

Probabilidad I

Santiago de Diego
Braulio Valdivielso
Francisco Luque

Agradecimientos

¡Un saludo a mi gente de Tinder, se os quiere!

Resumen

Contents

Agradecimientos	3
Resumen	5
1 Conjuntos y funciones σ-aditivas	9
1.1 Introducción: Espacio medible	9
1.2 Límites en sucesiones de conjuntos	10
1.2.1 Límites en sucesiones monótonas	10
1.2.2 Límites superiores e inferiores	10
1.2.3 Límite de una sucesión de conjuntos	11
1.3 Funciones sobre conjuntos	12
1.4 Concepto de σ -aditividad	12
1.5 Continuidad en funciones sobre conjuntos	13
1.6 Medidas y probabilidades	15
1.6.1 Teorema de extensión de Carathéodory	16
1.7 Funciones medibles	18
1.8 Convergencia en probabilidad y convergencia casi segura	21
1.8.1 Definición de convergencia en probabilidad	21
1.8.2 Definición de convergencia uniforme en u	22
1.8.3 Propiedades	22
1.8.4 Definición de convergencia casi segura	23
1.9 Esperanza matemática	24
1.9.1 Esperanza matemática	25
1.9.2 Teorema de la convergencia monótona	26
1.9.3 Teorema de Fatou-Lebesgue	27
1.9.4 Teorema de la convergencia dominada	27
1.9.5 Desigualdades de Hölder, Schwarz, Minkowsky, básica, Tchevychev, Markov	27
1.9.6 Convergencia en R -medida	27
1.9.7 Función de distribución, distribución inducida y función generatriz de momentos	27

1.10	Apuntes escritos a recolocar	27
1.10.1	Lema de Fatou	27
1.10.2	Lema de Fatou-Lebesgue	27
1.11	Espacios de probabilidad y variables aleatorias	28
1.11.1	Algunas desigualdades interesantes	28
1.12	Funciones características y funciones de distribución	29
1.12.1	Función característica	29
1.13	Leyes de la probabilidad y tipos de leyes	29

Chapter 1

Conjuntos y funciones σ -aditivas

1.1 Introducción: Espacio medible

Concepto de σ -álgebra

Sea un conjunto Ω , y sea \mathcal{A} una σ -álgebra sobre Ω . Se dice que \mathcal{A} es una σ -álgebra si cumple las siguientes propiedades:

1. $\Omega \in \mathcal{A}$
2. Si $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$
3. Si $A_1, A_2, A_3 \dots \in \mathcal{A} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

Es decir, una σ -álgebra es una clase de conjuntos cerrada para las operaciones complementario y unión numerable. Existen una serie de propiedades inmediatas derivadas de las propiedades que definen a la σ -álgebra:

$$\emptyset \in \mathcal{A}$$

\mathcal{A} es cerrada para la operación intersección numerable.

Espacio medible

Al par (Ω, \mathcal{A}) se le llama espacio medible, y a los conjuntos que pertenecen a \mathcal{A} se les denomina conjuntos medibles.

1.2 Límites en sucesiones de conjuntos

1.2.1 Límites en sucesiones monótonas

Si tomamos la relación de inclusión entre conjuntos (\subseteq) como una relación de orden, podemos hablar sin ningún tipo de problema de sucesiones monótonas. En este sentido, una sucesión creciente de conjuntos sería una sucesión $\{A_n\}_{n \in \mathbb{N}}$ en la que se cumple que $A_i \subseteq A_{i+1}, \forall i \in \mathbb{N}$. De forma análoga se podría ver qué es una sucesión decreciente de conjuntos.

Resulta que existe una forma intuitiva de definir el límite de una sucesión monótona de conjuntos. En particular, el límite de una sucesión creciente de conjuntos $\{A_n\}_{n \in \mathbb{N}}$ se puede definir como

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{k \in \mathbb{N}}^{\infty} A_k$$

Esta definición aprovecha la relación de orden entre los conjuntos de la sucesión para afirmar que *si un elemento está en el último conjunto de la sucesión, entonces está en todos los anteriores, y por tanto en **todos***, por ello se puede utilizar una intersección numerable (que está perfectamente definida) para formalizar el concepto.

Con una intuición análoga se define el límite de una sucesión decreciente de conjuntos. Si $\{A_n\}_{n \in \mathbb{N}}$ es una sucesión decreciente de conjuntos, entonces:

$$\lim_{n \rightarrow \infty} A_n = \bigcap_{k \in \mathbb{N}}^{\infty} A_k$$

1.2.2 Límites superiores e inferiores

No solo se puede hablar de límites en sucesiones monótonas. Para definir los límites en sucesiones arbitrarias de conjuntos tenemos que recurrir a los conceptos de límite inferior y límite superior. La intuición de estos límites superior e inferior pasan por el concepto de las colas de la sucesión.

Dada una sucesión de conjuntos $\{A_n\}_{n \in \mathbb{N}}$ podemos considerar el conjunto

$$B_n = \bigcap_{k=n}^{\infty} A_k$$

y este conjunto contiene aquellos elementos que están en **todos** los A_k para $k \geq n$. Es fácil probar que la sucesión $\{B_n\}_{n \in \mathbb{N}}$ es una sucesión creciente de conjuntos, y por tanto se puede obtener el límite $\lim_{n \rightarrow \infty} B_n$. Informalmente, ese límite es un conjunto que contiene a todos los elementos de A_n que están en todos los conjuntos A_k a partir de cierto $n \in \mathbb{N}$. Definimos el límite inferior de A_n como

$$\liminf A_n = \lim_{n \rightarrow \infty} B_n = \bigcup_{n \in \mathbb{N}} \bigcap_{k=n}^{\infty} A_n$$

Análogamente, podríamos definir la sucesión de conjuntos $\{C_n\}_{n \in \mathbb{N}}$ como

$$C_n = \bigcup_{k=n}^{\infty} A_n$$

Cada C_n contiene todos los elementos que están presentes en algún A_k para $k \geq n$. Es fácil también ver que la sucesión $\{C_n\}_{n \in \mathbb{N}}$ es una sucesión decreciente de conjuntos, y por tanto su límite también está bien definido. Se puede definir entonces el límite superior de $\{A_n\}_{n \in \mathbb{N}}$ como

$$\limsup A_n = \lim_{n \rightarrow \infty} C_n = \bigcap_{n \in \mathbb{N}} \bigcup_{k=n}^{\infty} A_n$$

Informalmente se puede pensar en este límite superior de A_n como el conjunto de los elementos que están en infinitos conjuntos de la sucesión.

