

# Probabilidades

Alexander A. Ramírez M. ([alexanderramirez.me](http://alexanderramirez.me))

27/06/2016

## Contents

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Presentación . . . . .	1
1.2	Definición frecuentista . . . . .	2
1.2.1	¿Cómo medir? . . . . .	3
1.3	Definición Axiomática . . . . .	3
1.3.1	Espacio medible (measurable) . . . . .	4
1.3.2	Medida de Probabilidad . . . . .	4
1.3.3	Espacio de probabilidad . . . . .	4
1.3.3.1	Ejemplos e interpretación . . . . .	4
1.3.4	Propiedades . . . . .	5
1.3.5	Conjuntos Independientes . . . . .	7
1.3.6	Probabilidad total . . . . .	7
1.3.7	Probabilidad condicional . . . . .	8
1.3.8	.. otra propiedad? . . . . .	8
1.4	Variable Aleatoria . . . . .	9
1.5	Función de Distribución . . . . .	9
<b>2</b>	<b>Repaso de Teoría de la medida</b>	<b>9</b>
2.1	Límite de conjuntos . . . . .	9
2.2	Lema de Fatou . . . . .	10
2.3	Medida de Lebesgue . . . . .	10
2.4	Desigualdades de Cola . . . . .	10
2.5	$\sigma$ -álgebra generada por una clase . . . . .	11
2.6	Integral Lebesgue-Stieltjes . . . . .	11
<b>3</b>	<b>Bibliografía</b>	<b>11</b>

## 1 Introducción

### 1.1 Presentación

Inspirado en las clases con el Prof: Ricardo Rios. Gracias Profesor.

La disyuntiva fundamental es la exactitud vs. la precisión. ¿De qué cosas podemos hablar con exactitud o con qué nivel de precisión podemos hablar?. Es un compromiso entre lo completamente determinado y lo no determinado sino aproximado con un nivel de precisión.

Lanzar un objeto como una taza de un octavo piso en contraposición a lanzar una hoja de papel de la misma altura. La pregunta es ¿en cuanto tiempo llega el objeto al suelo?.

En general podemos iniciar con relaciones determinísticas como:

$$Y = f(x)$$

Donde a cada valor de  $x$  le corresponde un único valor de la función  $f(x)$ .

En otros casos hay un componente “aleatorio” sobre el cual podemos hablar de forma aproximada. En estos casos nos interesa comprometer un nivel de precisión. Es decir, nos interesa acotar o controlar ese componente. La relación evoluciona a la forma:

$$Y = f(x) + \epsilon$$

Los modelos de probabilidades se basan en ese elemento y en el estudio de  $\epsilon$ , sus propiedades y finalmente qué podemos decir de  $Y$ , nuestra variable dependiente basados en dichas propiedades.

De forma intuitiva podemos definir un Modelo de Probabilidad como las propiedades que podemos deducir e inferir basados en la forma en que se comporta  $f(x)$  y  $\epsilon$ . Por otra parte hay muchos problemas en donde tampoco se conoce  $f$ . Es decir, no se conoce el proceso generador de los datos y sólo podemos hacer aproximaciones de  $f$  basados en su forma y en propiedades relacionadas con la forma de  $f$ .

En ese punto llegamos a las distribuciones de probabilidad que cumplen con las características propias de las funciones de probabilidad o funciones de medida. Si pensamos que  $f$  se aproxima a una distribución de probabilidad conocida entonces podemos aplicar una serie de herramientas propias de dichas distribuciones como estimar sus parámetros (si existen) y a partir de dichos estimadores conocer intervalos de confianza y aplicar pruebas de hipótesis.

Esos pasos que acabo de describir son el enlace entre la Probabilidad y la Estadística. Ahí es donde se une la Probabilidad con la Estadística. Basados en las propiedades de la probabilidad podemos hacer estadística, que consiste en inferir a partir de los datos provistos; con la dificultad de que los datos tienen errores, en su recolección, interpretación, etc.

Nos interesa saber cuando  $|Y - Y_0| > \delta$  o  $|\epsilon| > \delta$ , es decir cuando el error es mayor que un nivel  $\delta$ , nos interesa acotar el error y conocer con qué frecuencia ocurre eso. Nos interesa controlar  $\epsilon$ , de esta forma logramos algún nivel de precisión.

