbb{R}$ una función de distribución finita, $R=(a_1,b_1|\times\cdots\times(a_n,b_n|$ un rectángulo en $\mathbb{R}^n$ y $R^{(j)}=\left(a_1^{(j)},b_1^{(j)}\right|\times\cdots\times\left(a_n^{(j)},b_n^{(j)}\right|$ una colección finita de rectángulos en $\mathbb{R}^n$, ajenos por parejas, tal que $R=\bigcup_{j=1}^{m}R^{(j)}$, entonces:

$$\mu_F(R)=\sum_{j=1}^{m}\mu_F\left(R^{(j)}\right)$$

3. Sea $F : \mathbb{R}^n \mapsto \mathbb{R}$ es una función de distribución finita, $R = (a_1, b_1| \times \cdots \times (a_n, b_n|$ un rectángulo en $\mathbb{R}^n$ y $R^{(j)} = \left(a_1^{(j)}, b_1^{(j)}\right| \times \cdots \times \left(a_n^{(j)}, b_n^{(j)}\right|$ una colección finita de rectángulos en $\mathbb{R}^n$ tal que $R \subset \bigcup_{j=1}^{m} R^{(j)}$, entonces:

$$\mu_F(R) \leq \sum_{j=1}^{m} \mu_F\left(R^{(j)}\right)$$

4. Sea $F : \mathbb{R}^n \mapsto \mathbb{R}$ una función de distribución finita y $R_1, \ldots, R_k$ y $R^{(1)}, \ldots, R^{(m)}$ dos colecciones finitas de rectángulos en $\mathbb{R}^n$, todos de la forma:

$$\left(a_1^{(1)}, b_1^{(1)}\right| \times \cdots \times \left(a_n^{(1)}, b_n^{(1)}\right|$$

y tales que $R_1, \ldots, R_k$ son ajenos por parejas, $R^{(1)}, \ldots, R^{(m)}$ son ajenos por parejas y $\bigcup_{i=1}^{k} R_i = \bigcup_{j=1}^{m} R^{(j)}$.

Entonces:

$$\sum_{i=1}^{k} \mu_F(R_i) = \sum_{j=1}^{m} \mu_F\left(R^{(j)}\right)$$

5. Sea $F : \mathbb{R}^n \mapsto \mathbb{R}$ una función de distribución finita, $R = (a_1, b_1| \times \cdots \times (a_n, b_n|$ un rectángulo en $\mathbb{R}^n$ y $R^{(i)} = \left(a_1^{(i)}, b_1^{(i)}\right| \times \cdots \times \left(a_n^{(i)}, b_n^{(i)}\right|$ una colección infinita de rectángulos en $\mathbb{R}^n$ tal que $R \subset \bigcup_{i=1}^{\infty} R^{(i)}$, entonces:

$$\mu_F(R) \leq \sum_{i=1}^{\infty} \mu_F\left(R^{(i)}\right)$$

6. Sea $F : \mathbb{R}^n \mapsto \mathbb{R}$ una función de distribución finita. Entonces existe una única medida finita μ_F, definida sobre los conjuntos borelianos de $\mathbb{R}^n$, tal que:

$$\mu_F((-\infty, x_1] \times \cdots \times (-\infty, x_n]) = F(x_1, \ldots, x_n)$$

para cualquier $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

7. Sea $F : \mathbb{R}^n \mapsto \mathbb{R}$ una función de distribución. Entonces existe un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y una familia $X_1, \ldots, X_n$ de variables aleatorias reales definidas sobre Ω tal que F es la función de distribución conjunta de $X_1, \ldots, X_n$.

8. Si μ es una medida finita definida sobre los conjuntos borelianos de $\mathbb{R}^n$, la función F : $\mathbb{R}^n \mapsto \mathbb{R}$ definida por:

$$F(x_1, \ldots, x_n) = \mu((-\infty, x_1] \times \cdots \times (-\infty, x_n])$$

es una función de distribución finita y la medida μ_F que genera sobre los borelianos, por ser única, coincide con μ. De esta forma, toda medida finita μ definida sobre los conjuntos borelianos de $\mathbb{R}^n$ está generada por una función de distribución finita en n variables.

Esperanza de variables aleatorias

Definición 58. *Si X es una variable aleatoria no negativa, definimos la esperanza de X, $E[X]$, de la siguiente manera:*

$$E[X] = \int_{\Omega} X dP$$

Definición 59. *Diremos que la variable aleatoria X tiene esperanza finita si $E[|X|] < \infty$. En ese caso definimos su esperanza de la siguiente manera:*

$$E[X] = E[X^+] - E[X^-]$$

La Esperanza de una variable aleatoria tiene las mismas propiedades que la integral, es decir, se tiene la linealidad, el teorema de la convergencia monótona, el lema de Fatou, etcétera.