A partir de estas definiciones, es fácil comprobar que

$$\liminf A_n \subseteq \limsup A_n$$

1.2.3 Límite de una sucesión de conjuntos

Diremos que una sucesión de conjuntos tiene límite si su límite inferior y superior coinciden, y el límite tendrá como valor efectivamente el de estos límites. Es decir:

$$\lim A_n = \liminf A_n = \limsup A_n$$

en caso de que límite superior e inferior coincidan.

1.3 Funciones sobre conjuntos

Una vez definidos los conceptos sobre conjuntos con los que vamos a trabajar, pasamos a definir las funciones sobre conjuntos. Vamos entonces a definir lo que es una función de conjunto. Sean los espacios medibles (Ω, \mathcal{A}) y (Ω', \mathcal{A}') , definimos la función:

$$\begin{aligned} X : \mathcal{A} &\rightarrow \mathcal{A}' \\ A &\longrightarrow X(A) \end{aligned}$$

A raíz de esta definición podemos definir también lo que se conoce como función inversa. Dada una función X , la función inversa de X , X^{-1} , asigna a cada conjunto $A' \in \mathcal{A}'$ el conjunto $A \in \mathcal{A}$ tal que $X(A) = A'$. La propiedad básica que cumplen las funciones inversas es que preservan las operaciones e inclusiones de conjuntos.

1.4 Concepto de σ -aditividad

Sea un conjunto Ω y una σ -álgebra \mathcal{A} sobre Ω . Definimos la función de conjunto:

$$\varphi : \mathcal{A} \rightarrow \mathbb{R}$$

$$\varphi(A) \text{ es único}$$

Se dice que φ es aditiva si $\varphi\left(\sum_1^n A_k\right) = \sum_1^n \varphi(A_k)$

Se dice que φ es σ -aditiva si $\varphi\left(\sum_1^\infty A_k\right) = \sum_1^\infty \varphi(A_k)$

Función subaditiva

Sea φ definida como antes. Se dice que φ es **subaditiva** si $\varphi(A \cup B) \leq \varphi(A) + \varphi(B)$

Lema 1.4.1 Si $\exists B : \varphi(B) < \infty \Rightarrow \varphi(\emptyset) = 0$

Teorema 1.4.1 Si φ es σ -aditiva y $\sum \varphi(A_i) < \infty \Rightarrow \sum (|\varphi(A_i)|) < \infty$

Demostración 1.4.1 Tomemos las sucesiones:

Si $\varphi(A_n) \geq 0 \Rightarrow A_n^+ = A_n$ y $A_n^- = \emptyset$

Si $\varphi(A_n) < 0 \Rightarrow A_n^+ = \emptyset$ y $A_n^- = A_n$

Entonces $\varphi\left(\sum A_n^+\right) = \sum \varphi(A_n^+)$ y $\varphi\left(\sum A_n^-\right) = \sum \varphi(A_n^-)$, ambas finitas.

Sumando ambas cantidades obtenemos que $\sum (|\varphi(A_i)|) < \infty$ ■

Teorema 1.4.2 *Si φ es σ -aditiva, $\varphi \geq 0$, entonces: φ es no decreciente y subaditiva*

Demostración 1.4.2 *Veamos que es no decreciente ($A \subseteq B \Rightarrow \varphi(A) \leq \varphi(B)$). Podemos escribir $B = (B \cap A) + (B \cap \bar{A}) = A + (B \cap \bar{A})$. Entonces $\varphi(A) \leq \varphi(A) + \varphi(B \cap \bar{A}) = \varphi(B)$.*

1.5 Continuidad en funciones sobre conjuntos

Una vez introducido el concepto de límite para una sucesión de conjuntos, vamos a tratar de definir la continuidad para funciones de conjunto. Tendremos tres tipos de continuidad, cada uno relacionado con un tipo de sucesiones de conjuntos de las que hemos definido anteriormente. En esta sección trabajaremos con una función $\varphi : \Omega \rightarrow \Omega'$

Diremos que φ es continua por abajo si cumple que, dada una sucesión creciente de elementos $A_n \uparrow A$, se tiene que

$$\lim \varphi(A_n) = \varphi(A)$$

Por otra parte, diremos que φ es continua por arriba si cumple que dada una sucesión decreciente de elementos $A_n \downarrow A$, se tiene que

$$\lim \varphi(A_n) = \varphi(A)$$

Por último, diremos que una función es continua si lo es por arriba y por abajo.

Teorema 1.5.1 *Teorema de continuidad para funciones sobre conjuntos*

Sea φ una función σ -aditiva. Entonces, φ es aditiva y continua. Inversamente, si φ es aditiva y, o bien continua por abajo, o finita y continua en \emptyset , entonces φ es σ -aditiva.

Demostración 1.5.1 *Por un lado, sea φ una función σ -aditiva. Entonces es trivialmente aditiva. Ahora, veamos que es continua por abajo y por arriba. Sea $A_n \uparrow A$, entonces:*

$$A = \lim A_n = \bigcup A_n = A_1 + (A_2 - A_1) + (A_3 - A_2) + \dots$$

Unión de conjuntos disjuntos. Por tanto:

$$\varphi(A) = \varphi(\lim A_n) = \lim_{n \rightarrow \infty} \{\varphi(A_1) + \varphi(A_2 - A_1) + \dots + \varphi(A_n - A_{n-1})\} = \lim \varphi(A_n)$$

Veamos la continuidad por arriba. Sea $A_n \downarrow A$, tomamos A_{n_0} tal que $\varphi(A_{n_0})$ es finito. Entonces $A_{n_0} - A_n \uparrow A_{n_0} - A$, y por el apartado anterior tenemos la convergencia desde abajo, por tanto:

$$\varphi(A_{n_0}) - \varphi(A) = \varphi(\lim(A_{n_0} - A_n)) = \lim \varphi(A_{n_0} - A_n) = \varphi(A_{n_0}) - \lim \varphi(A_n)$$

De donde se deduce que $\varphi(A) = \lim \varphi(A_n)$.