Después de entender hasta donde puede llegar  $\epsilon$  entonces nos interesa que ocurra pocas veces. En términos relativos que sea poco frecuente y a la postre que su probabilidad sea baja. Es decir, que si repetimos el experimento muchas veces la frecuencia en que el error  $\epsilon$  supera la cota ocurre pocas veces. Entonces nuestro modelo queda:

$$P(|Y - Y_0| > \delta) < \eta$$

Alternativamente, también lo podemos formular de forma análoga. Si el error es menor a una cota superior eso ocurre con una probabilidad alta. De hecho lo veremos más adelante, los intervalos de confianza se definen de esa forma.

Entonces tenemos que si  $|\epsilon| < \delta$ . Es decir, el error está en un intervalo  $(\epsilon - \delta, \epsilon + \delta)$ . Esto nos interesa que ocurra con la mayor probabilidad posible. Pero podemos colocar un umbral como sigue:  $P(|\epsilon| < \delta) > \eta$ .

Finalmente lo que se desea es controlar el error y se utilizan las herramientas de la probabilidad para este propósito.

## 1.2 Definición frecuentista

La forma de presentar los datos por excelencia es un histograma, sin embargo esa representación no nos coloca en una misma escala.

Las probabilidades nos ofrece un marco de comparación de magnitudes y nos coloca en una misma escala entre 0 y 1.

La probabilidad de un evento se puede escribir mediante la función indicatriz 1. Para  $A \subseteq \Omega$  y  $x_1, \dots, x_n$  observaciones de un experimento, se puede definir la probabilidad como sigue:

$$P(A)_{(x_1, \dots, x_n)} = \frac{1}{n} \sum_{i=1}^n 1_A(x_i)$$

Donde  $1_A(x)$  se define:

$$1_A(x) = \begin{cases} 0 & x \in A \\ 1 & x \notin A \end{cases}$$

El reto es que este tipo de probabilidad depende profundamente de los datos y es conocido que la calidad de los datos es baja.

En general estamos tratando con eventos con elementos de azar e incertidumbre. Se trata de encontrar lo común en lo no conocido o no controlado. Es encontrar consistencia entre los resultados y estudiar los compartamientos similares.

Si tratamos, por ejemplo, con la distancia, tal que:

$$\begin{aligned} d(P(A)_{(x_1, \dots, x_n)}; P(A)) &< \delta \\ P(d(P(A)_{(x_1, \dots, x_n)}; P(A)) < \delta) &> 1 - \epsilon \\ P(d(P(A)_{(x_1, \dots, x_n)}; P(A)) > \delta) &< \epsilon \end{aligned}$$

Al final se trata de medir, se trata de conseguir el tamaño relativo de un evento, el tamaño relativo de un conjunto.

Probabilidad: Ciencia del diablo para justificar el pecado.

### 1.2.1 ¿Cómo medir?

A partir de conjuntos elementales ¿Cómo hablar del tamaño de algo?.

Conjuntos elementales  $\{\mathcal{C}\}$

Sea  $\ell : \{\mathcal{C}\} \rightarrow \mathcal{R}^+ \cup \{\emptyset\}$ .

$\ell(\mathcal{C})$  está bien definida,  $(\mathcal{C}_\infty \cap \mathcal{C}_\epsilon) = \emptyset$  entonces:

$$\ell(\mathcal{C}_\infty \cup \mathcal{C}_\epsilon) = \ell(\mathcal{C}_\infty) + \ell(\mathcal{C}_\epsilon)$$

Si  $\{\mathcal{C}_i\}$ ,  $\forall i, i \geq n$  y  $(\mathcal{C}_i \cap \mathcal{C}_j) = \emptyset$  para  $i \neq j$  y sea  $\mathcal{C} = \bigcup_{j=1}^n \mathcal{C}_j$  tenemos:

$$\ell(\mathcal{C}) = \sum_{i=1}^n \ell(\mathcal{C}_i)$$

Buscar Axioma de elección

## 1.3 Definición Axiomática

$(\Omega, \mathcal{F}, P)$

La probabilidad parte de una definición axiomática que tiene base en la teoría de la medida. Los tres elementos principales son  $\Omega$  que es el **espacio muestral**,  $\mathcal{F}$  que es una  $\sigma$ -álgebra sobre  $\Omega$  y  $P$  que es la medida de probabilidad con las propiedades de la medida de Lebesgue.