Sea $\mu_X : \mathfrak{B}\left(\overline{\mathbb{R}}\right) \mapsto \mathbb{R}$ la medida de probabilidad definida por:

$$\mu_X(B) = P[X \in B]$$

$\left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right), \mu_X\right)$ es un espacio de probabilidad, el cual podemos completar, obteniendo una nueva $\sigma-$álgebra, la cual denotaremos por $\mathfrak{B}_X\left(\overline{\mathbb{R}}\right)$, y una medida definida sobre $\mathfrak{B}_X\left(\overline{\mathbb{R}}\right)$, para la cual no cambiaremos la notación; es decir, la denotaremos por μ_X.

Tenemos así un espacio de probabilidad completo $\left(\overline{\mathbb{R}}, \mathfrak{B}_X\left(\overline{\mathbb{R}}\right), \mu_X\right)$, de manera que podemos definir la integral de las funciones medibles $f : \left(\overline{\mathbb{R}}, \mathfrak{B}_X\left(\overline{\mathbb{R}}\right), \mu_X\right) \mapsto \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$ como con cualquier otra medida de probabilidad y los resultados expuestos acerca de la integral son válidos también en este caso.

Para evitar confusiones, utilizaremos el término variable aleatoria únicamente para las funciones medibles $X : (\Omega, \Im, P) \to \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$.

RESULTADOS

1. Si X es una variable aleatoria, entonces $E[f(X)] = \int_{\overline{\mathbb{R}}} f d\mu_X$ para cualquier función medible no negativa $f : \left(\overline{\mathbb{R}}, \mathfrak{B}_X\left(\overline{\mathbb{R}}\right)\right) \mapsto \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$.

2. Si X es una variable aleatoria, entonces $E[f(X)] = \int_{\overline{\mathbb{R}}} f d\mu_X$ para cualquier función integrable $f : \left(\overline{\mathbb{R}}, \mathfrak{B}_X\left(\overline{\mathbb{R}}\right), \mu_X\right) \mapsto \left(\overline{\mathbb{R}}, \mathfrak{B}\left(\overline{\mathbb{R}}\right)\right)$.

3. Para cualquier variable aleatoria X, el conjunto $\{x \in \mathbb{R} : P[X = x] > 0\}$ es a lo más infinito numerable.

4. Sean X y Y dos variables aleatorias independientes de esperanza finita, entonces XY también tiene esperanza finita y $E[XY] = E[X]E[Y]$.

5. Sean $X_1, \ldots, X_n$ n variables aleatorias independientes de esperanza finita, entonces $\prod_{k=1}^{n} X_k$ también tiene esperanza finita y:

$$E\left[\prod_{k=1}^{n} X_k\right] = \prod_{k=1}^{n} E[X_k]$$

6. Sea X una variable aleatoria, entonces X tiene esperanza finita si y sólo si:

$$(R)\int_0^\infty [1 - F_X(x)]\, dx < \infty \text{ y } (R)\int_{-\infty}^0 F_X(x)dx < \infty$$

y, en ese caso, se tiene:

$$E[X] = (R)\int_0^\infty [1 - F_X(x)]\, dx - (R)\int_{-\infty}^0 F_X(x)dx$$

Donde (R) significa que las integrales se pueden tomar en el sentido de Riemann.

Cuando se tiene $(R)\int_0^\infty P[X > x]\, dx = \infty$ y $(R)\int_{-\infty}^0 P[X < x]\, dx < \infty$, se define $E[X] = \infty$, mientras que cuando $(R)\int_0^\infty P[X > x]\, dx < \infty$ y $(R)\int_{-\infty}^0 P[X < x]\, dx = \infty$, se define $E[X] = -\infty$. Cuando ambas integrales sean divergentes, entonces la esperanza de X no está definida.

Varianza y covarianza

Definición 60. *Sea X una variable aleatoria de esperanza finita. Se define la* ***varianza*** *de X, $Var(X)$, mediante la relación:*

$$Var(X) = E\left[(X - E(X))^2\right]$$

Definición 61. *A la raíz cuadrada no negativa de la varianza se le llama la* ***desviación estándar*** *de X.*

Definición 62. *Diremos que una variable aleatoria X tiene varianza finita si se cumplen las siguientes dos condiciones:*

i. X tiene esperanza finita.

ii. $(X - E[X])^2$ tiene esperanza finita.

Definición 63. *Sean X y Y dos variables aleatorias de varianza finita. Se define la **covarianza** de X y Y, $Cov(X,Y)$, mediante la relación:*

$$Cov(X,Y) = E\left[(X - E[X])(Y - E[Y])\right] = E[XY] - E[X]E[Y]$$

RESULTADOS

1. Una variable aleatoria X tiene varianza finita si y sólo si X^2 tiene esperanza finita.

2. Sea X una variable aleatoria de esperanza finita, entonces:

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

3. Sean X y Y dos variables aleatorias de varianza finita. Entonces, XY tiene esperanza finita.

4. Si X y Y son dos variables aleatorias de varianza finita y a y b son dos números reales cualesquiera, entonces $aX + bY$ tiene varianza finita.

5. Sean X y Y dos variables aleatorias independientes de varianza finita, entonces $Cov(X,Y) = 0$.

6. Sean X y Y dos variables aleatorias de varianza finita y a y b dos números reales cualesquiera. Entonces $Cov(aX, bY) = abCov(X,Y)$.