Inversamente, sea φ una función aditiva. Si φ es continua por abajo, tenemos

$$\varphi\left(\sum_n A_n\right) = \varphi\left(\lim \sum_{k=1}^n A_k\right) = \lim \varphi\left(\sum_{k=1}^n A_k\right) = \lim \sum_{k=1}^n \varphi(A_k) = \sum_n \varphi(A_n)$$

Y por tanto es σ -aditiva. Si es finita y continua en \emptyset , entonces se obtiene la σ -aditividad de:

$$\varphi\left(\sum_n A_n\right) = \varphi\left(\sum_{k=1}^n A_k\right) + \varphi\left(\sum_{k=n+1}^{\infty} A_k\right) = \sum_{k=1}^n \varphi(A_k) + \varphi\left(\sum_{k=n+1}^{\infty} A_k\right)$$

Y tenemos que

$$\varphi\left(\sum_{k=n+1}^{\infty} A_k\right) \rightarrow \varphi(\emptyset) = 0$$

■

Una vez demostrado este teorema, vamos a ver un teorema que nos relaciona las propiedades del supremo e ínfimo de una función σ -aditiva con los conjuntos sobre los que está dicha función definida:

Teorema 1.5.2 *Teorema del supremo e ínfimo*

Sea φ una función σ -aditiva sobre una σ -álgebra \mathcal{A} . Entonces, existen $C, D \in \mathcal{A}$ tales que $\varphi(C) = \sup \varphi$ y $\varphi(D) = \inf \varphi$

Demostración 1.5.2 *Probaremos la existencia del conjunto C . La del conjunto D es análoga. Si $\varphi(A) = \infty$ para algún $A \in \mathcal{A}$, entonces podemos establecer $A = C$ y la demostración del teorema es trivial. Entonces, supongamos que $\varphi < \infty$ y dado que el valor $-\infty$ está excluido, φ es finita.*

Entonces, existe una sucesión $\{A_n\} \subset \mathcal{A}$ tal que $\varphi(A_n) \rightarrow \sup \varphi$. Sea $A = \cup A_n$ y para cada n , consideramos la partición de A en 2^n conjuntos A_{nm} de la forma $\cap_{k=1}^n A'_k$, donde $A'_k = A_k$ o $A - A_k$. Para $n < n'$, cada conjunto A_{nm} es una suma finita de conjuntos $A_{n'm'}$. Sea ahora B_n la suma de los conjuntos A_{nm} para los cuales φ es no negativa. Si no hay ninguno, entonces $B_n = \emptyset$

1.6 Medidas y probabilidades

Definición de medida: Una función de conjuntos μ_0 es una medida si verifica:

- Es σ -aditiva, es decir, $\mu_0(\cup A_j) = \sum \mu_0(A_j)$
- Es no decreciente $A \subset B \Rightarrow \mu^\circ(A) \leq \mu^\circ(B)$
- $\mu^\circ(\emptyset) = 0$

A la tupla $(\Omega, \mathcal{A}, \mu_{\mathcal{A}})$ se le denomina espacio de medida.

Definición de medida exterior: Una medida exterior es una función de conjuntos positiva y σ -subaditiva, es decir, no se cumple la primera propiedad. La σ -subaditividad implica que $\mu(A \cup B) \leq \mu(A) + \mu(B)$.

Para que una medida exterior fuera una medida tendría que ser σ -aditiva. Es decir, la falla la primera condición. Sí que es positiva ya que $\mu^\circ(\emptyset) = 0$ y es creciente. Esta medida exterior se aplica a cualquier conjunto. Va a haber unos subconjuntos en los que la función se comporte como si fuera aditiva.

Una vez definido el concepto de medida, vamos a dar ahora el de probabilidad. Una probabilidad \mathcal{P} sobre un espacio medible (Ω, \mathcal{A}) es una medida que además cumple que $\mathcal{P}(\Omega) = 1$. Por tanto, tiene las siguientes propiedades:

$$\mathcal{P}(\emptyset) = 0$$

$$\forall A \in \mathcal{A}, 0 \leq \mathcal{P}(A) \leq 1$$

Es una función σ -aditiva

De forma análoga al concepto de espacio de medida, podemos definir ahora el de espacio probabilístico. Un espacio probabilístico es una tupla formada por un conjunto Ω , una σ -álgebra sobre ese conjunto, \mathcal{A} , y una función de probabilidad \mathcal{P} .

Teorema 1.6.1 *Teorema de sucesión*

1. Si $A_n \uparrow A \Rightarrow \mathcal{P}(A_n) \uparrow \mathcal{P}(A)$
2. $\mathcal{P}(\liminf A_n) \leq \liminf \mathcal{P}(A_n)$
3. Si $A_n \downarrow A \Rightarrow \mathcal{P}(A_n) \downarrow \mathcal{P}(A)$
4. $\mathcal{P}(\limsup A_n) \geq \limsup \mathcal{P}(A_n)$
5. Si $A_n \rightarrow A \Rightarrow \mathcal{P}(A_n) \rightarrow \mathcal{P}(A)$

Demostración 1.6.1 *Falta la demostración del teorema de sucesión***1.6.1 Teorema de extensión de Carathéodory**

Una vez vistas las definiciones anteriores de medida y medida exterior, vamos a ver un teorema fundamental, que nos servirá para justificar la forma en la que haremos el cálculo de probabilidades a posteriori. Veamos dos lemas antes que nos serán necesarios para la demostración del teorema principal.