### 1.3.1 Espacio medible (measurable)

$(\Omega, \mathcal{F})$

Sea  $\Omega$  el espacio de los eventos puntuales (individuales) o **espacio muestral**.  $\Omega$  es el conjunto de todos los posibles resultados de un experimento. Por otra parte sea  $\mathcal{F}$  tal que  $\mathcal{F} \equiv \sigma$ -álgebra y  $\mathcal{F} \subseteq \wp(\Omega)$  donde  $\wp(\Omega)$  es el conjunto partes de  $\Omega$ , es decir,  $\mathcal{F}$  cumple con las propiedades siguientes:

1.  $\Omega \in \mathcal{F}$
2.  $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$
3.  $\{A_n\}_{n=1}^{\infty} \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

A la  $\sigma$ -álgebra en  $\Omega$ , es decir el par  $(\Omega, \mathcal{F})$  se le denomina espacio de medida (measurable)<sup>1</sup>.

$\sigma$  viene de numerable, profesor RR dixit

### 1.3.2 Medida de Probabilidad

$P$

Por otra parte definamos una **medida de probabilidad** como una función que va de  $\mathcal{F}$  a los números reales en el intervalo  $[0, 1]$ ,  $P : \mathcal{F} \Rightarrow [0, 1]$ .  $P$  es una función que asigna una probabilidad a cada evento y cumple las propiedades siguientes:

1.  $P(\Omega) = 1$
2.  $\{A_n\}_{n=1}^{\infty} \in \mathcal{F} \wedge A_i \cap A_j = \emptyset$  para  $i \neq j$  entonces,  $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$

$P$  es una medida de probabilidad en el espacio de medida  $(\Omega, \mathcal{F})$ .

En el Durrett se define: “Una medida como una función no negativa ‘countably additive set function’,  $\mu : \mathcal{F} \Rightarrow \mathcal{R}$  con:

- (i)  $\mu(A) \geq \mu(\emptyset) = 0, \forall A \in \mathcal{F}$ , y
- (ii) Si  $A_i \in \mathcal{F}$  es una secuencia numerable de conjuntos disjuntos, es decir,  $A_i \cap A_j = \emptyset, \forall i, j, i \neq j$  entonces:

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i)$$

Si  $\mu(\Omega) = 1$ , entonces  $\mu$  es una medida de probabilidad.

Esta definición nos proporciona el salto entre medir la unión de conjuntos disjuntos como la suma de sus medidas.

### 1.3.3 Espacio de probabilidad

$(\Omega, \mathcal{F}, P)$

Un espacio de probabilidad está conformado por un espacio medible y una medida de probabilidad. Estos conforman la terna  $(\Omega, \mathcal{F}, P)$ .

#### 1.3.3.1 Ejemplos e interpretación

Las operaciones sobre conjuntos tienen una interpretación ligeramente distinta en esta notación. Los eventos puntuales pertenecen a un evento. Es decir, un evento puntual es un elemento de un conjunto que es un evento.

---

<sup>1</sup>La palabra medible no existe.

- Sea  $\Omega$  nuestro **espacio muestral**, definido como el conjunto de posibles resultados de un experimento como el lanzamiento de un dado  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .
- Un evento  $A \in \mathcal{F}$ , puede ser, que el resultado de un experimento salga un número par, entonces  $A = \{2, 4, 6\}$ , o por otra parte un número impar entonces,  $A = \{1, 3, 5\}$ .  $A \subset B$  implica que la ocurrencia de  $A$  implica la ocurrencia de  $B$ .  $A \cap B$  significa que ambos eventos  $A$  y  $B$  ocurren.  $A \cup B$  significa que alguno de los eventos ocurre, ya sea  $A$  o  $B$ .  $A^C$  es el evento contrario a  $A$ . Es decir, si  $A$  ocurre  $A^C$  no ocurre. Por otra parte, si  $A^C$  ocurre entonces  $A$  no ocurre.
- Un evento puntual es un resultado particular de un experimento, normalmente se denota como  $w \in \Omega$ . En el caso del lanzamiento de un dado, que salga el número 4, entonces  $w = 4$ . Un evento sería el conjunto  $A = \{4\}$ .  $\omega \in A$  significa que  $A$  ocurre cuando  $\omega$  ocurre.  $\omega \notin A$  significa que  $A$  no ocurre cuando  $\omega$  ocurre.
- Existe un evento particular que es el evento nulo conformado por el conjunto vacío  $\emptyset$ .  $A \cap B = \emptyset$  significa que  $A$  y  $B$  son **disjuntos**, y no pueden ocurrir simultáneamente.