7. Sean $X, X_1, \ldots, X_n$ $n+1$ variables aleatorias de esperanza finita. Entonces:

a) $Var(X) = 0$ si y sólo si existe una constante c tal que $P[X = c] = 1$.

b) $Var(aX + b) = a^2 Var(X)$ para cualquier pareja $a, b \in \mathbb{R}$.

c) Si $X_1, \ldots, X_n$ tienen varianza finita, entonces $\sum_{i=1}^n X_i$ también tiene varianza finita y:

$$Var(\textstyle\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i) + 2\sum_{\{i,j \in \{1,\ldots,n\}: i<j\}} Cov(X_i, X_j)$$

8. Sean $X_1, \ldots, X_n$ n variables aleatorias independientes y de varianza finita, entonces $\sum_{i=1}^n X_i$ también tiene varianza finita y:

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

9. **Desigualdad de Cauchy-Schwarz:** Sean X y Y dos variables aleatorias cualesquiera, entonces:

$$E[|XY|] \leq \sqrt{E[X^2]}\sqrt{E[Y^2]}$$

Además, si X y Y tienen varianza finita, entonces $|E[XY]| = \sqrt{E[X^2]}\sqrt{E[Y^2]}$ si y sólo si existen constantes a y b tales que por lo menos una de ellas es distinta de cero y $P[aX + bY = 0] = 1$.

10. Sean X y Y dos variables aleatorias de varianza finita. Entonces:

$$|Cov(X,Y)| \leq \sqrt{Var(X)}\sqrt{Var(Y)}$$

Además, la igualdad se cumple si y sólo si existen constantes a, b y c tales que a y b no son ambas cero y $P[aX + bY = c] = 1$.

Regularidad de la medidas sobre los conjuntos borelianos de $\mathbb{R}^n$

RESULTADOS

1. Sea μ una medida finita, definida sobre los conjuntos borelianos de $\mathbb{R}^n$, Entonces, para cualquier conjunto boreliano B de $\mathbb{R}^n$ se tiene:

$$\mu(B) = \text{ínf}\{\mu(O) : B \subset O \text{ y } O \text{ es un abierto de } \mathbb{R}^n\}$$

2. Sea μ una medida finita, definida sobre los conjuntos borelianos de $\mathbb{R}^n$, Entonces, para cualquier conjunto boreliano B de $\mathbb{R}^n$ se tiene:

$$\mu(B) = \sup\{\mu(K) : K \subset B \text{ y } K \text{ es un compacto de } \mathbb{R}^n\}$$

Conjuntos compactos

Definición 64. *Sea (X,d) un espacio métrico.*

i) Diremos que $K \subset X$ es compacto si para cualquier familia infinita $\{G_\gamma : \gamma \in \Gamma\}$ de subconjuntos abiertos de X tales que $K \subset \bigcup_{\gamma\in\Gamma} G_\gamma$, existe un conjunto finito $T \subset \Gamma$ tal que $K \subset \bigcup_{\gamma\in T} G_\gamma$.

ii) Diremos que $K \subset X$ es secuencialmente compacto si para toda sucesión $(x_n)_{n\in\mathbb{N}}$ de elementos de K existe una subsucesión que converge a algún elemento de K.

RESULTADOS

1. Sea (X,d) un espacio métrico y K un subconjunto de X. Las siguientes propiedades son equivalentes:

a) K es compacto.

b) K es secuencialmente compacto.

Teorema de consistencia de Kolmogorov

El teorema de Kolmogorov muestra que la propiedad de σ-aditividad de una función de probabilidad es consistente en el sentido de que si se asume como válida para el caso finito, entonces se puede extender al caso infinito.

Se puede enunciar de la manera siguiente:

Teorema 3 (Teorema de Kolmogorov). *Sea Γ un conjunto infinito y supongamos que para cada subconjunto finito u de Γ, con n elementos, se tiene una función de distribución en n variables $F_u : \mathbb{R}^n \to \mathbb{R}$ de tal forma que si v y w son dos subconjuntos finitos de Γ tales que $v \subset w$, entonces la función de distribución F_v coincide con la distribución marginal que se obtiene de F_w restringiéndola a las coordenadas en u. Entonces, existe un espacio de probabilidad $(\Omega, \Im, P)$ y una familia de variables aleatorias reales $\{X_t : t \in \Gamma\}$ definidas sobre Ω tal que si $u = \{t_1, \ldots, t_n\}$ es cualquier subconjunto finito de Γ, entonces la función de distribución conjunta de $X_{t_1}, \ldots, X_{t_n}$ es F_u.*

Obsérvese que, al inicio, no se tienen variables aleatorias ya que esto asumiría que se tiene un espacio de probabilidad en el cual están definidas y es precisamente ese espacio de probabilidad el que se construye en la demostración del teorema.