Definición: Un conjunto $A \in S(\Omega)$ es μ° -medible si se cumple que $\mu^\circ(D) = \mu^\circ(AD) + \mu^\circ(A^cD), \forall D \in S(\Omega)$

Definición: Una extensión exterior μ° de una medida μ se define como:

$$\mu^\circ(A) = \inf \sum_j \mu(A_j), \quad A \subset \bigcup_j A_j, \quad \text{recubrimiento}$$

Teorema 1.6.2 *La extensión exterior μ° de una medida μ sobre un álgebra es una extensión de μ a una medida exterior*

Demostración 1.6.2 *Primero, veamos que es una extensión. Esto es trivial, dado que $\mu^\circ(A) = \inf \sum \mu(A) = \mu(A)$. Ahora, veamos que es una medida exterior.*

Veamos que es subaditiva, es decir, $\mu^\circ(\bigcup_j A_j) \leq \sum_j \mu^\circ(A_j)$. Sea entonces, para cada A_j , un recubrimiento $\bigcup_k A_{kj} \supset A_j$. Entonces:

$$\sum_k \mu(A_{kj}) \leq \frac{\varepsilon}{2^j} + \mu^\circ(A_j)$$

$$\mu^\circ(\cup_j A_j) \leq \sum_{j,k} \mu(A_{kj}) \leq \sum_j \left(\mu^\circ(A_j) + \frac{\varepsilon}{2^j} \right) \leq \sum_j (\mu^\circ(A_j)) + \varepsilon$$

Claramente es no decreciente, ya que si $A \subseteq B$, todo recubrimiento de B lo será también de A , y por tanto, $\mu^\circ(A) \leq \mu^\circ(B)$

Por último, $\mu^\circ(\emptyset) = \mu(\emptyset) = 0$

Teorema 1.6.3 Sea μ° una medida exterior sobre la σ -álgebra \mathcal{A} , y sea \mathcal{A}° la clase de conjuntos μ° -medibles:

- \mathcal{A}° es un σ -campo
- μ° en \mathcal{A}° es una medida

Demostración 1.6.3 Primero, veamos que \mathcal{A}° es un álgebra. Sea $A \in \mathcal{A}^\circ$, entonces $A^c \in \mathcal{A}^\circ$ dado que la definición de μ° -medibilidad es simétrica. Sean entonces $A, B \in \mathcal{A}^\circ$. Entonces:

$$\begin{aligned} \mu^\circ(D) &= \mu^\circ(AD) + \mu^\circ(A^cD) = \\ &= \mu^\circ(ABD) + \mu^\circ(AB^cD) + \mu^\circ(A^cBD) + \mu^\circ(A^cB^cD) \leq \\ &\leq \mu^\circ(ABD) + \mu^\circ(AB^cD \cup A^cBD \cup A^cB^cD) = \\ &= \mu^\circ(ABD) + \mu^\circ(AB)^cD \end{aligned}$$

Dado entonces que \mathcal{A}° es cerrada bajo la operación complementario e intersección finita, lo es para la unión finita, así que nos encontramos ante un álgebra. Veamos ahora que μ° es aditiva en \mathcal{A}° . Sean $A, B \in \mathcal{A}^\circ$ disjuntos. Entonces:

$$\mu^\circ(A + B) = \mu^\circ((A + B)A) + \mu^\circ((A + B)A^c) = \mu^\circ(A) + \mu^\circ(B)$$

Y teniendo que $\mu^\circ(A) \geq \mu^\circ(\emptyset) = 0$, μ° sobre \mathcal{A}° es una función aditiva y positiva. Para completar la prueba, veamos que μ° es σ -aditiva. Sean $A_n \in \mathcal{A}^\circ$, $A = \sum A_n$. Tomamos los conjuntos $B_n = \sum_{k=1}^n A_k \in \mathcal{A}^\circ$. Tenemos que:

$$\mu^\circ(D) \geq \mu^\circ(B_n D) + \mu^\circ(B_n^c D) \geq \sum_{k=1}^n \mu^\circ(A_k D) + \mu^\circ(A^c D)$$

Haciendo $n \rightarrow +\infty$

$$\mu^\circ(D) \geq \sum_{n=1}^{+\infty} \mu^\circ(A_n D) + \mu^\circ(A^c D) \geq \mu^\circ(AD) + \mu^\circ(A^c D)$$

Además queda demostrado que $A \in \mathcal{A}^\circ$, así que tenemos que \mathcal{A}° es un σ -campo. Por último, tomando $D = A$, tenemos que $\mu^\circ(A) \geq \sum \mu^\circ(A_n) \Rightarrow \mu^\circ(A) = \sum \mu^\circ(A_n) \Rightarrow \mu^\circ$ es una medida. ■

Teorema 1.6.4 *de extensión de Carathéodory*

Dada una medida μ sobre un álgebra \mathcal{C} , existe una extensión μ° sobre la σ -álgebra minimal sobre \mathcal{C} ($\mathcal{A}(\mathcal{C})$). Además, si μ es finita, la extensión es única.

Demostración 1.6.4 $\forall A \in \mathcal{C}, \forall D \in \mathcal{C}, \forall \varepsilon > 0, \exists \{A_j\} \subset \mathcal{C}$ recubrimiento de D tal que

$$\mu^\circ(D) + \varepsilon \geq \sum \mu(A_j) \geq \sum \mu(AA_j) + \sum \mu(A^c A_j) \geq \mu^\circ(AD) + \mu^\circ(A^c D) \Rightarrow$$

$$\stackrel{\varepsilon \rightarrow 0}{\Rightarrow} \mu^\circ(D) \geq \mu^\circ(AD) + \mu^\circ(A^c D), A \in \mathcal{A}^\circ \Rightarrow \mathcal{C} \subseteq \mathcal{A}^\circ, \mathcal{A}(\mathcal{C}) \subseteq \mathcal{A}^\circ$$

Por los dos teoremas que hemos demostrado antes, la restricción μ° sobre $\mathcal{A}(\mathcal{C})$ es una extensión de μ a una medida sobre $\mathcal{A}(\mathcal{C})$. Ahora, veamos que es única. Para ello, sean μ_1, μ_2 extensiones de μ . Sea entonces el conjunto $\mathcal{M} = \{A : \mu_1(A) = \mu_2(A)\} \Rightarrow \mathcal{C} \subset \mathcal{M}$. Tenemos entonces que:

$$\Omega \in \mathcal{C} \Rightarrow \Omega \in \mathcal{M} \Rightarrow \mu_1(\Omega) = \mu_2(\Omega) = \mu(\Omega) < \infty \Rightarrow \mu_1, \mu_2 \text{ finitos}$$

$$A \in \mathcal{M} \Rightarrow \mu_1(A^c) = \mu(\Omega) - \mu_1(A) = \mu(\Omega) - \mu_2(A) = \mu_2(A^c) \Rightarrow A^c \in \mathcal{M}$$

$$A, B \in \mathcal{M} \Rightarrow \mu_1(A+B) = \mu_1(A) + \mu_1(B) = \mu_2(A) + \mu_2(B) = \mu_2(A+B) \Rightarrow A+B \in \mathcal{M}$$

Entonces, de momento \mathcal{M} es un campo.