### 1.3.4 Propiedades

Vamos a desarrollar algunas propiedades que se derivan de la definición previa. Todos los conjuntos nombrados en las propiedades siguientes pertenecen a  $\mathcal{F}$ .

$$(1) P(\emptyset) = 0$$

Por la segunda propiedad de  $\mathcal{F}$  si  $\Omega \in \mathcal{F} \Rightarrow \Omega^C = \emptyset \in \mathcal{F}$ . Definamos  $A_n : A_n = \emptyset, \forall n$ . Por la segunda propiedad de  $P$ ,  $\emptyset = \{A_n\}_{n=1}^{\infty} \in \mathcal{F} \wedge A_i \cap A_j = \emptyset$  para  $i \neq j$  entonces,  $P(\emptyset) = P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(\emptyset)$ .

Supongamos que  $\sum_{n=1}^{\infty} P(\emptyset) > 0$  es decir,  $P(\emptyset) \neq 0$  entonces  $\sum_{n=1}^{\infty} P(A_n) > 0$  y  $P(\bigcup_{n=1}^{\infty} A_n) > 0$  lo cual implica que  $\bigcup_{n=1}^{\infty} A_n \neq \emptyset$ , es decir, debe existir  $i$  tal que  $A_i \neq \emptyset$ , lo cual contradice la definición inicial de  $A_n : A_n = \emptyset, \forall n$ . Por lo tanto, por reducción al absurdo,  $P(\emptyset) = 0$ .

Que la probabilidad de un conjunto  $A$  sea 0,  $P(A) = 0$ , no quiere decir que  $A = \emptyset$ . De forma análoga que la probabilidad de un evento  $B$  sea 1,  $P(B) = 1$ , no quiere decir que  $B = \Omega$

$$(2) P(A \cup B) = P(A) + P(B)$$

Sea  $A$  y  $B$  conjuntos tal que  $A, B \in \mathcal{F}$  y  $A$  y  $B$  son disjuntos,  $A \cap B = \emptyset$ . Definamos la serie  $\{A_n\} \forall n, n \geq 1$  tal que,  $A_1 = A; A_2 = B; A_{n \geq 3} = \emptyset$ . Tenemos que por (2) propiedad de  $P$ ,  $\{A_n\}_{n=1}^{\infty} \in \mathcal{F} \wedge A_i \cap A_j = \emptyset$  para  $i \neq j$  entonces,  $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ . Sabemos por definición que  $\bigcup_{n=1}^{\infty} A_n = A \cup B$ . Entonces  $P(A \cup B) = P(A) + P(B)$ .

Otra manera con la propiedad (1) me permite calcular de la forma siguiente:  $P(\bigcup_{n=1}^{\infty} A_n) = P(A \cup B \cup \emptyset \cup \dots \cup \emptyset)$ , por definición de la unión  $P(A \cup B \cup \emptyset \cup \dots \cup \emptyset) = P(A \cup B)$  y también  $P(A \cup B \cup \emptyset \dots \emptyset) = P(A) + P(B) + P(\emptyset) + \dots + P(\emptyset)$ . Por (1),  $P(A) + P(B) + P(\emptyset) + \dots + P(\emptyset) = P(A) + P(B) + 0 + \dots + 0 = P(A) + P(B)$  Entonces nos queda  $P(A \cup B) = P(A) + P(B)$

$$(3) P(A) = 1 - P(A^C)$$

Sea  $A$  tal que  $A \cup A^C = \Omega, \forall A \subseteq \mathcal{F}$ . Sabemos que un conjunto intersectado con su complemento resulta el vacío, esto es,  $A \cap A^C = \emptyset$  entonces  $P(\Omega) = P(A \cup A^C)$  y por (2)  $P(A \cup A^C) = P(A) + P(A^C)$ . Por definición de  $P$ ,  $P(\Omega) = 1$  entonces nos queda que  $P(A) + P(A^C) = 1$ . Despejando nos queda que  $P(A) = 1 - P(A^C)$  y también  $P(A^C) = 1 - P(A)$ .