Por ejemplo, en el caso del proceso de Wiener, el cual modela el movimiento, en una dimensión, de un grano de polen que se coloca sobre agua (movimiento browniano), lo que se tiene de inicio es lo siguiente:

Imaginando que el grano de polen se mueve en un plano cartesiano infinito, considerando la proyección del movimiento sobre uno de los ejes, digamos el horizontal, se tiene una partícula que se mueve sobre una línea recta, así que, en cada tiempo $t \geq 0$, la partícula se encuentra en una posición que denotaremos por W_t. Esta posición es aleatoria ya que el movimiento del grano de polen se debe a los choques que recibe de las moléculas de agua. Lo que se busca entonces es un espacio de probabilidad en el cual se pueda definir una familia de variables aleatorias reales $\{W_t : t \in [0, \infty)\}$ de tal manera que W_t represente la posición de la particula en el tiempo t. Para esto, se parte de las propiedades que se observan en un movimiento browniano, o mejor dicho, de propiedades que se aproximan a las observaciones. Estas propiedades consisten en lo siguiente:

i. El grano de polen no se mueve a saltos, sino de una manera continua; así que, cualquiera que sea su movimiento, la función $t \to W_t$ es continua.

ii. El punto de inicio del movimiento se toma como el origen de la linea recta sobre el cual se mueve la partícula. Así que $W_0 = 0$.

iii. El movimiento que sigue la partícula a partir de su posición en un tiempo $t > 0$, es independiente de como haya sido en el intervalo $[0, t]$. Así que, Si $t_1 < t_2 < \cdots < t_n$ son números reales positivos, entonces las variables aleatorias $W_{t_1} - W_0$, $W_{t_2} - W_{t_1}$, $\ldots$, $W_{t_n} - W_{t_{n-1}}$ son independientes.

iv. Si $0 \leq s < t$, la distribución de $W_t - W_s$ es normal com parámetros $\mu = 0$ y $\sigma^2 = t - s$.

Si $t_1 < t_2 < \cdots < t_n$ son números reales positivos, sabiendo que las variables aleatorias $W_{t_1} - W_0$, $W_{t_2} - W_{t_1}$, $\ldots$, $W_{t_n} - W_{t_{n-1}}$ son independientes, podemos encontrar la función de densidad conjunta de $W_{t_1}, W_{t_2}, \ldots, W_{t_n}$ considerando la transformación $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ dada por:

$y_k = x_1 + x_2 + \cdots + x_k$, para $k \in \{1, \ldots, n\}$.

cuya inversa, ϕ, está dada por:

$x_1 = y_1$

$x_k = y_k - y_{k-1}$, para $k \in \{2, \ldots, n\}$.

En efecto, se tiene:

$(W_{t_1}, \ldots, W_{t_n}) = \varphi\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}\right)$

Así que:

$f_{W_{t_1},\ldots,W_{t_n}}(y_1, \ldots, y_n)$

$= |J_\phi(y_1, \ldots, y_n)|\, f_{W_{t_1},W_{t_2}-W_{t_1},\ldots,W_{t_n}-W_{t_{n-1}}}\left(\phi(y_1, \ldots, y_n)\right)$

$= \frac{1}{\sqrt{(2\pi)^n t_1 (t_2-t_1)\cdots(t_n-t_{n-1})}} \exp\left\{-\frac{1}{2t_1} y_1^2\right\} \exp\left\{-\frac{1}{2(t_2-t_1)}(y_2 - y_1)^2\right\} \cdots \exp\left\{-\frac{1}{2(t_n-t_{n-1})}(y_n - y_{n-1})^2\right\}$

$= \frac{1}{\sqrt{(2\pi)^n t_1 (t_2-t_1)\cdots(t_n-t_{n-1})}} \exp\left\{-\frac{1}{2}\left[\frac{1}{t_1} y_1^2 + \frac{1}{t_2-t_1}(y_2 - y_1)^2 + \cdots + \frac{1}{t_n-t_{n-1}}(y_n - y_{n-1})^2\right]\right\}$

Algo similar se puede hacer cuando $t_1 = 0$.

En otras palabras, para cada subconjunto finito $\{t_1, t_2, \ldots, t_n\}$ de números reales no negativos, se tiene la función de densidad conjunta, y, por lo tanto, la función de distribución conjunta, de $W_{t_1}, W_{t_2}, \ldots, W_{t_n}$.

De lo que se trata entonces es de que, partiendo del conocimiento de esas distribuciones conjuntas, se pueda construir un espacio de probabilidad en el cual se pueda definir toda la familia de variables aleatorias $\{W_t : t \in [0, \infty)\}$ de tal manera que para cada subconjunto finito de números reales no negativos $\{t_1, t_2, \ldots, t_n\}$ la función de distribución conjunta de $W_{t_1}, W_{t_2}, \ldots, W_{t_n}$ sea la que se tenía al inicio.