Sean ahora $\{A_n\} \subset \mathcal{M}$ monótona, entonces $\mu_1(\lim A_n) = \lim \mu_1(A_n) = \lim \mu_2(A_n) = \mu_2(\lim A_n) \Rightarrow \mathcal{M}$ cerrada frente al límite de sucesiones monótonas $\Rightarrow \mathcal{M}$ σ -álgebra que contiene a \mathcal{C} , por tanto, contiene a la σ -álgebra minimal sobre \mathcal{C}

$$\mathcal{A}(\mathcal{C}) \subseteq \mathcal{M} \Rightarrow \mu_1 = \mu_2 \text{ en } \mathcal{A}(\mathcal{C})$$

1.7 Funciones medibles

Una vez definidos los conceptos sobre espacios de medida y espacios probabilísticos, vamos a aproximarnos al concepto de función medible. Para ello, daremos dos definiciones de función medible, para demostrar más adelante que estas dos definiciones son equivalentes. Empecemos con la definición

constructiva de función medible. Dado que nos interesa que el codominio sea \mathbb{R} , trabajaremos con funciones definidas entre un espacio medible (Ω, \mathcal{A}) y $(\mathbb{R}, \mathcal{B})$, donde \mathcal{B} representa la σ -álgebra de Borel.

Primero, se define la función puntual $X : \Omega \rightarrow \mathbb{R}$, que asigna a cada $\omega \in \Omega$ un número $x = X(\omega)$ único. Llamaremos a esta función variable aleatoria. Utilizaremos de aquí en adelante la siguiente notación:

- $[X]$ = dominio de la variable aleatoria X .
- $[X \leq Y] = \{\omega : X(\omega) \leq Y(\omega)\}$
- $[X = x] = \{\omega : X(\omega) = x\}$

Sea entonces la función $X = \sum_j x_j I_{A_j}$, donde A_j son conjuntos medibles y I_{A_j} denota la función indicadora de dicho conjunto. Estas funciones se llaman funciones elementales, y cuando el número de valores distintos que toma la función X es finito, se conocen como funciones simples. Entonces, la definición constructiva de función medible es la que sigue:

Definición constructiva de función medible: Una función es medible si es límite de una sucesión de funciones simples $\{X_n\}$ convergentes.

Esta definición que hemos dado es constructiva, y por tanto, nos será muy útil para la definición constructiva de las integrales. No obstante, para enunciar y demostrar las propiedades de las funciones medibles, suele ser más útil la definición descriptiva siguiente:

Definición descriptiva de función medible: Sea una función $\varphi : \mathcal{A} \rightarrow \mathcal{B}$. Se dice que φ es medible si $\forall B \in \mathcal{B} \Rightarrow \varphi^{-1}(B) \in \mathcal{A}$. Es decir, una función se dice medible si la imagen inversa de todo conjunto medible es medible.

No obstante, se puede dar, a raíz de esta, otra definición más económica:

Para la definición anterior, es suficiente con exigir la medibilidad de las imágenes inversas de los elementos de una subclase \mathcal{C} para la cual la σ -álgebra minimal sobre \mathcal{C} sea \mathcal{B}

Veamos ahora que las dos definiciones que hemos dado son equivalentes.

Teorema 1.7.1 *Teorema de medibilidad*

Las definiciones constructiva y descriptiva de una función medible son

equivalentes, y la clase de funciones medibles es cerrada bajo las operaciones usuales del análisis.

Demostración 1.7.1 Sean X_n funciones medibles en el sentido descriptivo. Entonces todos los conjuntos de la forma

$$[\inf X_n < x] = \cup [X_n < x], [-X_n < x] = [X_n > -x]$$

son medibles, y por tanto, las funciones

$$\sup X_n = -\inf(-X_n), \liminf(X_n) = \sup(\inf_{k \geq n} X_k)$$

$$\limsup X_n = -\liminf(-X_n)$$

son medibles en el sentido descriptivo. Por un lado, las funciones de este tipo son claramente cerradas bajo las operaciones de supremo, ínfimo, y límites. Además, toda función simple es medible en el sentido descriptivo, ya que todos los conjuntos $[X \leq x] = \sum_{x_j \leq x} A_j$ son medibles. Entonces, el límite de

sucesiones convergentes de funciones simples son medibles, y por tanto las funciones medibles en sentido constructivo lo son en sentido descriptivo.

Veamos la otra implicación, es decir, que las funciones medibles en sentido descriptivo lo son en sentido constructivo. Sea una función X medible en sentido descriptivo. Entonces, las funciones

$$X_n = -nI_{[X < -n]} + \sum_{-n2^n+1}^{n2^n} \frac{k-1}{2^n} I_{\left[\frac{k-1}{2^n} \leq X < \frac{k}{2^n}\right]} + nI_{X \geq n}, n \in \mathbb{N}$$

son simples. Entonces, dado que

$$|X_n(\omega) - X(\omega)| < \frac{1}{2^n} \quad \text{para} \quad |X(\omega)| < n$$

y

$$X_n(\omega) = \pm n \quad \text{para} \quad X(\omega) = \pm\infty$$

se tiene entonces que $X_n \rightarrow X$ y esto, con lo anterior, prueba la equivalencia de las dos definiciones de función medible. ■

Lema 1.7.1 Sea $X : A \longrightarrow B$

X es medible $\iff X^{-1}(S) \in A, S \in B$

Teorema 1.7.2 Sea $X : \Omega \longrightarrow \mathbb{R}$, y $g : \mathbb{R} \longrightarrow \mathbb{R}$ continua. Entonces, $g(X)$ es medible

Demostración 1.7.2 Si X es elemental, tenemos que $g(X) = g\left(\sum_j x_j I_{A_j}\right) =$

$\sum_j g(x_j) I_{A_j}$ función elemental, y por tanto, medible.