Queda claro que  $\forall A \in \mathcal{F}, P(A) = 1 - P(A^C) \leq 1$ , además es positiva, lo que nos queda que  $0 \leq P(A) \leq 1$ .

$$(4) P(A \cup B) = 1 - P(A^C \cap B^C)$$

Por la propiedad anterior sabemos que  $P(C) = 1 - P(C^C)$ , sea  $C = A \cup B$  entonces,  $P(C) = P(A \cup B) = 1 - P((A \cup B)^C)$ , por propiedades de conjuntos sabemos que  $(A \cup B)^C = (A^C \cap B^C)$ , entonces  $P(A \cup B) = 1 - P(A^C \cap B^C)$ .

$$(5) A \subseteq B \Rightarrow P(A) \leq P(B)$$

Intuitivamente si escribimos a  $B$  como la unión de dos conjuntos disjuntos y uno de los elementos de la unión es  $A$ . Entonces  $B$  es el resultado de la suma entre dos cantidades positivas que incluyen a  $A$  por lo que  $P(A) \leq P(B)$ .

Se  $A \subseteq B$  es verificable que  $A \cup B = B$ . También sabemos que  $(A \cup B) \cap \Omega = A \cup B$ . También podemos escribir  $\Omega$  en función de la unión de un conjunto y su complemento, es decir,  $A \cup A^C = \Omega$ , entonces podemos decir que  $(A \cup B) \cap (A \cup A^C) = B$ , por ley de conjuntos queda  $A \cup (B \cap A^C) = B$  y se puede verificar que  $A \cap (B \cap A^C) = \emptyset$  y por propiedades de  $P$ , la probabilidad  $P(A \cup (A^C \cap B)) = P(A) + P(A^C \cap B) = P(B)$ . Como  $P(B)$  se escribe como la suma de  $P(A)$  más otra cantidad positiva entonces  $P(A) \leq P(B)$ .

Otra forma de escribir a  $B$  es la siguiente  $B = A \cup (B - A)$ . Es verificable que  $B - A = B \cap A^C$ .

$$(6) \quad A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$$

Sea  $A, B \in \mathcal{F}$  tenemos que por las propiedades de  $\mathcal{F}$  entonces,  $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$  y  $B \in \mathcal{F} \Rightarrow B^C \in \mathcal{F}$ . Por otra parte si  $A^C, B^C \in \mathcal{F} \Rightarrow A^C \cup B^C \in \mathcal{F}$ . Finalmente si  $A^C \cup B^C \in \mathcal{F}$  entonces  $(A^C \cup B^C)^C \in \mathcal{F}$ . Por la ley de De Morgan  $(A^C \cup B^C)^C = A \cap B$  entonces  $A \cap B \in \mathcal{F}$ .

$$(7) \quad P(A \cap B^c) = P(A) - P(A \cap B)$$

Sea  $A$ , lo podemos expresar de la forma siguiente,  $A \cap \Omega = A$ , pero  $\Omega$  la podemos escribir como la unión disjunta de un conjunto  $B$  y su complemento  $B^C$ ,  $\Omega = B \cup B^C$ , entonces,  $A \cap (B \cup B^C) = A$ , por propiedad de la intersección  $(A \cap B) \cup (A \cap B^C) = A$ , sabemos propiedades de  $P$  que  $P((A \cap B) \cup (A \cap B^C)) = P(A \cap B) + P(A \cap B^C) = P(A)$ , si despejamos nos queda que  $P(A \cap B^c) = P(A) - P(A \cap B)$ .

En general si contamos con una secuencia de conjuntos disjuntos tal que  $\{A_n\}, \forall n : n \geq 1$  y  $\bigcup_{n=1}^{\infty} A_n = \Omega$  y  $A_i \cap A_j = \emptyset, \forall i, j : i \neq j$  entonces para cualquier conjunto  $A$ ,  $P(A) = \bigcup_{n=1}^{\infty} P(A \cap A_n) = \sum_{n=1}^{\infty} P(A \cap A_n)$ .

$$(8) \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Esta propiedad es mejor conocida como el principio de inclusión y exclusión.

Por la propiedad (7) podemos escribir a  $A$  y  $B$  como la unión de intersecciones. Sea  $A = (A \cap B) \cup (A \cap B^C)$  y  $B = (B \cap A) \cup (B \cap A^C)$ ,  $P(A) = P(A \cap B) + P(A \cap B^C)$  y  $P(B) = P(B \cap A) + P(B \cap A^C)$ . Por otra parte  $A \cup B = (A \cap B) \cup (A \cap B^C) \cup (B \cap A) \cup (B \cap A^C)$ , por conmutatividad de la intersección queda  $A \cup B = (A \cap B) \cup (A \cap B^C) \cup (B \cap A^C)$ . Al sacar la probabilidad  $P(A \cup B) = P(A \cap B) + P(A \cap B^C) + P(B \cap A^C)$ , por definición de  $A$  queda,  $P(A \cup B) = P(A) + P(B \cap A^C)$  si sumamos y restamos  $P(B \cap A)$  entonces,  $P(A \cup B) = P(A) + P(B \cap A^C) + P(B \cap A) - P(B \cap A)$  y por la definición de  $B$  queda,  $P(A \cup B) = P(A) + P(B) - P(B \cap A)$ .