Regresando a la formulación general del teorema de Kolmogorov, la idea de la demostración es la siguiente:

Para cada subconjunto finito de Γ, se tiene una función de distribución finito dimensional, de tal manera que se satisface la condición de consistencia que se formula en el enunciado del teorema. Se considera entonces el producto cartesiano de tantas copias de $\mathbb{R}$ como elementos tenga Γ, es decir $\mathbb{R}^\Gamma$. Para cada subconjunto finito u de Γ se expresa $\mathbb{R}^\Gamma$ como el producto cartesiano $\mathbb{R}^u \times \mathbb{R}^{\Gamma-u}$ y se genera, sobre $\mathbb{R}^u$, una medida de probabilidad a partir de la distribución finito dimensional correspondiente al conjunto u (es decir, se genera una medida sobre los borelianos de $\mathbb{R}^u$). De lo que se trata entonces es de obtener una medida sobre $\mathbb{R}^\Gamma$ juntando todas las medidas que se obtienen sobre los conjuntos $\mathbb{R}^u$, donde u corre sobre todos los subconjuntos finitos de Γ. Esto se logra definiendo primero una cuasi medida sobre el álgebra de subconjuntos de $\mathbb{R}^\Gamma$ formada por la familia de todos los conjuntos que pertenecen a $\mathfrak{B}(\mathbb{R}^u) \times \mathbb{R}^{\Gamma-u}$ para algún subconjunto finito u de Γ. Después se aplica el teorema de extensión de Carathéodory para obtener una medida sobre la σ-álgebra generada por esa álgebra y los conjuntos de medida exterior cero.

Recordemos que si $\mathbb{F}$ es un conjunto cualquiera y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$, una quasi medida es una función no negativa $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ finitamente aditiva, σ-subaditiva y tal que $\mu(\emptyset) = 0$.

Además, si $\mathbb{F}$ es un conjunto cualquiera, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \mathbb{R}$ una función no negativa y finitamente aditiva, entonces las siguientes propiedades son equivalentes:

i) μ es σ-subaditiva.

ii) μ es σ-aditiva.

iii) Para cualquier sucesión creciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$ tales que $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$, se tiene $\mu\left(\bigcup_{n=1}^\infty A_n\right) = \text{lím}_{n\rightsquigarrow\infty} \mu(A_n)$.

iv) Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$ tales que $\bigcap_{n=1}^\infty A_n \in \mathcal{A}$, se tiene $\mu\left(\bigcap_{n=1}^\infty A_n\right) = \text{lím}_{n\rightsquigarrow\infty} \mu(A_n)$.

v) Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{A}$ tales que $\bigcap_{n=1}^\infty A_n = \emptyset$, se tiene $\text{lím}_{n\to\infty} \mu(A_n) = 0$.

Finalmente, recordemos el teorema de extensión de Carathéodory:

Si $\mathbb{F}$ es un conjunto cualquiera, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu_0 : \mathcal{A} \mapsto \overline{\mathbb{R}}$ una quasi medida, entonces existe una medida $\mu : \Im \mapsto \overline{\mathbb{R}}$ tal que $\mu(A) = \mu_0(A)$ para cualquier $A \in \mathcal{A}$, donde $\Im$ es una σ-álgebra que contiene a $\sigma(\mathcal{A})$ y a los conjuntos de medida exterior cero.

La medida μ se construye definiendo primero una medida exterior definida sobre todos los subconjuntos de $\mathbb{F}$:

$$\mu_e(A) = \text{ínf}\left\{\sum_j \mu_0(A_j) : A_1, A_2, \ldots \text{ es cubierta de } A\right\}$$

Después se definen los conjuntos medibles como aquellos conjuntos $E \subset \mathbb{F}$ para los cuales se cumple la siguiente relación:

$$\mu_e(A) = \mu_e(A \cap E) + \mu_e(A \cap E^c)$$

para cualquier conjunto $A \subset \mathbb{F}$.

La medida $\mu(E)$ de un conjunto medible E se define entonces como la medida exterior de E.

Demostración del teorema de Kolmogorov

Paso 1.

Denotemos por U a la familia de subconjuntos finitos de Γ.

Sean $\Omega = \mathbb{R}^\Gamma = \{f : \Gamma \mapsto \mathbb{R}\}$y, para cada $u \in U$, $\mathbb{R}^u = \{f : u \mapsto \mathbb{R}\}$.

Por definición, un elemento $\omega \in \Omega$ es una función de Γ en $\mathbb{R}$, sin embargo podemos también imaginar a ω como un vector el cual tiene una coordenada para cada $t \in \Gamma$.

Paso 2.

Sabemos que, para cada $u \in U$, con n elementos, se tiene una función de distribución en n variables $F_u : \mathbb{R}^u \to \mathbb{R}$.

Esta función de distribución F_u genera una medida de probabilidad P_u definida sobre los conjuntos borelianos de $\mathbb{R}^u$ tal que:

$$P_u\left((-\infty, u_1] \times \cdots \times (-\infty, u_n]\right) = F_u(u_1, \ldots, u_n)$$

para cualquier $(u_1, \ldots, u_n) \in \mathbb{R}^u$.

Paso 3.

Para cada $u = \{u_1, \ldots, u_n\} \in U$, denotemos por Π_u a la función $\Pi_u : \Omega \mapsto \mathbb{R}^u$ definida por $\Pi_u(\omega) = (\omega(u_1), \ldots, \omega(u_n))$ y, para pareja $u, v \in U$ tal que $u \subset v$, denotemos por Π_{vu} a la función $\Pi_{vu} : \mathbb{R}^v \mapsto \mathbb{R}^u$ definida por $\Pi_{vu}(f) = f_u$, donde $f_u : u \mapsto \mathbb{R}$ es la restricción de $f : v \mapsto \mathbb{R}$ a u.