Si X es el límite de funciones elementales:

$$g(X) = g(\lim X_n) = \lim g(X_n) = \lim \sum_j g(x_j) I_{A_j}$$

■

Teorema 1.7.3 Sean X función medible y g función boreliana. Entonces, $g(X)$ es medible.

Demostración 1.7.3

$$(gX)^{-1}(B) = X^{-1}(g^{-1}(B)) \in \mathcal{A}$$

■

1.8 Convergencia en probabilidad y convergencia casi segura

Vamos a establecer ahora criterios que nos permitan comparar variables aleatorias. Dado un espacio de probabilidad $(\Omega, \mathcal{A}, \mathcal{P})$, se dice que dos variables aleatorias son equivalentes si:

$$X \mathcal{R} Y \Leftrightarrow \mathcal{P}[X = Y] = 1$$

Análogamente:

$$X(\omega) = Y(\omega) \forall \omega \in \Omega \setminus \Lambda, \mathcal{P}(\Lambda) = 0$$

1.8.1 Definición de convergencia en probabilidad

Si $P[|X_n - X| \geq \varepsilon] \rightarrow 0$, entonces se dice que $X_n \xrightarrow{\mathbb{P}} X$, es decir, X_n converge en probabilidad a X

Lema 1.8.1 $X_n \xrightarrow{\mathcal{P}} X \wedge X_n \xrightarrow{\mathcal{P}} Y \implies X \mathcal{R} Y$

Demostración 1.8.1

$$|X - Y| = |X - X_n - Y + X_n| \leq |X_n - X| + |X_n - Y|$$

$$P[|X - Y| \geq \varepsilon] \leq P[|X_n - X| \geq \frac{\varepsilon}{2}] + P[|X_n - Y| \geq \frac{\varepsilon}{2}] \rightarrow 0$$

Y por tanto, $X \mathcal{R} Y$ ■

1.8.2 Definición de convergencia uniforme en u

Se dice que una variable aleatoria X converge uniformemente en u si se cumple que $\mathcal{P}[|X_{n+u} - X_n| \geq \varepsilon] \rightarrow 0$

Lema 1.8.2 *La convergencia en probabilidad implica la convergencia uniforme*

Demostración 1.8.2

$$|X_{n+u} - X_n| \leq |X_{n+u} - X| + |X_n - X| \Rightarrow$$

$$\Rightarrow \mathcal{P}[|X_{n+u} - X_n| \geq \varepsilon] \leq \mathcal{P}[|X_{n+u} - X| \geq \frac{\varepsilon}{2}] + \mathcal{P}[|X_n - X| \geq \frac{\varepsilon}{2}] \rightarrow 0$$

1.8.3 Propiedades

1. $X_n \xrightarrow{\mathcal{P}} X \wedge Y_n \xrightarrow{\mathcal{P}} Y \implies X_n + Y_n \xrightarrow{\mathcal{P}} X + Y$
2. $X_n \xrightarrow{\mathcal{P}} X \implies KX_n \xrightarrow{\mathcal{P}} KX$
3. $X_n \xrightarrow{\mathcal{P}} K$ (es decir, que degenera), entonces $X_n^2 \xrightarrow{\mathcal{P}} K^2$.
Para demostrarlo basta notar que: $X_n^2 - K^2 = (X_n + K)(X_n - K)$
4. $X_n \xrightarrow{\mathcal{P}} a \wedge Y_n \xrightarrow{\mathcal{P}} b \implies X_n Y_n \xrightarrow{\mathcal{P}} a \cdot b$

Para demostrarlo tenemos que notar que:

$$X_n Y_n = \frac{(X_n + Y_n)^2 - X_n^2 - Y_n^2}{2} \xrightarrow{\mathcal{P}} \frac{(a + b)^2 - a^2 - b^2}{2} = ab$$

5. $X_n \xrightarrow{\mathcal{P}} 1 \implies \frac{1}{X_n} \xrightarrow{\mathcal{P}} 1$

1.8. CONVERGENCIA EN PROBABILIDAD Y CONVERGENCIA CASI SEGURA 23

$$6. X_n \xrightarrow{\mathcal{P}} a \wedge Y_n \xrightarrow{\mathcal{P}} b \implies X_n Y_n^{-1} \xrightarrow{\mathcal{P}} ab^{-1}$$

$$7. X_n \xrightarrow{\mathcal{P}} X \wedge Y_n \xrightarrow{\mathcal{P}} Y \implies X_n Y_n \xrightarrow{\mathcal{P}} XY$$

Para demostrarlo basta notar que: $(X_n - X)(Y_n - Y) \xrightarrow{\mathcal{P}} 0$ y luego $X_n Y_n - XY_n - X_n Y - XY \xrightarrow{\mathcal{P}} 0$ (Por la propiedad siguiente)

$$8. X_n \xrightarrow{\mathcal{P}} X, \text{ entonces } YX_n \xrightarrow{\mathcal{P}} YX$$

Teorema 1.8.1 Si $X_n \xrightarrow{\mathcal{P}} X$ y $g(\cdot)$ es continua, entonces se cumple que:

$$g(X_n) \xrightarrow{\mathcal{P}} g(X)$$

Demostración 1.8.3 $\mathcal{P}[|X| \geq k] < \frac{\varepsilon}{2} \forall \varepsilon > 0, \exists k$. Consideramos el compacto $[-k, k]$, $A = [|X| < k]$, $B = [|X_n - X| < \delta]$ y $C = [|g(X_n) - g(X)| < \varepsilon]$. Tenemos que ver si $\mathcal{P}(B) \rightarrow 1 \Rightarrow \mathcal{P}(C) \rightarrow 1$:

$$\begin{aligned} AB \subset C &\Rightarrow C^c \subset A^c \cup B^c \Rightarrow \mathcal{P}(C^c) \leq \mathcal{P}(A^c) + \mathcal{P}(B^c) \leq \\ &\leq \mathcal{P}[|X| \geq k] + \mathcal{P}[|X_n - X| \geq \delta] \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