Hay dos casos particulares que vale la pena destacar.

- (i) Sea  $P(A) = P(B) = 1$ . Entonces  $P(A \cup B) \geq P(A) = 1 \Rightarrow P(A \cup B) = 1$ , despejando (8) nos queda que  $P(A \cap B) = P(A) + P(B) - P(A \cup B)$ ,  $P(A \cap B) = 1 + 1 - 1 = 1$ . Es decir, si  $P(A) = P(B) = 1 \Rightarrow P(A \cup B) = P(A \cap B) = 1$ .
- (ii) Sea  $P(A) = P(B) = 0$ . Entonces  $P(A \cap B) \leq P(A) = 0 \Rightarrow P(A \cap B) = 0$ , despejando en (8) nos queda que  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ,  $P(A \cup B) = 0 + 0 - 0 = 0$ . Es decir, si  $P(A) = P(B) = 0 \Rightarrow P(A \cup B) = P(A \cap B) = 0$ .

La forma más general del principio de inclusión y exclusión es la siguiente:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right)$$

$$(9) \quad \text{Desigualdad de Boole} - P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$$

$$(10) \quad \text{Desigualdad de Bonferroni} - \text{Sean } A_1, A_2, \dots, A_n \text{ eventos y } A = \bigcup_{i=1}^n A_i. \text{ Muestre que } 1_A \leq \sum_{i=1}^n 1_{A_i}, \text{ etc. y luego tome el valor esperado para concluir:}$$

$$\begin{aligned}
P(\cup_{i=1}^n A_i) &\leq \sum_{i=1}^n P(A_i) \\
P(\cup_{i=1}^n A_i) &\geq \sum_{i=1}^n P(A_i) - \sum_{j < i} P(A_i \cap A_j) \\
P(\cup_{i=1}^n A_i) &\leq \sum_{i=1}^n P(A_i) - \sum_{j < i} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k)
\end{aligned}$$

En general, si detenemos el principio de inclusión y exclusión después de un número par (impar) de sumas, obtenemos una cota inferior (superior).

$$(11) \quad P(\cup_{i=1}^n A_i) \geq \max_i P(A_i)$$

(12) Sea  $A_1 \subseteq A_2 \subseteq \dots \subseteq A_{n+1}$  definamos  $A_N = \cup_{n=1}^N A_n$  y  $B_n = A_n - \cup_{i=1}^{n-1} A_i$  donde  $B_i \cap B_j \neq \emptyset$  para  $i \neq j$  y  $A_1 = B_1$ . Entonces  $\cup_{n=1}^N A_n = \cup_{n=1}^N B_n$ , como la sucesión  $A_n$  es creciente queda  $B_k = A_k - A_{k-1}$  entonces:

$$\begin{aligned}
P(B_k) &= P(A_k - A_{k-1}) \\
&= P(A_k) - P(A_{k-1}) \\
P(A_k) &= P(A_k - A_{k-1} \cup A_{k-1}) \\
&= P(A_k) - P(A_{k-1}) + P(A_{k-1}) \\
P(\cup_{n=1}^N A_n) &= P(A_N) \rightarrow P(\lim_{n \rightarrow \infty} \cup_{n=1}^N A_n) = P(\lim_{N \rightarrow \infty} A_N) = \lim_{N \rightarrow \infty} P(A_N)
\end{aligned}$$

$$\begin{aligned}
P(\cup_{n=1}^N A_n) &= P(\cup_{n=1}^N B_n) \\
&= \sum_{i=1}^N P(B_n) \\
&= \sum_{i=1}^N P(A_n) - P(A_{n-1}) \\
&= P(A_N)
\end{aligned}$$

### 1.3.5 Conjuntos Independientes

Con respecto a la función de probabilidad  $P(A \cap B) = P(A)P(B)$  si  $A \perp B$ .

$$A \perp B \Leftrightarrow A^C \perp B \Leftrightarrow A \perp B^C \Leftrightarrow A^C \perp B^C$$