Obsérvese que Π_u y Π_{vu} son simplemente proyecciones sobre un espacio de menos coordenadas al del dominio.

Paso 4.

Sabemos que si $u, v \in U$ y $u \subset v$, entonces la función de distribución F_u coincide con la distribución marginal que se obtiene de F_v restringiéndola a las coordenadas de u; es decir, si $u = \{u_1, \ldots, u_n\}$ y $v = \{u_1, \ldots, u_n, v_1, \ldots, v_m\}$, entonces:

$$F_u(u_1, \ldots, u_n) = \text{lím}_{(v_1,\ldots,v_m)\to(\infty,\ldots,\infty)} F_v(u_1, \ldots, u_n, v_1, \ldots, v_m)$$

Así que:

$$P_u\left((-\infty, u_1] \times \cdots \times (-\infty, u_n]\right) = P_v\left((-\infty, u_1] \times \cdots \times (-\infty, u_n] \times \mathbb{R} \times \cdots \times \mathbb{R}\right)$$

Pero:

$$(-\infty, u_1] \times \cdots \times (-\infty, u_n] \times \mathbb{R} \times \cdots \times \mathbb{R} = \Pi_{vu}^{-1}\left((-\infty, u_1] \times \cdots \times (-\infty, u_n]\right)$$

Por lo tanto:

$$P_u\left((-\infty, u_1] \times \cdots \times (-\infty, u_n]\right) = P_v\left(\Pi_{vu}^{-1}\left((-\infty, u_1] \times \cdots \times (-\infty, u_n]\right)\right)$$

Como los conjuntos de la forma $(-\infty, u_1] \times \cdots \times (-\infty, u_n]$ generan la $\sigma-$álgebra de Borel en $\mathbb{R}^u$, se tiene entonces:

$$P_u(B_u) = P_v\left(\Pi_{vu}^{-1}(B_u)\right)$$

para cualquier conjunto boreliano B_u de $\mathbb{R}^u$.

Paso 5.

Para cada $u \in U$, denotemos por $\mathcal{B}_u$ a la σ-álgebra de los conjuntos borelianos de $\mathbb{R}^u$ y definamos:

$$\Im_0 = \{\Pi_u^{-1}(B_u) : u \in U \text{ y } B_u \in \mathcal{B}_u\}$$

Cada elemento de $\Im_0$ es un subconjunto de Ω, es decir está formado por vectores cada uno de los cuales tiene una coordenada para cada $t \in \Gamma$; lo que caracteriza a esos vectores es que restringiéndonos a las coordenadas que corresponden a los elementos de u, se obtiene un elemento de B_u. También puede pensarse $\Pi_u^{-1}(B_u)$ como $B_u \times \mathbb{R}^{\Gamma - u}$; es decir, restringiéndonos a las coordenadas correspondientes a u, $\Pi_u^{-1}(B_u)$ es B_u, mientras que restringiéndonos a las coordenadas correspondientes a $\Gamma - u$, es $\mathbb{R}^{\Gamma - u}$.

Paso 6.

Demostremos que $\Im_0$ es un álgebra de subconjuntos de Ω:

Obviamente, $\Omega \in \Im_0$ y si $E \in \Im_0$ entonces $E^c \in \Im_0$. Por otra parte, si $E = \Pi_u^{-1}(A_u) \in \Im_0$ y $F = \Pi_v^{-1}(B_v) \in \Im_0$, sea $w = u \cup v$, $A_w = \Pi_{wu}^{-1}(A_u)$ y $B_w = \Pi_{wv}^{-1}(B_v)$, entonces:

$E \cup F = \Pi_w^{-1}(A_w \cup B_w) \in \Im_0$.

Por lo tanto, $\Im_0$ es un álgebra de subconjuntos de Ω.

Paso 7.

Definamos $P : \Im_0 \mapsto [0,1]$ de la siguiente manera:

$$P\left(\Pi_u^{-1}(B_u)\right) = P_u(B_u)$$

Observemos en primer lugar que P está bien definida. En efecto, supongamos que $\Pi_u^{-1}(B_u) = \Pi_v^{-1}(A_v)$, entonces, definiendo $w = u \cup v$, se tiene:

$$\Pi_u^{-1}(B_u) = \Pi_w^{-1}\left(\Pi_{wu}^{-1}(B_u)\right)$$

$$\Pi_v^{-1}(A_v) = \Pi_w^{-1}\left(\Pi_{wv}^{-1}(A_v)\right)$$

Así que:

$$\Pi_{wu}^{-1}(B_u) = \Pi_{wv}^{-1}(A_v)$$

Por lo tanto:

$$P_u(B_u) = P_w\left(\Pi_{wu}^{-1}(B_u)\right) = P_w\left(\Pi_{wv}^{-1}(A_v)\right) = P_v(A_v)$$

Paso 8.

$\Omega = \Pi_u^{-1}(\mathbb{R}^u)$ para cualquier $u \in U$.