1.8.4 Definición de convergencia casi segura

Una sucesión de variables aleatorias, X_n , converge con probabilidad 1, o de forma casi segura, a una variable aleatoria X (que puede degenerar en una constante K) cuando se cumple que:

$$P(\lim_{n \rightarrow \infty} X_n = X) = 1$$

De esta forma interpretamos que $X_n \xrightarrow{c.s.} X$ cuando la probabilidad de que en el límite la sucesión de variables aleatorias y aquella a la que converge sean iguales es uno

Teorema 1.8.2

$$\begin{aligned} X_n \xrightarrow{c.s.} X &\implies X_n \xrightarrow{\mathcal{P}} X \\ X_n \xrightarrow{\mathcal{P}} X &\implies \exists X_{n_k} : X_{n_k} \xrightarrow{c.s.}_{k \rightarrow \infty} X \end{aligned}$$

Demostración 1.8.4

$$0 = \lim_{n \rightarrow \infty} P \bigcup_{m \geq n} [|X_m - X| \geq \epsilon] \geq \lim_{n \rightarrow \infty} P[|X_n - X| \geq \epsilon]$$

■

Lema 1.8.3 Sea \mathbb{X}_n finita. Entonces se tiene que:

$$\mathbb{X}_n \xrightarrow{c.s.} \mathbb{X} \iff \mathbb{X}_n \text{ es de cauchy}$$

Demostración 1.8.5

\implies

$\exists n_0 : \forall n \geq n_0$ se tiene que:

$$|\mathbb{X}_n(\omega) - \mathbb{X}(\omega)| < \frac{\epsilon}{2} \quad \omega \in A \quad P(A) = 1$$

$$|\mathbb{X}_n(\omega) - \mathbb{X}_m(\omega)| \leq |\mathbb{X}_n(\omega) - \mathbb{X}(\omega)| + |\mathbb{X}_m(\omega) - \mathbb{X}(\omega)| < \epsilon$$

\impliedby

$\mathbb{X}_n(\omega)$ de Cauchy

$$\lim \mathbb{X}_n(\omega) = \mathbb{X}(\omega) \quad \omega \in A$$

■

1.9 Esperanza matemática

Para el cálculo de probabilidades, nos será muy útil el concepto de integral sobre funciones de conjunto. Este concepto nos servirá para calcular lo que se conoce como la esperanza matemática de una variable aleatoria X . En este apartado trabajaremos sobre el espacio de probabilidad $(\Omega, \mathcal{A}, \mathcal{P})$. Comenzaremos definiendo la integral para las funciones simples, para dar luego una definición de integral para funciones no negativas y por último para funciones cualesquiera.

Sea entonces $\{A_k\} \in \mathcal{A}$, tal que $\sum_k A_k = \Omega$, partición medible del espacio.

Sea entonces la función simple $X = \sum_{j=1}^m x_j I_{A_j}$, $x_j \geq 0$. La integral de la función X se define como:

$$\int_{\Omega} X d\mathcal{P} = \sum_{j=1}^m x_j \mathcal{P}_{A_j}$$

Ahora, para cualquier función no negativa X , se define la integral de la función como:

$$\int_{\Omega} X d\mathcal{P} = \lim \int_{\Omega} X_n d\mathcal{P}$$

Donde $\{X_n\} \rightarrow X$. Finalmente, la integral en Ω de una función medible X se define como:

$$\int_{\Omega} X d\mathcal{P} = \int_{\Omega} X^+ d\mathcal{P} - \int_{\Omega} X^- d\mathcal{P}$$

donde $X^+ = XI_{[X \geq 0]}$ y $X^- = -XI_{[X < 0]}$. Si $\int_{\Omega} X d\mathcal{P}$ es finita, es decir, si los dos términos de la diferencia anterior son finitos, entonces se dice que X es integrable en Ω . Ahora, una vez definida la integral, vamos a ver algunas de sus propiedades. Tenemos primero una serie de propiedades relacionadas con la aditividad de la integral. Sean X, Y dos funciones medibles, entonces (no se demostrarán las propiedades triviales):

$$\int (X + Y) d\mathcal{P} = \int X d\mathcal{P} + \int Y d\mathcal{P}$$

Demostración 1.9.1 Sean $X = \sum x_j I_{A_j}$ y $Y = \sum y_k I_{B_k}$. Entonces

$$X + Y = \sum_k \sum_j x_j I_{A_j B_k} + \sum_k \sum_j y_k I_{A_j B_k} = \sum_k \sum_j (x_j + y_k) I_{A_j B_k}$$

Si calculamos ahora las integrales:

$$\int (X+Y) d\mathcal{P} = \sum_k \sum_j (x_j + y_k) \mathcal{P}(A_j B_k) = \sum_j x_j \mathcal{P}(A_j) + \sum_k y_k \mathcal{P}(B_k) = \int X d\mathcal{P} + \int Y d\mathcal{P} \blacksquare$$

$$\int_{A+B} X d\mathcal{P} = \int_A X d\mathcal{P} + \int_B X d\mathcal{P}$$

$$\int cX d\mathcal{P} = c \int X d\mathcal{P}$$

Veamos ahora algunas propiedades relacionadas con el orden:

$$X \geq 0 \rightarrow \int X d\mathcal{P} \geq \int 0 = 0$$

$$X \geq Y \rightarrow \int X d\mathcal{P} \geq \int Y d\mathcal{P}$$

$$X \stackrel{c.s.}{=} Y \rightarrow \int X d\mathcal{P} = \int Y d\mathcal{P}$$

■

1.9.1 Esperanza matemática

La esperanza matemática se define de forma distinta si estamos trabajando con una v.a. discreta o continua. En el caso de una v.a. discreta, sea $p(x_i)$

su función de probabilidad. Definimos la esperanza matemática como:

$$E(\mathbb{X}) = x_1 p(\mathbb{X} = x_1) + \dots + x_n p(\mathbb{X} = x_n) = \sum_{i=1}^n x_i p(x_i)$$