$\{A_{ij}\} \subseteq \mathcal{F}$  es independiente.

$$\{A_{1k}, \dots, A_{ik}\}$$

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

Teorema de la clase monótona. Está en la página 42 de Introducción a la teoría de la medida de Ileana Iribarren

### 1.3.6 Probabilidad total

$A_i \in \mathcal{F}$  y  $A_i \cap A_j = \emptyset$  si  $i \neq j$

$$\Omega = \uplus_{i=1}^n A_i$$

Sea  $A \in \mathcal{F}$

$$P(A) = \sum_{i=1}^n P(A \cap A_i)$$

### 1.3.7 Probabilidad condicional

Proporciones de una proporción. Cambio de sistema de referencia.

Sea  $A, B \in \mathcal{F}$ ,  $P(A | B) = \alpha$ , tal que,  $P(A \cap B) = \alpha P(B)$ . Si  $P(B) \neq 0$ ,  $\alpha = \frac{P(A \cap B)}{P(B)}$ . Si  $P(B) = 0$  entonces  $\alpha \in \mathbb{R}$

$$P(A) = \sum_{i=1}^n P(A | A_i) P(A_i)$$

$$\begin{aligned} P(A | B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{P(B|A)P(A)}{\sum_{i=1}^n P(B|A_i)P(A_i)} \end{aligned}$$

Prueba que la probabilidad condicional es una medida de probabilidad.

Si  $P(B) \neq 0$ ,

$$P(\Omega | B) = \frac{P(\Omega \cap B)}{P(B)} = 1$$

Para  $A \cap C = \emptyset$

$$P(A \cup C | B) = P(A | B) + P(C | B)$$

Como  $A \cap B \subseteq B$ ,

$$0 \leq \frac{P(A \cap B)}{P(B)} \leq 1$$

$$P(\cdot | B) : \mathcal{F} \rightarrow [0, 1]$$

### 1.3.8 .. otra propiedad?

Sea  $A_1 \subseteq A_2 \subseteq \dots \subseteq A_n$

$$A = \cup_{i=1}^n A_i = \lim A_i$$

$$\begin{aligned} A &= A_1 \cup (A_2 - A_1) \cup \dots \cup (A_{n+1} - A_n) \cup \dots \\ &= A_1 \cup \bigcup_{i=1}^{\infty} (A_{i+1} - A_i) \end{aligned}$$

$$\begin{aligned} P(A) &= P(A_1) + \sum_{i=1}^{\infty} \frac{P(A_{i+1} - A_i)}{P(A_{i+1} - A_i)} \\ &= P(A_1) + \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_{i+1} - A_i) \end{aligned}$$

$\mathcal{F}$  es completa si contiene todos los conjuntos de medida nula de  $P$ .

Buscar Medida exterior

Buscar Medida Producto

Desigualdad de Bonferroni

Para controlar la incertidumbre no hay fórmulas exactas.



## 1.4 Variable Aleatoria

## 1.5 Función de Distribución

# 2 Repaso de Teoría de la medida

## 2.1 Límite de conjuntos

Sea  $A_1, \dots, A_n \subseteq \Omega$ ,

$$\lim_{n \rightarrow \infty} \sup(A_n) = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

$$\lim_{n \rightarrow \infty} \inf(A_n) = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

La unión de una secuencia de conjuntos es el ínfimo y la intersección es el supremo.

$$B_k = \bigcup_{n=k}^{\infty} A_n \Rightarrow B_{k+1} \subseteq B_k, \forall k.$$

$$C_k = \bigcap_{n=k}^{\infty} A_n \Rightarrow C_k \subseteq C_{k+1}, \forall k.$$

Si son monótonas:

$$\lim_{n \rightarrow \infty} \sup(A_n) = \lim_{n \rightarrow \infty} \inf(A_n)$$

Definición  $\{a_n\}$ :

Sucesión decreciente  $\downarrow$

$$\lim_{n \rightarrow \infty} \sup(a_n) = \inf_{k \geq 1} \sup_{n \geq k} (a_n)$$

Sucesión creciente  $\uparrow$

$$\lim_{n \rightarrow \infty} \inf(a_n) = \sup_{k \geq 1} \inf_{n \geq k} (a_n)$$

$$\lim_{n \rightarrow \infty} \inf(a_n) \leq \lim_{n \rightarrow \infty} \sup(a_n)$$

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sup(a_n) = \lim_{n \rightarrow \infty} \inf(a_n)$$