Así que:

$$P(\Omega) = P\left(\Pi_u^{-1}(\mathbb{R}^u)\right) = P_u(\mathbb{R}^u) = 1$$

Paso 9.

Mostremos que P es finitamente aditiva:

En efecto, Si $E = \Pi_u^{-1}(A_u)$ y $F = \Pi_v^{-1}(B_v)$ son elementos de $\Im_0$, ajenos, sea $w = u \cup v$, $A_w = \Pi_{wu}^{-1}(A_u)$ y $B_w = \Pi_{wv}^{-1}(B_v)$, entonces A_w y B_w son ajenos y $E \cup F = \Pi_w^{-1}(A_w \cup B_w)$, así que:

$P(E \cup F) = P_w(A_w \cup B_w) = P_w(A_w) + P_w(B_w)$

$= P_u(A_u) + P_v(B_v) = P(E) + P(F)$

Paso 10.

Mostremos ahora que P es σ-subaditiva:

Para esto, basta con demostrar que si tenemos una sucesión decreciente $(E_i)_{i\in\mathbb{N}}$, de elementos de $\Im_0$, tal que $\bigcap_{i=1}^{\infty} E_i = \emptyset$, entonces $\lim_{i\to\infty} P(E_i) = 0$.

Para cada $i \in \mathbb{N}$, sea $v_i \in U$ y $A_i \in \mathcal{B}_{v_i}$ tales que $E_i = \Pi_{v_i}^{-1}(A_i)$, y definamos $u_i = \bigcup_{j=1}^{i} v_j$ y $B_i = \Pi_{u_i v_i}^{-1}(A_i)$. Entonces $B_i \in \mathcal{B}_{u_i}$, $E_i = \Pi_{u_i}^{-1}(B_i)$ y la sucesión de conjuntos $(u_n)_{n\in\mathbb{N}}$ es creciente.

Supongamos que $\varepsilon = \lim_{i\to\infty} P(E_i) > 0$.

Para cada $i \in \mathbb{N}$, $P_{u_i}(B_i)$ puede ser aproximada por medidas de compactos contenidos en B_i, en particular existe un subconjunto compacto de $\mathbb{R}^{u_i}$, contenido en B_i, tal que:

$$P_{u_i}(B_i) - P_{u_i}(K_i) < \tfrac{\varepsilon}{2^{i+1}}$$

Sea $F_i = \Pi_{u_i}^{-1}(K_i)$.

Obviamente se tiene $F_i \subset E_i$ para cualquier $i \in \mathbb{N}$, así que si demostramos que $\bigcap_{i=1}^{\infty} F_i \neq \emptyset$, habremos demostrado que $\bigcap_{i=1}^{\infty} E_i \neq \emptyset$, llegando así a una contradicción.

Para cada $j \in \mathbb{N}$ definamos:

$$H_j = \bigcap_{i=1}^{j} F_i$$

Entonces:

$$H_j \subset \bigcap_{i=1}^{j} E_i = E_j$$

Y se tiene:

$P(E_j) - P(H_j) = P(E_j) - P\left(\bigcap_{i=1}^{j} F_i\right) = P\left(\bigcup_{i=1}^{j} (E_j - F_i)\right)$

$\leq \sum_{i=1}^{j} P(E_j - F_i) \leq \sum_{i=1}^{j} P(E_i - F_i) = \sum_{i=1}^{j} [P(E_i) - P(F_i)]$

$= \sum_{i=1}^{j} [P_{u_i}(B_i) - P_{u_i}(K_i)] < \sum_{i=1}^{j} \frac{\varepsilon}{2^{i+1}} < \frac{\varepsilon}{2}$

Así que:

$$P(H_j) > P(E_j) - \tfrac{\varepsilon}{2} \geq \tfrac{\varepsilon}{2} > 0$$

Por lo tanto $H_j \neq \emptyset$ para cualquier $j \in \mathbb{N}$.

Para cada $j \in \mathbb{N}$, sea $x^{(j)} \in H_j$, entonces $x^{(j)} \in F_i = \Pi_{u_i}^{-1}(K_i)$ para $i \in \{1, \ldots, j\}$. Por lo tanto, $x_{u_i}^{(j)} = \Pi_{u_i} x^{(j)} \in K_i$ para $i \in \{1, \ldots, j\}$.

(Obsérvese que, para cualesquiera $i, j \in \mathbb{N}$, $x_{u_i}^{(j)}$ tiene como coordenadas las coordenadas de $x^{(j)}$ que corresponden a los elementos de u_i. Recordemos, además, que $u_1 \subset u_2 \subset u_3 \subset \cdots$)

Visto de otra manera, fijando $i \in \mathbb{N}$ se tiene $x_{u_i}^{(j)} \in K_i$ para cualquier $j \in \{i, i+1, \ldots\}$.

En particular, $\left(x_{u_1}^{(j)}\right)_{j \in \mathbb{N}}$ es una sucesión en K_1, así que tiene por lo menos una subsucesión convergente $\left(x_{u_1}^{(m_{(1,j)})}\right)_{j \in \mathbb{N}}$, donde se puede asumir que $m_{(1,1)} > 1$.