En el caso de una v.a. continua, sea $f(x)$ la función de densidad. Definimos la esperanza matemática como:

$$E(\mathbb{X}) = \int_{-\infty}^{\infty} x f(x) dx$$

1.9.2 Teorema de la convergencia monótona

Una vez vista la definición de integral y esperanza matemática y algunas de sus propiedades, vamos a enunciar y demostrar un teorema de convergencia que nos será de mucha utilidad para el estudio de variables aleatorias. Veamos su enunciado y demostración:

Teorema 1.9.1 *Teorema de la convergencia monótona para funciones medibles no negativas*

Sea $\{X_n\} \geq 0$ tal que $X_n \uparrow X$. Entonces, se tiene que $\int X_n \uparrow \int X$

Demostración 1.9.2 *Tomamos las sucesiones $X_{kn} \uparrow X_k$. La sucesión $Y_n = \max_{k \geq n} X_{kn}$ es una sucesión de funciones simples no negativas y no decreciente, y además*

$$X_{kn} \leq Y_n \leq X_n \rightarrow \int X_{kn} \leq \int Y_n \leq \int X_n$$

Ahora, cuando $n \rightarrow \infty$, tenemos que

$$X_k \leq \lim Y_n \leq X \rightarrow \int X_k \leq \int \lim Y_n \leq \int X$$

Por último, cuando $K \rightarrow \infty$, obtenemos

$$X \leq \lim Y_n \leq X \rightarrow \lim \int X_n \leq \int \lim Y_n \leq \lim \int X_n$$

De donde extraemos que $\lim Y_n = X$ y que $\int X = \lim \int X_n$ ■

1.9.3 Teorema de Fatou-Lebesgue**1.9.4 Teorema de la convergencia dominada**

Sea \mathbb{X}_n una sucesión de variables aleatorias t.q. $\mathbb{X}_n \rightarrow \mathbb{X}$. Si $\exists Y$ t.q. $|\mathbb{X}_n| \leq Y$. Entonces se tiene que \mathbb{X} es integrable y:

$$\lim \int \mathbb{X}_n dP = \int \mathbb{X} dP$$

1.9.5 Desigualdades de Hölder, Schwarz, Minkowsky, básica, Tchevychev, Markov**1.9.6 Convergencia en R-medida****1.9.7 Función de distribución, distribución inducida y función generatriz de momentos****1.10 Apuntes escritos a recolocar****1.10.1 Lema de Fatou**

Lema 1.10.1 Sean Y, Z dos funciones integrables (pueden no mantener su signo), entonces:

- Si $Y_n \leq X_n \implies \int \liminf X_n \leq \liminf \int X_n$
- Si $X_n \leq Z_n \implies \limsup \int X_n \leq \int \limsup X_n$

1.10.2 Lema de Fatou-Lebesgue

Lema 1.10.2 Sean Y, Z dos funciones integrables (pueden no mantener su signo) y X_n una sucesión de variables aleatorias, entonces:

- Si $X_n \leq Y \forall n \implies E(\limsup X_n) \geq \limsup E(X_n)$
- Si $X_n \geq Z \forall n \implies$ si la sucesión es convergente y acotada se cumple que: $E(\lim X_n) \geq \lim E(X_n)$

1.11 Espacios de probabilidad y variables aleatorias

Lema 1.11.1 *de Borel-Cantelli*

Sea (Ω, \mathcal{A}, P) , $A_n \in \mathcal{A}$. Entonces, si $\sum P A_n < \infty \implies P(\limsup A_n) = 0$

Demostración 1.11.1

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \bigcap_{n=1}^{\infty} B_n$$

$$P(\limsup A_n) = P \lim B_n = \lim P B_n \leq \lim_n P \bigcup_{k=n}^{\infty} A_k \leq \lim_n \sum_{k=n}^{\infty} P A_k = 0$$

El recíproco no es cierto ■

Lema 1.11.2 Sea $Y \leq X_n \leq Z$ con Y, Z integrables y tal que $X_n \xrightarrow{c.s.} X$. Entonces se tiene que $\int X_n \rightarrow \int X$

1.11.1 Algunas desigualdades interesantes

Lema 1.11.3 $|a+b|^r \leq C_r |a|^r + C_r |b|^r$ con $C_r = 1$ si $r = 1$ y $C_r = 2^r$ si $r > 1$

Entonces se tiene $E|X + Y|^r \leq C_r E|X|^r + C_r E|Y|^r$

Desigualdad de Hölder-Schwartz

$$E|XY| \leq (E|X|^r)^{\frac{1}{r}} (E|Y|^s)^{\frac{1}{s}} \text{ con } r > s; \frac{1}{r} + \frac{1}{s} = 1$$

Desigualdad de Markov

Sea X una v.a. y sea g una función borel. Entonces se tiene que si g es par y no decreciente en $[0, \infty) \forall a > 0$, entonces:

$$\frac{Eg(X) - g(a)}{\sup g(X)} \leq P[|X| \geq a] \leq \frac{Eg(X)}{g(a)}$$

1.12 Funciones características y funciones de distribución

1.12.1 Función característica

Función puntual que se define sobre la recta real, no negativa, no decreciente, continua por la izquierda y finita. Para nosotros será la función que está entre 0 y 1.

Verifica estas tres propiedades:

- $0 \leq F[a, b] \leq \infty$
- $F[a, b] \rightarrow 0, a \rightarrow b$
- Si $a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots \leq a_n \leq b_n$, entonces:

$$\sum_{k=1}^n F[a_k, b_k] - \sum_{k=1}^{n-1} F[b_k, a_{k+1}] = F[a_1, b_n]$$

Las funciones de distribución representan medidas

Teorema 1.12.1 *La relación $\mu[a, b] = F[a, b]$ establece una correspondencia uno a uno (da igual pasar de una a otra)*

$$\int_{\mathbb{R}} g(X) dP_X(x) = \int_{\mathbb{R}} g P g = \int_{\Omega} g(X(\omega)) dP(\omega)$$

1.13 Leyes de la probabilidad y tipos de leyes