Ejercicio:

$$\lim_{n \rightarrow \infty} \inf(a_n) \subseteq \lim_{n \rightarrow \infty} \sup(a_n)$$

## 2.2 Lema de Fatou

$$P(\liminf_{n \rightarrow \infty} (a_n)) \leq \liminf_{n \rightarrow \infty} (P(a_n))$$

$$P(\liminf_{n \rightarrow \infty} (A_n)) \leq \liminf_{n \rightarrow \infty} (P(A_n)) \leq \limsup_{n \rightarrow \infty} (P(A_n)) \leq P(\limsup_{n \rightarrow \infty} (A_n))$$

Consecuencia del Lema de Fatou es la continuidad de la medida de probabilidad.

$$A_n \uparrow A \Rightarrow P(A_n) \uparrow P(A)$$

$$A_n \downarrow \emptyset \Rightarrow P(A_n) \downarrow 0$$

## 2.3 Medida de Lebesgue

$$\ell([a, b]) = b - a > 0$$

$\sigma$ -álgebra de Borel.  $\Omega = [0, 1]$ ,  $\mathcal{F} = \mathcal{B}(0, 1)$  y una función de medida  $\lambda$ .

Si  $A$  es no medible Lebesgue entonces:

$$\lambda(A) = \frac{1}{n} \sum_{i=1}^n 1_A(x_i) = ?$$

$$A = (a, b) \subset [0, 1]$$



$$P(A)_{(x_1, \dots, x_n)} \sim \lambda(A)$$

$$\Omega = \{c, s\} \cup \{canto\}, \mathcal{F} = \mathcal{P}(\Omega), P(\{c\}) = P(\{s\}) = \frac{1}{2}, P(\{canto\}) = 0$$

Cuando se modela se corta en algún lugar. Hay que tener cuidado sobre lo que se deja por fuera.  
En el modelo usual de la moneda, por ejemplo, se deja por fuera la caída de una moneda de canto.

## 2.4 Desigualdades de Cola

Nos preparamos para dominar la incertidumbre mediante las desigualdades de cola y controlar el error.

$$\begin{aligned} P(|f(x) - y| > \epsilon) &< \delta \\ P(X > \epsilon) &< \delta \\ P(X \leq \epsilon) &\geq 1 - \delta \end{aligned}$$

Uno se llena de desigualdades en teoría de probabilidades. Lo que queremos es controlar la incertidumbre, por lo que no hay respuestas exactas.

Hilbert - Teoría del Error Probabilístico - Consecuencia de Teoremas Límites

## 2.5 $\sigma$ -álgebra generada por una clase

$(\Omega, \mathcal{F})$

La completitud de la sigma álgebra Sea  $\mathcal{C} \subseteq \mathcal{P}(\Omega)$

$\sigma(\mathcal{C}) = \{\cap \mathcal{F} : \mathcal{F} \equiv \sigma\text{-álgebra y } \mathcal{C} \subseteq \mathcal{F}\}$ . (Axioma de elección escondido)

Borelianos + Conjuntos de medida nula.

Si  $P : \mathcal{F} \rightarrow [0, 1]$  es una medida de probabilidad, Donde  $\hat{\mathcal{F}}$  es la  $\sigma$ -álgebra generada por la unión de  $\mathcal{F}$  y los conjuntos de medida nula  $N$ , y se define de la manera siguiente  $\hat{\mathcal{F}} = \sigma(\mathcal{F} \cup N)$  y  $N$  son los conjuntos de medida nula donde  $N = \{A \subseteq \Omega : P(A) = 0\}$ .

$\mathcal{B}([0, 1]) = \sigma\{(a, b) : a < b, a, b \in \mathbb{R}\}$  la menor  $\sigma$ -álgebra que contiene los intervalos abiertos.

$\mathcal{P} = \lambda \equiv$  medida de Lebesgue,  $\lambda[a, b] = b - a$  y  $\mathcal{M}$  la  $\sigma$ -álgebra de Lebesgue, tal que  $\mathcal{M} = \sigma(\mathcal{B}([0, 1]) \cup \{N(\lambda)\})$ .

## 2.6 Integral Lebesgue-Stieltjes

## 3 Bibliografía

- Durrett
- Dudley - Real Analysis
- Feller I y II
- Billingsley
- Wet the odds
- Teoría de la medida
- Ulam, Feynman - Proyecto Manhattan
- Kac, Probability, Number Theory
- Gnedenko, Elementary theory of probability