A su vez, $\left(x_{u_2}^{(m_{1,j})}\right)_{j \in \mathbb{N}}$ es una sucesión en K_2, así que tiene por lo menos una subsucesión convergente $\left(x_{u_2}^{(m_{(2,j)})}\right)_{j \in \mathbb{N}}$, donde se puede asumir que $m_{(2,1)} > 2$.

Además:

$$\Pi_{u_2 u_1}(x_{u_2}^{(m_{(2,j)})}) = x_{u_1}^{(m_{(2,j)})}$$

$$\text{lím}_{j \to \infty} x_{u_1}^{(m_{(2,j)})} = \text{lím}_{j \to \infty} x_{u_1}^{(m_{(1,j)})}$$

Así que, si $x_{u_1} = \text{lím}_{j \to \infty} x_{u_1}^{(m_{(1,j)})}$ y $x_{u_2} = \text{lím}_{j \to \infty} x_{u_2}^{(m_{(2,j)})}$, entonces $\Pi_{u_2 u_1}(x_{u_2}) = x_{u_1}$.

Continuando de la misma forma, obtenemos, para cada $i \in \mathbb{N}$, una sucesión convergente $\left(x_{u_i}^{(m_{(i,j)})}\right)_{j \in \mathbb{N}}$ en K_i, de tal manera que, si $x_{u_i} = \text{lím}_{j \to \infty} x_{u_i}^{(m_{(i,j)})}$, entonces $\Pi_{u_{i+1} u_i}(x_{u_{i+1}}) = x_{u_i}$.

Hemos obtenido entonces una sucesión $(x_{u_i})_{i \in \mathbb{N}}$ tal que $x_{u_i} \in K_i$ para cualquier $i \in \mathbb{N}$ y si $i, j \in \mathbb{N}$, con $i < j$, entonces $\Pi_{u_j u_i}(x_{u_j}) = x_{u_i}$.

Definamos la función $y : \cup_{i=1}^{\infty} u_i \to \mathbb{R}$ de la siguiente manera:

$$y(t) = x_{u_i}(t) \text{ si } t \in u_i$$

Denotando, para cada $j \in \mathbb{N}$, por $\Pi^{(u_j)}$ a la proyección de $\mathbb{R}^{\cup_{i=1}^{\infty} u_i}$ sobre $\mathbb{R}^{u_j}$, este elemento $y \in \mathbb{R}^{\cup_{i=1}^{\infty} u_i}$, así definido, tiene la siguiente propiedad:

$$\Pi^{(u_j)}(y) = x_{u_j} \in K_j$$

Para tener definido un punto en $\bigcap_{i=1}^{\infty} F_i$ únicamente resta completar y definiendo arbitrariamente las coordenadas que corresponden a $\Gamma - \cup_{i=1}^{\infty} u_i$; por ejemplo definamos $\omega_0 \in \Omega$ de la siguiente manera:

$$\omega_0(t) = \begin{cases} y(t) & \text{si } t \in \cup_{i=1}^{\infty} u_i \\ 0 & \text{en otro caso} \end{cases}$$

Entonces $\Pi_{u_i}(\omega_0) = x_{u_i} \in K_i$ para cualquier $i \in \mathbb{N}$, así que $\omega_0 \in \Pi_{u_i}^{-1}(K_i) = F_i$ para cualquier $i \in \mathbb{N}$, es decir:

$$\omega_0 \in \bigcap_{i=1}^{\infty} F_i$$

Así que $\bigcap_{i=1}^{\infty} F_i \neq \emptyset$, lo cual establece la contradicción mencionada.

Por lo tanto $\text{lím}_{i\to\infty} P(E_i) = 0$, así que P es σ-subaditiva.

De acuerdo con el teorema de extensión de Carathéodory, existe una única medida de probabilidad P definida sobre $\Im = \sigma(\Im_0)$ tal que $P(\Pi_u^{-1}(B_u)) = P_u(B_u)$ para cualquier $u \in U$ y $B_u \in \mathcal{B}_u$.

Para $t \in \Gamma$, sea $X_t : \Omega \mapsto \mathbb{R}$ definida por $X_t(\omega) = \omega(t)$.

Entonces:

$$[X_t \leq x] = \Pi_{\{t\}}^{-1}((-\infty, x]) \in \Im_0$$

para cualquier $t \in \Gamma$ y $x \in \mathbb{R}$.

Así que X_t es $\Im$-medible para cualquier $t \in \Gamma$.

Además, para cualquier $u = \{t_1, \ldots, t_n\} \in U$ y $(x_1, \ldots, x_n) \in \mathbb{R}^u$, se tiene:

$P[X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n] = P(\Pi_u^{-1}((-\infty, x_1] \times \cdots \times (-\infty, x_n]))$

$= P_u((-\infty, x_1] \times \cdots \times (-\infty, x_n]) = F_u(x_1, \ldots, x_n)$.

Así que F_u es la función de distribución conjunta de $X_{t_1}, \ldots, X_{t_n}$.

■